DOCUMENT RESUME

ED 124 582 TM 005 338

AUTHOR Napier, John D.
TITLE The Ability of Elementary School Teachers to Assess Moral Thought.
PUB DATE [Apr 76]
NOTE 34p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Elementary School Students; *Elementary School Teachers; Evaluation; Evaluative Thinking; *Individual Characteristics; Manuals; *Moral Development; Moral Values; Predictor Variables; Statistical Analysis; *Student Evaluation; Teacher Background; Test Reliability; *Tests; Thought Processes
IDENTIFIERS Kohlberg Moral Development Test; Moral Knowledge Test; *Porter and Taylor Global Rater Manual

ABSTRACT

A published rater manual used by sixty elementary school teachers to assess moral thought statements was found ineffective. The study also examined what factors were related to their assessing ability. The subjects took a pretest using only descriptions of the six moral stages originally researched by Lawrence Kohlberg. Next the teachers were given the self training rater manual and a posttest. Correlations of the demographic and personality variables studied with scores on the pre and posttests showed no significant correlations with the posttest and only one significant (but moderate) correlation with the pretest scores. Analysis of variance tests on the pre and posttests showed Stage 1 and 3 scores significantly different but showed no overall difference. (Author/RC)

The Ability of Elementary School Teachers to Assess Moral Thought

John D. Napier

University of Georgia

A Paper Presented to the

Special Interest Group/Research in Social Studies Education

at the

Annual Meeting of the American Educational Research Association

San Francisco

1976

Abstract

The study examined whether sixty elementary school teachers could assess moral thought statements using a published rater manual and what factors were related to their assessing ability. The subjects took a pretest using only descriptions of the six moral stages originally researched by Lawrence Kohlberg. Next the teachers were given the self training rater manual and a posttest. Correlations of the demographic and personality variables studied with scores on the pre and posttests showed no significant correlations with the posttest and only one significant (but moderate) correlation with the pretest scores. Analysis of variance tests on the pre and posttests showed Stage 1 and 3 scores significantly different but showed no overall difference. The major conclusion was that the rater manual used was ineffective.

The Ability of Elementary School Teachers to Assess Moral Thought

The work of Lawrence Kohlberg and his associates has provided a description of moral thought development (Figure 1) as well as instructional procedures for developing moral thought (Kohlberg & Turiel, 1971). This construct of moral thought development and program for moral education has been suggested as a medium for value education in social studies curriculum (e.g., Fraenkel, 1973; Ferguson & Friesen, 1974; Joyce, 1972; Kohlberg, 1973). One of the instructional techniques used in most programs employing Kohlberg's ideas was assessment of moral thought prior to or during instruction (Rest, 1974). The assessment of moral thought is a complex task. If teachers cannot assess the stages of moral thought before instruction then the successful assessment of stages of moral thought during instruction is doubtful. Further, if teachers cannot assess the stages of moral thought then the successful use of the approaches advocated by Kohlberg and his associates for moral education is also doubtful.

Porter and Taylor (1972) have written a global rater manual as a self training aid for assessing the stages of moral thought statements. The purpose of this study

was to investigate whether teachers could use the Porter and Taylor global rater manual to correctly assess the stage of different moral thought statements, and what factors were related to their assessing abilities.

Stage Scoring Moral Thought Statements

The measurement system used to determine the stage of moral thought involved the qualitative coding of responses to hypothetical moral dilemmas. Two procedures have been used by Kohlberg and his associates to stage score moral thought statements. The first procedure was a sentence coding process. The second procedure was a global rating process. In both, a rater guide was used to determine a profile for a subject consisting of a stage or a combination of stages. Besides a stage profile, each subject was given a Moral Maturity Score. The Moral Maturity Score was the sum of the product of the percentage of responses at each stage in a given profile multiplied by the ordinal value of the stage times 100. The range of the Moral Maturity Score was from 100 (all Stage 1 responses) to 600 (all Stage 6 responses). The second procedure for stage scoring appeared the easiest to use and did not result in subject profiles which differed markedly from those profiles produced when the sentence coding process was used (Blatt & Kohlberg, 1973).

The use of rating guides aided in the classification of qualitative data, and the skill used in classifying qualitative data (like the data produced by the responses to hypothetical moral dilemmas) was content analysis. Although not much has been reported about variables influencing the ability to use content analysis, inferences from the abilities required and variables found important in similar research techniques yielded several variables of probable importance (Hyman, 1970; Richards, Dohrenwend & Klein, 1965). First, intelligence seemed important, especially verbal intelligence. Second, the demographic variables of sex, age, social class, educational background, and experience could be important. Finally, the ability to relate to others, or in other words, the ability to be receptive of others, also seemed important.

The research of Rest, Kohlberg, and Turiel (1969) suggested one other variable of probable importance. Their research showed that the moral stage of an individual influenced comprehension of moral thoughts of others. Therefore, the moral stage might correlate to assessing the moral thoughts of others.

In summary, the Porter and Taylor self training global rater manual was examined to determine if it aided teachers in assessing moral thought statements.

This manual was the only published guide available to aid teachers, and it represented the easiest form of stage scoring moral thought statements. Second, the variables of sex, age, social class, educational background, experience, verbal intelligence, receptive ability, and moral stage were suggested as possible correlates to assessing ability.

Methods

Subjects

The subjects of the study were sixty teachers involved in an experimental Master of Education program at a middle Atlantic state university. Although the teachers were volunteers for the experimental program, they did not volunteer for this particular study. The majority (83%) of the teachers taught in elementary schools while the remainder taught in junior high schools. Most of the teachers (63%) taught all subjects with the others equally distributed among various disciplines. All the teachers were engaged in an elementary social studies education program during the time of this study. Other demographic and personality characteristics of the teachers are given in Table 1.

Definitions

Demographic variables. The demographic variables

investigated were sex, age, and experience. Experience was defined as teaching experience. The variables of social class and educational background were of less importance because the teachers were all college graduates and all were considered on similar socioeconomic levels. These latter two variables were eliminated from investigation since differentiation between teachers was difficult.

Personality variables. The personality variables studied were the ability to relate to others (receptive ability), verbal intelligence, and moral maturity (moral stage). The ability to relate to others was defined as being cognitively open. Milton Rokeach's definition was accepted as an adequate description of the meaning of verbal receptive ability (Rokeach, 1960, p. 57):

> ". . . the extent to which the person can receive, evaluate, and act on relevant information received from the outside on its own intrinsic merits, unencumbered by irrelevant factors in the situation arising from within the person or from the outside."

Verbal intelligence was defined as the verbal score from the Graduate Records Examination. Moral maturity was defined as the Moral Maturity Score obtained from the global rating process from assessing moral thought statements.

Description of moral stages. The description of moral stages referred to the stage and sequence chart developed by Kohlberg and his associates (Kohlberg & Turiel, 1971, pp. 415-6), and was similar to that presented in Figure 1.

Self training global rater manual. The rater manual referred to an adapted version of one of the stories of the guide written by Porter and Taylor (1972, pp. 11-22).

Measurement Instruments

Moral maturity measure. The Kohlberg measure for determining the moral stages of individuals was constructed from four of the nine hypothetical dilemmas and interview questions found in his dissertation (Kohlberg, 1958, pp. 361-75). Stories II, V, VI, and IX were rewritten into an "adult" form, updated, and made into a paper and pencil openended questionnaire. These particular stories were chosen because the situations they presented were unfamiliar to the teachers. The validity questions of this measure was tentatively answered by Kohlberg on the basis of construct validity. The reliability of the nine dilemmas had been reported ranging from .76 to .31 among the situations with the median value being .51 (Kohlberg, 1969, pp. 369-97).

The responses to the measure were stage scored with the aid of the global rater guide found in the Appendix of Kohlberg's dissertation (Kohlberg, 1958, pp. 376-428). Each stage profile was converted into a Moral Maturity Score (see earlier discussion). Interrater reliability coefficients using Pearson product moment correlations were obtained on a random sample (30%) of the responses of the teachers. Three different forms of reliability were obtained. First, a reliability of .91 was found on assigning major stages to the teachers. Second, a reliability of .76 was found on assigning major and minor stages to the teachers. Third, a reliability of .70 was found on assigning Moral Maturity Scores to the teachers.

Receptive ability measure. The Rokeach dogmatism test form E was used. The measure had construct validity and a test-retest reliability ranging from .93 to .68. The original scoring used by Rokeach had a scale of +3, +2, +1, -1, -2, -3 in which +3 represented a totally close-minded response and -3 represented a totally open-minded response. In this study the scale was modified to +1 through +6 in which +1 represented a totally close-minded response and +6 represented a totally open-minded response. This change reversed the scale so that the higher scores indicated open-mindedness and the lower scores

indicated close-mindedness (Rokeach, 1960, pp. 71-97).

Verbal intelligence measure. The Graduate Records Examination test of verbal ability had both reliability and validity as inferred from the information from the total testing program. The program developers found that all tests in the program had reliability coefficients greater than .90 on the Kuder-Richardson Formula 20. The basis for the validity of the various test was their predictive validity for groups (Buros, 1972, pp. 1030-2).

Assessing ability measure. The testing instrument used to measure the ability to assess moral thought statements (termed Moral Knowledge Test) was constructed from typical stage responses to story III of Kohlberg's hypothetical moral dilemmas (Kohlberg, 1958, pp. 363-4). There were 24 items with each of the six stages represented by 4 items. In each stage, 2 of the 4 items represented "Orientation to Intentions and Consequences," and 2 of the 4 items represented "Motives for Engaging in Moral Action." "Orientations to Intentions and Consequences" and "Motives for Engaging in Moral Action" were the concepts of moralization that were used in the test. For each concept at each stage, 1 of the 2 items had a do choice and the other had a don't choice. These 24 items had an equal distribution of stage statements, an equal

distribution of do and don't choices, and a sample of the concepts used in moralization (see Note 1). These items were adapted from the work of James Rest. This test had content validity based on the sampling of moral statements representing the types of stages, choices, and concepts used in moral thought statements. A form of "face" validity was, also, assumed because both Rest and Kohlberg agreed that these statements represented one of the six stages, one of the choices, and one of two concepts of moral thought described above (Kohlberg, 1969, pp. 370-82). The reliability of the measure was determined by using the Cronbach alpha projected to a standard test (100 items). A reliability coefficient of .62 was found for the scores on the pretest, and a coefficient of .48 was found for the scores on the posttest. Although these reliability coefficients were low, they were adequately reliable for a group size of sixty teachers (Thorndike & Hagen, 1969, pp. 194-5).

The test produced two types of scores for each teacher. First, the test produced an overall correct score which represented the number of correctly assigned stage statements. This correct score ranged from 0 to 24. Second, the test yielded correct scores for each stage level ranging from 0 to 4. Thus an analysis of overall and individual stage

performance was possible.

Research Design and Procedures

Data collection. The data of age, sex, teaching experience, and verbal score on the Graduate Records Examination were obtained from the information folders for the experimental Master of Education program. The data on Moral Maturity Score, Rokeach dogmatism score, and correct scores on Moral Knowledge Test I (pretest) and Moral Knowledge Test II (posttest) were obtained over a period of three weekend sessions of the program. At the first session, the Rokeach measure was given, and the Kohlberg measure for determining moral maturity was handed out to be returned at the second session. At the second session, the teachers were given Moral Knowledge Test I (pretest) using just the stage descriptions materials provided. The description was given to the teachers in order to preclude their having to memorize the descriptions of Kohlberg's stages. Prior to administering Moral Knowledge Test I, the teachers were given information about Kohlberg's approach to moral development and moral education. Specifically, the teachers were shown a filmstrip (Kohlberg & Selman, 1972), given two articles (Kohlberg & Mayer, 1972; Kohlberg & Whitten, 1972), and directed in discussions during two weekend sessions. Finally, most of the

teachers (88%) had tried to intuitively assess the moral thoughts of their students prior to the pretest, but had expressed difficulties about their attempts. After the pretest was returned, the teachers were instructed to take the same test home and use both the stage descriptions and the Porter and Taylor global rater manual provided and to assess the stages of the same 24 moral thought items (Moral Knowledge Test II). Two weeks later, the teachers returned Moral Knowledge Test II at the third class session[2].

Two observations about the data collection method were made. First, according to Kohlberg, the responses to hypothetical moral dilemmas and questions uncovered the underlying structure of moral thought and unless an individual understood this way of thinking, it could not be used properly (Kohlberg, 1971). So, taking the Kohlberg measure under an unsupervised condition was assumed to have no effect on the resulting Moral Maturity Scores for the teachers. Second, the testing influence from the Moral Knowledge Test I to the Moral Knowledge Test II was not deemed a contaminating factor. If the teachers had been given only the Moral Knowledge Test II, the teachers would have been able to do the exercise as many times as they wished before returning it because they were given two weeks to complete the exercise. Thus a

prior exercise (Moral Knowledge Test I) was assumed to have no real influence on the results of Moral Knowledge Test II.

As with any social science research, the problem of missing data occurred. Despite all efforts, a complete data bank for all teachers was not obtained. Review of the effects of the missing data resulted in the conclusion that the missing data represented random loss.

Statistical analyses. The first question studied was

1. Is there a significant linear correlation between the selected demographic and personality variables of the teachers and their ability to assess moral thought statements and different stages of moral thought statements?

To test this question, Pearson product moment correlations on age, teaching experience, verbal intelligence, Rokeach score, and Moral Maturity Score with overall and stage correct scores for Moral Knowledge Test I and II were made. Since sex was a true dichotomy, point biserial correlations were used to test the relationship with correct scores for Moral Knowledge Test I and II. Procedures used were found in Bruning and Kintz (1968).

The second and third questions studied were

2. Is there a significant difference between Moral Knowledge Test I and II in the ability of the teachers to assess moral thought statements and different stages of moral thought statements?
3. Is there a significant difference within Moral Knowledge Test I and II in the ability of the teachers to assess different stages of moral thought statements?

To test these questions, a three factorial analysis of variance test with n observations per cell, an analysis of simple main effects test and a Newman-Kuels Sequential Range test were used. The factors of Test and Stage were assumed fixed effects while the factor of Subject was assumed a random effect. The n observations per cell were the 4 items at each stage found in the Moral Knowledge Test. Procedures used were found in Winer (1971).

All tests of the null hypotheses associated with each question used a two tail test at the p = .05 level of significance.

Results

Descriptive Statistics

The descriptive statistics for the dependent and independent variables are presented in Table 1. The demographic variables indicated that the teachers were, on the average, female, around 33 years of age, and had 8 years of teaching experience. The personality variables indicated that the teachers were more than one standard deviation below the norm mean of 500 on the Graduate Records Examination for verbal ability. Although no norms were available for the Rokeach dogmatism test, the design of the measure provided for an open-minded score to fall within the range of 140 to 240. Therefore, as a group these teachers were open-minded. The Moral Maturity Score also had no norm reference. But, in conjunction with the Moral Stage, it was noted that as a group these teachers were thinking on the principle level of moral thought in many instances. The score for Moral Knowledge Test I showed that the teachers correctly identified only one-third of the 24 moral statements. The results for scores of Moral Knowledge Test II were almost exactly the same as for Moral Knowledge Test I.

The frequency distribution of stage scores assigned by the sixty teachers to each moral thought statement in

Moral Knowledge Test I and II are given in Table 2. The table was constructed to reorder the 24 test items in order to group by and identify according to stage, choice and concept. The frequency distribution for Moral Knowledge I demonstrated that these teachers did not correctly assign stage scores to each of the 4 items within each stage equally. Also, there was a difference in correct stage score assignment between stages. In addition, when teachers incorrectly stage scored different stage level moral thought statements, they assigned stage scores to any of the other stage levels. The same general observations were made by inspecting the frequency distribution for Moral Knowledge Test II. Further discusssion of these results is made in the next section.

Statistical Tests

First question. Pearson product moment correlations and point biserial correlations of demographic and personality variables to the overall scores of Moral Knowledge Test I and II produced only one significant result. Verbal ability ($r = .43$) significantly correlated with the overall score for Moral Knowledge Test I. No significant correlations were found with overall scores for Moral Knowledge Test II.

Pearson product moment correlations were also done

between the demographic and personality variables and scores for each stage within each Moral Knowledge Test. The results tended to conform to the ones found in the correlations with overall scores for the Moral Knowledge tests. For Moral Knowledge Test I, verbal ability significantly correlated with Stage 2 ($r = .29$), Stage 3 ($r = .31$), and Stage 6 ($r = .35$). Again, there were no significant correlations with verbal ability and scores for Moral Knowledge Test II. Four other significant correlations were uncovered. For Moral Knowledge Test I, the Moral Maturity Score significantly correlated with Stage 2 ($r = .26$). For Moral Knowledge Test II, teaching experience significantly correlated with Stage 6 ($r = .34$), and the Rokeach dogmatism score significantly correlated with Stage 1 ($r = .26$) and Stage 4 ($r = .28$). Point biserial correlations between sex and scores for each stage were not done because of the insignificant correlations with overall scores for Moral Knowledge Test I and II.

In summary, the correlation tests demonstrated that verbal ability was the only significant correlation with performance on Moral Knowledge Test I, and no significant correlations were found for performance on Moral Knowledge Test II. This finding for overall performance was

supported by the findings for correlations with stage performance in Moral Knowledge Test I and II. However, it must be noted that none of the significant correlating variables had more than a moderate level of correlation with the overall and stage scores for Moral Knowledge Test I and II.

Second and third questions. The results of the analysis of variance test is presented in Table 3. The analysis between the main effects of Test showed no significant difference between scores for Moral Knowledge Test I and II.

Because the second question required a test for differences between Test at each Stage level, and since the interaction of Test x Stage was significant (Table 3), a test of simple main effects was performed. Table 4 presents the results of these tests. The analysis of Test at each Stage level showed that Stages 1 and 3 were significantly different between Moral Knowledge Test I and II. Referring to Table 2, the teachers improved on Stage 1 items but declined in correctly assigning stage scores for Stage 3 items.

Since the third question asked for the determination of difference between Stages at each Test level and the interaction between Test x Stage was significant (Table 3),

a test of simple main effects for Stages at Test levels was performed (Table 4). For both Test levels there was a significant difference between Stages. A Newman-Keuls test was performed on the total correct stage scores for both Moral Knowledge Test I and II (Table 5). The results for Moral Knowledge Test I showed (a) Stage 3 was significantly different from Stage 5 and 4; (b) Stage 1, 2, and 6 were significantly different from Stage 5. No other significant differences were noted. The results for Moral Knowledge Test II showed (a) Stage 1 was significantly different from Stage 2, 3, 4, 5 and 6; (b) Stage 6 and 3 were significantly different from Stage 4.

In summary, the overall test for difference between Moral Knowledge Test I and II showed that the performance on the two tests were not different. Inspection of performance on stages between and within each Moral Knowledge Test indicated that there were some differences. These latter findings support the conclusion that the Porter and Taylor global rater guide did have some effect on the performances between Moral Knowledge I and II, but did not aid the teachers in changing overall performance.

Discussion

This study supported the generalization that teachers cannot assess moral thought statements with an adequate

degree of correctness by using the Porter and Taylor global rater manual. However, information concerning factors related to the ability to assess moral thought statements was meager. None of the demographic and personality variables examined were strongly related to assessing abilities in either Moral Knowledge Test I or II. These variables were studied as being related linearly to the ability to assess moral thought statements. It could be that they do not relate in a linear fashion, or there are other more strongly correlated factors that were not investigated.

A rater manual is supposed to equalize raters so that differences in demographic and personality variables are controlled. Although the variable of verbal ability, which significantly correlated with Moral Knowledge Test I, was removed with Moral Knowledge Test II, the teachers did not improve in their assessing abilities. There was something influencing the teachers in such a way as to cause them to do poorly in assessing moral thought statements. Tests on differences in stage preformance between and within tests indicated that the Porter and Taylor global rater manual had an effect. However, the effect was merely to help rate Stage 1 moral statements better and Stage 3 statements worse between the two tests. At present, the exact reason

for this change is not evident. What is evident is that the Porter and Taylor manual was inadequate in helping these teachers successfully use the global rating process.

As a result, research needs to be done on (a) ways to assist teachers to properly assess moral thought statements (i.e., a better rater manual and training); (b) a determination of factors relating to assessing ability. Until these research problems are solved, the use of the advocated instructional procedures requiring diagnosis of moral thoughts of students cannot be used by teachers like the ones in this study. It is imperative, therefore, that social studies educators refrain from indicating that teachers can correctly assess the moral thought stage of their students. For if teachers tried to assess the moral thought stage of their students, they could incorrectly stage score Stage 1 moral thought as Stage 6 or Stage 6 as Stage 1!

References

Blatt, M. M. & Kohlberg, L. The effects of classroom moral discussion upon children's level of moral judgment. In Collected papers on moral development and moral education. Cambridge: Center for Moral Education, Harvard University, 1973.

Bruning, J. L. & Kintz, B. L. Computational handbook of statistics. Glenview, Ill.: Scott, Foresman, 1968.

Buros, O. K. (Ed.). The seventh mental measurement yearbook. Rohway, N. J.: Quinn & Boden, 1972.

Hyman, H. H. Interviewing in social research. Chicago: University of Chicago Press, 1970.

Fraenkel, J. R. Helping students think and value. Englewood Cliffs, N. J.: Prentice-Hall, 1973.

Ferguson, P. & Friesen, J. W. Values theory and teaching: The problem of autonomy versus determinism. Theory and Research in Social Education. 1974, 1, 1-24.

Joyce, B. R. New strategies for social education. Chicago: Science Research Asspciates, 1972.

Kohlberg, L. The development of modes of moral thinking and choice in the years ten to sixteen. Unpublished Ph.D. dissertation, University of Chicago, 1958.

Kohlberg, L. Stages and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally, 1969.

Kohlberg, L. From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), Cognitive development and epistemology. New York: Academic Press, 1971.

Kohlberg, L. Moral development and the new social studies. Social Education. 1973, 37, 369-375.

Kohlberg, L. & Mayer, R. Development as the aim of education. Harvard Educational Review. 1972, 42, 449-496.

Kohlberg, L. & Selman, R. First things: A strategy for teaching values. Guidance Associates Filmstrip, 1972.

Kohlberg, L. & Turiel, E. Moral development and moral education. In G. Lesser (Ed.), Psychology and educational practice. Glenview, Ill.: Scott, Foresman, 1971.

Kohlberg, L. & Whitten, P. Understanding the hidden curriculum. Learning. 1972, 1, 10-14.

Porter, N. & Taylor, N. How to assess the moral reasoning of students. Toronto: Ontario Institute for Studies in Education, 1972.

Rest, J. Developmental psychology as a guide to value education: A review of "Kohlbergian" programs. Review of Educational Research. 1974, 44, 241-259.

Rest, J., Kohlberg, L. & Turiel, E. Level of moral development as a determinant of preference and comprehension of moral judgments made by others. Journal of Personality. 1969, 37, 225-252.

Richards, S. A., Dohrenwend, B. S. & Klein, D. Interviewing: Its forms and functions. New York: Basic Books, 1965.

Rokeach, M. The open and closed mind. New York: Basic Books, 1960.

Thorndike, R. L. & Hagen, E. Measurement and evaluation in psychology and education (2nd ed.). New York: John Wiley & Sons, 1969.

Turiel, E. Developmental processes in the child's moral thinking. In P. Mussen (Ed.), Trends and issues in developmental psychology. New York: Holt, Rinehart & Winston, 1969.

Winer, B. J. Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill, 1971.

Footnotes

[1]In order to demonstrate the items on the Moral Knowledge Test, the following 4 examples are given. The examples are similar to the ones used on the Moral Knowledge Test and represent the four basic contents and four of the six stages used on the test:

Stage 1, Choice do, Concept of "Orientation to Intentions and Consequences":

Heinz should steal the drug for his dying wife. It isn't actually bad to take the drug. It isn't like Heinz didn't ask the druggist if he could pay later. The drug that Heinz would take is only worth $200. So, he isn't really taking a $2000 drug.

Stage 2, Choice don't, Concept of "Orientation to Intentions and Consequences":

Heinz should not take the drug for his dying wife. The druggist isn't wrong or evil, he just wants to sell the drug for a profit. That is what a person is in business to do, to make a profit.

Stage 3, Choice do, Concept of "Motives for Engaging in Moral Action":

No one will think Heinz is bad if he steals the drug for his dying wife, but his family will think he is an inhuman husband if he doesn't steal the drug. If Heinz lets his wife die, he will never be able to look anybody in the face again.

Stage 4, Choice don't, Concept of "Motive for Engaging in Moral Actions":

Heinz is desperate and he may not know he is wrong when he steals the drug for his dying wife. But he will know he did wrong after he is punished and sent to prison. He will always feel guilty for his dishonesty and lawbreaking.

For a more detailed description of the 24 items, the reader is referred to the discussion by Rest and Kohlberg (Kohlberg, 1969, pp. 370-82).

[2]No formal investigation was conducted as to whether the teachers used the Porter and Taylor self training global rater manual for Moral Knowledge Test II. Because of the simplicity of the adapted version of the Porter and Taylor manual, it was believed logical to assume that the teachers would use the manual. Informal discussions between course instructors and teachers after the experiment convinced this researcher that the logical assumption that teachers would use the manual was correct.

FIGURE 1

STAGES OF MORAL DEVELOPMENT*

PRE-CONVENTIONAL LEVEL

VALUE RESIDES IN EXTERNAL QUASI-PHYSICAL HAPPENINGS, IN PERSONS AND STANDARDS.

STAGE 1: PUNISHMENT AND OBEDIENCE ORIENTATION. EGOCENTRIC DEFERENCE TO SUPERIOR POWER OR PRESTIGE OR A TROUBLE-AVOIDING SET. OBJECTIVE RESPONSIBILITY.

STAGE 2: INSTRUMENTAL RELATIVIST ORIENTATION. RIGHT ACTION IS THAT OF INSTRUMENTALLY SATISFYING THE SELF'S NEEDS AND OCCASIONALLY THE NEEDS OF OTHERS. AWARENESS OF RELATIVISM OF VALUE TO EACH ACTOR'S NEEDS AND PERSPECTIVE. NAIVE EGALITARIANISM AND ORIENTATION TO EXCHANGE AND RECIPROCITY.

CONVENTIONAL LEVEL

MORAL VALUE RESIDES IN PERFORMING GOOD OR RIGHT ROLES, IN MAINTAINING THE CONVENTIONAL ORDER AND THE EXPECTANCIES OF OTHERS.

STAGE 3: "GOOD BOY--NICE GIRL" ORIENTATION. ORIENTATION TO APPROVAL AND TO PLEASING AND HELPING OTHERS. CONFORMITY TO STEREOTYPICAL IMAGES OF MAJORITY OR NATURAL ROLE BEHAVIOR, AND JUDGMENT BY INTENTIONS.

STAGE 4: LAW AND ORDER ORIENTATION. ORIENTATION TO "DOING DUTY" AND TO SHOWING RESPECT FOR AUTHORITY AND MAINTAINING THE GIVEN SOCIAL ORDER FOR ITS OWN SAKE. REGARD FOR EARNED EXPECTATIONS OF OTHERS.

POST-CONVENTIONAL LEVEL

MORAL VALUE RESIDES IN CONFORMITY BY THE SELF TO SHARED OR SHAREABLE STANDARDS, RIGHTS, OR DUTIES.

STAGE 5: SOCIAL-CONTRACT LEGALISTIC ORIENTATION. RECOGNITION OF AN ARBITRARY ELEMENT OR STARTING POINT IN RULES OR EXPECTATIONS FOR THE SAKE OF AGREEMENT. DUTY DEFINED IN TERMS OF CONTRACT, GENERAL AVOIDANCE OF VIOLATION OF THE WILL OR RIGHTS OF OTHERS, AND MAJORITY WILL AND WELFARE.

STAGE 6: UNIVERSAL ETHICAL PRINCIPLE ORIENTATION. ORIENTATION NOT ONLY TO ACTUALLY ORDAINED SOCIAL RULES BUT ALSO TO PRINCIPLES OF CHOICE INVOLVING APPEAL TO LOGICAL UNIVERSALITY AND CONSISTENCY. ORIENTATION TO CONSCIENCE AS A DIRECTING AGENT AND TO MUTUAL RESPECT AND TRUST.

*ADAPTED FROM TURIEL (1969, PP. 96-97)

Table 1

Descriptive Statistics for Variables under Study*

	AGE	TEACHING EXPERIENCE	GRE VERBAL	ROKEACH SCORE	MORAL MATURITY SCORE	STAGE	MORAL KNOWLEDGE I SCORE	MORAL KNOWLEDGE II SCORE
MEAN	32.9	8.2	394.3	153.4	424.5		7.97	7.85
MEDIAN	31.5	5.3	378.3	155.5	433.3		7.75	7.50
MODE	26.0	2.0	400.0	156.0	483.0	5(4)	8.00	7.00
STANDARD DEVIATION	7.5	7.4	91.0	22.9	55.6		2.54	2.25
RANGE	22.0 to 57.0	1.0 to 38.0	220.0 to 620.0	108.0 to 213.0	267.0 to 500.0	2(3) to 5(6)	3.00 to 14.00	3.00 to 13.00
NUMBER MISSING	0	1	6	2	3	3	0	0

*There were 9 Males and 51 Females in the Study

Table 2

Frequency Distribution of Stage Scores Assigned*

			MORAL KNOWLEDGE I STAGE ASSIGNED							MORAL KNOWLEDGE II STAGE ASSIGNED						
STAGE	CHOICE	CONCEPT	1	2	3	4	5	6	TOTAL	1	2	3	4	5	6	TOTAL
1	DO	INTENTION	3	25	12	3	13	4	60	1	31	17	0	9	2	60
1	DON'T	INTENTION	16	6	9	25	1	3	60	30	0	0	28	2	0	60
1	DO	MOTIVE	23	5	12	16	2	2	60	24	12	9	5	8	2	60
1	DON'T	MOTIVE	44	9	1	5	0	1	60	57	1	0	2	0	0	60
STAGE 1 TOTALS			86	45	34	49	16	10	240	112	44	26	35	19	4	240
2	DO	INTENTION	0	9	9	10	26	6	60	0	8	25	3	18	6	60
2	DON'T	INTENTION	13	24	8	9	5	1	60	16	14	1	27	1	1	60
2	DO	MOTIVE	9	25	13	9	4	0	60	7	28	14	1	7	3	60
2	DON'T	MOTIVE	9	28	9	4	4	6	60	11	27	14	3	5	0	60
STAGE 2 TOTALS			31	86	39	32	39	13	240	34	77	54	34	31	10	240
3	DO	INTENTION	3	19	24	4	8	2	60	1	23	29	2	3	2	60
3	DON'T	INTENTION	5	11	11	22	8	3	60	18	9	6	24	2	1	60
3	DO	MOTIVE	3	6	39	6	4	2	60	1	10	38	1	6	4	60
3	DON'T	MOTIVE	4	9	27	13	6	1	60	14	7	9	22	5	3	60
STAGE 3 TOTALS			15	45	101	45	26	8	240	34	49	82	49	16	10	240
4	DO	INTENTION	3	7	10	10	22	8	60	0	20	24	3	10	3	60
4	DON'T	INTENTION	5	11	12	28	1	3	60	2	1	24	30	3	0	60
4	DO	MOTIVE	4	14	18	11	8	5	60	0	12	24	9	8	7	60
4	DON'T	MOTIVE	17	6	11	18	4	4	60	27	5	8	13	5	2	60
STAGE 4 TOTALS			29	38	51	67	35	20	240	29	38	80	55	26	12	240
5	DO	INTENTION	3	5	5	6	29	12	60	0	3	7	3	30	17	60
5	DON'T	INTENTION	1	2	3	31	10	13	60	0	0	2	42	10	6	60
5	DO	MOTIVE	7	7	29	7	8	2	60	0	7	28	4	15	6	60
5	DON'T	MOTIVE	1	0	19	28	10	2	60	9	2	10	33	5	1	60
STAGE 5 TOTALS			12	14	56	72	57	29	240	9	12	47	82	60	30	240
6	DO	INTENTION	0	1	2	6	19	32	60	0	0	0	1	22	37	60
6	DON'T	INTENTION	1	5	9	9	17	19	60	1	4	2	8	23	22	60
6	DO	MOTIVE	5	5	5	9	20	16	60	6	6	7	6	19	16	60
6	DON'T	MOTIVE	2	5	13	9	17	14	60	10	5	13	14	8	10	60
STAGE 6 TOTALS			8	16	29	33	73	81	240	17	15	22	29	72	85	240

*The concept of "Intention" referred to "Orientations to Intentions and Consequences" and the concept of "Motive" referred to "Motive for Engaging in Moral Action."

Table 3

Analysis of Variance of Moral Knowledge Test I and II

Source	SS	df	MS	F
Subject	17.604	59	.2984	
A (Test)	.017	1	.0170	.094
A x Subject	10.629	59	.1802	
B (Stage)	11.022	5	2.2044	10.438*
B x Subject	62.290	295	.2112	
AB	2.665	5	.5330	3.125*
AB x Subject	50.314	295	.1706	
Within	481.750	2160	.2230	
Total	636.291	2879		

*Significant at p = .05

Table 4

Analysis of Simple Main Effects for Test-Stage Interaction

Source	SS	df	MS	F
A (Test) at B_1 (Stage 1)	1.408	1	1.408	7.814*
A at B_2 (Stage 2)	.169	1	.169	.938
A at B_3 (Stage 3)	.752	1	.752	4.173*
A at B_4 (Stage 4)	.300	1	.300	1.665
A at B_5 (Stage 5)	.019	1	.019	.105
A at B_6 (Stage 6)	.034	1	.034	.189
	(2.682)			
A x Subject	10.629	59	.1802	
B (Stage) at A_1 (Test 1)	5.047	5	1.009	4.777*
B at A_2 (Test 2)	8.640	5	1.728	8.182*
	(13.687)			
B x Subject	62.290	295	.2112	

*Significant at p = .05

Table 5

Newman-Keuls Sequential Range Test for Stages at Test Levels

Moral Knowledge I

T_j	57	67	81	86	86	101		
	B_5	B_4	B_6	B_2	B_1	B_3	r	$q_{95}(r,295)\sqrt{nMS_{bs}}$
B_5	—	10	24*	29*	29*	44*	6	29.12
B_4			14	19	19	34*	5	27.83
B_6				5	5	20	4	26.20
B_2					0	15	3	23.92
B_1						15	2	19.86

Moral Knowledge II

T_j	55	60	77	82	85	112		
	B_4	B_5	B_2	B_3	B_6	B_1	r	$q_{95}(r,295)\sqrt{nMS_{bs}}$
B_4	—	5	22	27*	30*	57*	6	29.12
B_5			17	22	25	52*	5	27.83
B_2				5	8	35*	4	26.20
B_3					3	30*	3	23.92
B_6						27*	2	19.86

*Significant at p = .05