

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND

7913060

STEGALL, MARY LIVINGSTON
THE EFFECT OF MORALITY CONTENT STORIES ON THE
YOUNG CHILD'S RESPONSES ON A MORAL
DEVELOPMENT INVENTORY.

THE UNIVERSITY OF NORTH CAROLINA AT
GREENSBORO, PH.D., 1978

University
Microfilms
International 300 N. ZEEB ROAD, ANN ARBOR, MI 48106

© 1979

MARY LIVINGSTON STEGALL

ALL RIGHTS RESERVED

THE EFFECT OF MORALITY CONTENT STORIES
ON THE YOUNG CHILD'S RESPONSES ON A
MORAL DEVELOPMENT INVENTORY

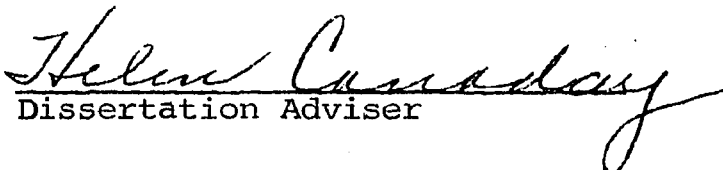
by

Mary Livingston Stegall

A Dissertation Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
1978

Approved by


Dissertation Adviser

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of the Graduate School at the University of North Carolina at Greensboro.

Dissertation
Adviser

Heleen Canaday

Committee Members

J. Allen Watson
Richard B. Johnson
Rosemary A. Nelson

August 10, 1978
Date of Acceptance by Committee

August 7, 1978
Date of Final Oral Examination

STEGALL, MARY LIVINGSTON. The Effect of Morality Content Stories on the Young Child's Responses on a Moral Development Inventory. (1978) Directed by: Dr. Helen Canady. Pp. 152.

The primary objective of the present study was to experimentally test the hypothesis that moral development could be fostered in first-grade children by reading them morality stories selected for their content similarity to the induction child-rearing techniques described by Hoffman and Saltzstein (1967). The induction technique involved instructing the child concerning the consequences of his misdeeds on others.

In order to assess the treatment effect of the morality stories, the researcher developed the Moral Development Inventory (M.D.I.) for measuring level of moral development in first-grade children. This instrument, modeled after a guilt inventory used in the Hoffman and Saltzstein (1967) study, consisted of four Inventory Stories. Each story contained a protagonist the age of the subjects to be tested. In each story, the protagonist committed a misdeed known only to himself. In the administration of the instrument, the subject was asked to complete each Inventory Story by choosing from three possible endings, each ending corresponding to either high, medium, or low level of guilt. A pilot study (N = 316 first graders) confirmed that the instrument was suitable for use with first-grade subjects. The M.D.I. was then administered to 73 first graders in Walton County, Florida. A numerical score (R = 4-12) was obtained for each subject.

Each subject's score was correlated with a respective morality rating from his teacher ($\underline{r} = .37, p < .001$), which resulted in the rejection of the null hypothesis that there would be no significant relationship between the subject's score on the M.D.I. and the teacher rating of the subject's morality. A positive relationship ($\underline{r} = .32, p < .01$) between the subject's M.D.I. score and his score on the language portion of the Comprehensive Tests of Basic Skills indicated that the M.D.I. contained a language factor. Split-half reliability ($\underline{r} = .63, p < .001$), test-retest reliability ($\underline{r} = .82$), and a non-significant F max statistic for homogeneity of variances justified the acceptance of the M.D.I. as a satisfactory research instrument for assessing the subjects in the experiment.

The subjects in the experiment were 27 first graders assigned to experimental conditions and 27 first graders assigned to control conditions. All subjects were from Walton County, Florida, schools. The experimental subjects were read one morality story each day for five consecutive days. The control subjects were read one nonmorality story in like manner. The subjects were administered the M.D.I. after the reading of the fifth treatment story. The experimental effects were assessed by a three-way ANOVA, which determined that there was a significant main effect for group, a significant effect for sex, and a significant effect for group by sex. The analysis indicated no significant differences in treatment effect on the individual Inventory

Stories, but there was a significant three-way (sex x group x story) interaction. Newman-Keuls post hoc comparisons on the mean scores revealed that the differences were primarily related to the very low scores of the control males and the lack of significant treatment effect on the females.

In summary, it was found that the treatment effect as measured by the M.D.I. was highly significant with the male subjects but not with the female subjects. Possible explanations were discussed. The research has definite implications for parents and educators. The induction technique for fostering moral development would seem to be an effective procedure for producing moral children, especially males.

ACKNOWLEDGMENTS

The author wishes to express her gratitude to her committee chairperson, Dr. Helen Canaday, for her time, guidance, and encouragement during the long course of this research investigation. Appreciation is also extended to Dr. Mildred Johnson, Dr. Rosemary Nelson, and Dr. J. Allen Watson for their helpful suggestions and constructive criticisms.

In addition, appreciation is expressed to the Walton County, Florida, School System for the tremendous amount of cooperation and kindness extended to me, and to my typist, Mrs. Elizabeth Hunt, whose gracious manner sustained me through many crises.

I dedicate this dissertation to the memory of my mother, Edith Pierce Livingston, whose untimely death prevented my sharing with her this culmination of the dream she held for me, and to my father, Frederick Elliott Livingston, whose love and support continues to sustain me in all achievements.

TABLE OF CONTENTS

	Page
APPROVAL PAGE	ii
ACKNOWLEDGMENTS	iii
LIST OF TABLES	
CHAPTER	
I. INTRODUCTION	1
Theoretical Orientation	4
Background for the Study	6
Statement of the Problem	9
Questions and Hypotheses	9
Implications of the Study	11
Limitations of the Study	12
Assumptions	15
Definitions of Terms	16
II. REVIEW OF LITERATURE	18
The Psychoanalytic Theory of Moral Development	21
The Cognitive-Developmental Theory of Moral Development	23
The Social Learning Theory of Moral Development	34
The Induction Method and Moral Internalization	39
The Phenomenon of Guilt	44
Summary	54
III. METHODOLOGY	56
Development of the Instrument	57
Procedures for Development of the Instrument	57
Pilot Study Related to Development of the Moral Development Inventory	58
Instrument Procedural Discussion	61
Administration of the Instrument	62
Scoring of the Guilt Inventory	65
Rating of Students' Overt Reaction to Transgression as Perceived by the Teacher	65
Analysis of the Instrument	66

TABLE OF CONTENTS (continued)

	Page
CHAPTER	
III. METHODOLOGY (continued)	
The Experiment	67
Procedures	68
Conducting the Experiment.	69
Analysis of the Experimental Data.	72
IV. RESULTS.	73
The Instrument	73
Pilot Study.	74
Analysis of the Instrument	75
Validity	77
Reliability.	80
Experiment	81
V. DISCUSSION	95
Pilot Study.	97
Validity of the Instrument	99
Reliability of the Instrument.	102
Experiment	104
VI. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS.	112
Summary	112
Conclusions.	117
Recommendations.	118
BIBLIOGRAPHY.	123
Appendix A Kohlberg's Definition of Moral Stages	135
Appendix B Story Completion Guilt Ratings as Adapted from Hoffman and Saltzstein (1967).	138
Appendix C Moral Inventory for First-Grade Children.	139
Appendix D Scoring Sheet for Moral Development Inventory	147
Appendix E Form for Teacher to Fill Out; One for Each Child--adapted from Hoffman and Saltzstein (1967)	148

TABLE OF CONTENTS (continued)

	Page
CHAPTER	
Appendix F Control Treatment Stories	150
Appendix G True-False Quiz for Control Treatment Stories	151

LIST OF TABLES

Table	Page
1	Summary Table of the Ss Total Scores on the M.D.I. 76
2	Number of Subjects Who Responded to First, Second, or Third Choice Order on Each of the Four Inventory Stories. 78
3	Chi-square Values for Order of Subjects' Responses to First, Second, and Third Choices on Each of Four Inventory Stories. 79
4	Number of Subjects by Classroom, Treatment Condition, Sex, and Removal Method. 82
5	Univariate Analysis of Variance on the Effects of Sex, Group Assignment, and Stories on Average Morality Score 84
6	Summary Table of Total M.D.I. Means (R = 4-12) and Inventory Story Means (R = 1-3). 85
7	Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score for Experimental Male Subjects for the Group x Sex x Story Interaction 87
8	Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score for Control Male Subjects for the Group x Sex x Story Interaction. . 87
9	Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score for Experimental Female Subjects for the Group x Sex x Story Interaction. 88
10	Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score for Control Female Subjects for the Group x Sex x Story Interaction. 88
11	Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score on Story 1 for the Group x Sex x Story Interaction. 90

LIST OF TABLES (Continued)

	Page
Table	
12 Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score on Story 2 for the Group x Sex x Story Interaction	90
13 Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score on Story 3 for the Group x Sex x Story Interaction	91
14 Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score on Story 4 for the Group x Sex x Story Interaction	93
15 Summary Table for Newman-Keuls Post Hoc Comparisons on Average Morality Score for the Group x Sex Interaction	94

CHAPTER I
INTRODUCTION

Recent societal concerns about morality have provided a strong impetus for renewed study of moral development.

It is probably more than coincidence that the student activism of the 1960's, the rising crime rate, and such heavily publicized events as Watergate and the street murder of Kitty Genovese have their parallels in the intensification of research on internalization of moral norms. (Hoffman, 1977, p. 275)

Concerned social critics such as David Reisman (Janis, 1969) and Erick Fromm (Janis, 1969) have pointed to the trend in modern society toward other-directedness rather than inner-directedness for the regulation of conduct. Among many people in powerful managerial and executive positions, deceit and exploitation, formerly inhibited through the individual's own moral code, are now deterred only by the threat of punishment. Across the United States, a pattern of crime has emerged that is both perplexing and appalling. "A new remorseless, mutant juvenile seems to have been born, and there is no more terrifying figure in America today" (The Youth Crime Plague, 1977, p. 18). The power blackout on July 13, 1977, plunged New York City into a darkness followed by an orgy of looting and wanton destructiveness that shocked the city and most of the world. Normal explanations seemed inadequate, as the people involved in the

crimes did not seem to be needy or hungry. "It was just like The Lord of the Flies. People resort to savage behavior when the brakes of civilization fail," wrote Ernest Dichter, a noted psychologist (Night of terror, p. 17). These indications pointed toward a time of moral crisis, described by Hall and Davis (1975) as a state of society in which people have, for whatever reason, lost their sense of what is right and what is wrong. Societies have dealt with this problem before. Solutions have ranged from stricter punishment to increased emphasis on moral education and were generally related to whichever theory of explanation was current.

When the study of morality was the exclusive domain of theology and philosophy, the origin of morality was attributed to the child's basic nature. Augustinian philosophy advocated that the child must be saved from his sinful nature by punishment. Followers of Locke maintained that the child's morality depended upon his training and experiences (Bijou, 1975). This latter view suggested the need for moral education.

Some people are convinced that the best process for reestablishing a foundation for ourselves in the moral domain is to reconsider and reevaluate the basic values that have been handed down through our religious and cultural traditions. (Hall & Davis, 1975, p. 19)

John Wilson, director of the Farmington Trust Research Unit in Oxford, England from 1965 until 1973, developed the first contemporary program for moral education (Hall & Davis, 1975).

Wilson's orientation was philosophical and centered on teaching the awareness of other people's feelings and emotions. This approach assumed that moral development could be fostered, a concept certainly not alien to that of John Dewey (1909), who advocated democratic experimental schools for moral and intellectual development seventy years ago. He believed that the aim of education was to aid in the building of a free and powerful character (Hall & Davis, 1975).

Lack of moral education in today's schools was suggested as a major contributor toward moral decline, but the mere phrase "moral education" stirred uneasiness in some educators (Sanborn, 1971). Social scientists had discouraged the idea of imposing middle-class educator values on all children, and educators were encouraged to accept variation in moral standards as a democratic ideal. Rather than creating a situation of greater individual freedom in the area of moral beliefs and ideals, the official laissez-faire policy produced what amounted to a shallowness of thought and commitment (Hall & Davis, 1975). Hersh and Paolitto (1977) presented strong possible implications for moral development in the schools. Strategies such as role playing, peer counseling, learning ethical philosophy, tutoring, interviewing, and moral discussions to stimulate moral development were suggested. Mills (1977), in answer to questions concerning problems of living in a society where there is no common religious perspective, suggested that

schools should inform the pupils about different traditions and should emphasize the salient points in the various moral beliefs. If possible, pupils should be introduced to people who claim that such traditions and beliefs have helped them in their moral lives. Mills further stressed the role of the teacher as a moral model before the students.

Teachers must show that they are not only concerned with academic problems of attributing moral appraisals, but that they are also concerned to put these moral evaluations into practice. There are two aspects to this: (a) the example of personal habits and behavior, and (b) a wider concern for morality in society.
(p. 167)

There appears to be new interest in re-assuming this responsibility (Hersh & Paolitto, 1977; Mills, 1977). It appears that the current increase in lack of moral development has become the stimulant. It was the purpose of this study to investigate the methods most effective for obtaining positive results.

Theoretical Orientation

There were three major theories which attempted to explain the origin of morality (Hoffman, 1970; McCandless & Evans, 1973). The concept of "original sin," which was included in most monotheistic religions, assumed early intervention by adults as essential to the ultimate salvation of the child. Classical psychoanalysis also accepted this concept. The young child was viewed as a bundle of libido drives which had to be subordinated by significant adults

in his life, precursors to the development of the child's own superego or internal control. In contrast, the doctrine of "innate purity" viewed adult society as a possible corrupting influence which had to be minimized in the early years of the child's development. Many cognitive-developmental psychologists accepted this doctrine, which placed great emphasis on the role of higher mental processes in moral development. The third doctrine, tabula rasa, assumed that the infant was infinitely malleable. Social learning theory was a descendant of this doctrine.

Each theoretical approach, in keeping with its philosophical base, defined morality and guided the empirical research in its own domain. Freudian research was concerned with the morality index guilt which resulted when moral standards were violated. The warning function of the conscience was described as associated with guilt (Fenichel, 1945). The belief was that the individual who had a highly developed sense of guilt would be a person of high moral character, self-controlled to avoid the punishment of internalized guilt (Redl & Wineman, 1951). Social learning theory emphasized strengthening "good" behaviors through reinforcement. Social learning theory also viewed the construct guilt as one of the measurable indices of moral internalization. This theory generated research which attempted to relate child-rearing practices to moral development (Hoffman, 1970). The cognitive developmentalists

believed that moral development was in large part a function of maturation within a context of general age-related experiences. In this latter view, the human organism matured in accord with his biological nature interacting with his environment. Although specific teaching was not the primary concern of the cognitive-developmental school, Kohlberg and Turiel (1971), proponents of this school, developed a general plan for moral education that resembles programmed learning.

Within the theoretical framework of each of the three major doctrines of morality, teaching has a definite, if variable, role in the child's moral development. Social learning theory, however, through its research relating child-rearing practices to moral development (Hoffman, 1970) and the Hoffman and Saltzstein (1967) research that found high moral development in children positively related to induction techniques of child-rearing, inspired this research. In the present study, the induction technique, (which is essentially a teaching method), was experimentally tested as a potential educational method for increasing the level of moral development in first-grade children.

Background for the Study

The fact that our ancestors lived and died for thousands of years without showing the slightest interest in children as children was supportive of the concept that children are

no more intrinsically interesting than sardines or prime numbers (Watson & Lindgren, 1973). The great interest in child development today has resulted from a special combination of values, attitudes, and precepts that characterized our society. The study of child development may be viewed as a byproduct of our strivings to become democratic. American parents, unsure of themselves as educators of children for a society that offered status flexibility in its social structure, consulted child specialists both about their child's behavior and about their own attitudes toward him (Dolto, 1955). This desire of parents to know what experiences would help children realize their best potential as adults has continued to be a compelling impetus for research. A strong current of interest in personality and socialization appeared in the 1950's, and of particular interest was the study of the effects of child-rearing practices on such personality variables and constructs as anxiety, hostility, aggression, need for achievement, and morality. This interest was due partly to the strong psychoanalytic orientation that was prevalent in the child guidance movement (Watson & Lindgren, 1973). Psychoanalytic theory, which stressed the importance of the parent, provided the main theoretic inspiration for most of the research in the role of parental practices in shaping and determining moral character (Hoffman, 1970). Researchers have generally focused on the child's anxiety over anticipated loss of

parental love as the prime motivator for the child's obtaining control over his behavior in accord with parental dictates (Sears, Maccoby, & Levin, 1957). Love withdrawal, with its highly punitive quality, was found to relate infrequently to moral indices in several studies (Burton, Maccoby, & Allinsmith, 1961; Grinder, 1962; Sears, Rau, & Alpert, 1965). Hoffman (1975a) reviewed the research on child-rearing antecedents of guilt and found support for the generalization that moral orientation was associated with the mother's frequent use of inductive discipline techniques which indicated the consequences of the child's behavior for others.

The effectiveness of induction as discipline, as compared to power assertion and love withdrawal, appears to be based less on the fear of punishment and more on the child's connecting its cognitive substance with his own resources for comprehending the necessities in the situation and controlling his own behavior accordingly. (Hoffman, 1970, p. 286)

Of the studies reviewed, all were correlational and provided no justification for inferring causality. There was some evidence to suggest that the child often affected the parents' behavior (Bell, 1971; Harper, 1971; Rheingold, 1969). A definite answer to the problem of causal direction required experimental work. The background literature on the variables believed to be associated with moral development and the existing dearth of experimental work were motivators for the present study.

Statement of the Problem

The present study was designed to determine the effect of morality content stories on first-grade children by measuring their responses to a Moral Development Inventory developed by the researcher. The problem considered in this study was suggested by Hoffman and Saltzstein's (1967) finding of a high relationship between inductive child-rearing techniques and high moral development in the child, as measured by the index guilt. The morality content stories were chosen to duplicate inductive techniques; that is, communicate information to the child concerning the painful consequences of his misdeeds for others. The concept that moral development could be fostered through literature was advocated by John Dewey more than seventy years ago (Hall & Davis, 1975), but it had not been experimentally tested.

A further purpose of this study was to develop an instrument for determining the level of moral development in first-grade children by measuring the index guilt. Guilt had been widely accepted as an indice for measuring the level of moral development (Ausubel, 1958; Freud, 1964; Hoffman & Saltzstein, 1967; Puiynski, 1975), and it was used as the dependent variable in this study.

Questions and Hypotheses

The basic questions in this study were the following:

1. Is it possible to foster moral development in the first-grade child by reading him morality content stories, chosen to duplicate the inductive method described by Hoffman and Saltzstein (1967)?

2. Is it possible to develop an instrument for assessing moral development in first-grade children which will satisfy the criteria for reliability and validity?

3. Is the testing of first-grade children on a multiple-choice test adversely affected by response mode?

4. Do the levels of guilt, described by Hoffman and Saltzstein (1967) (see Appendix B), appear in the first-grade child?

5. Do male and female first-grade children differ in level of moral development as measured by the Moral Development Inventory?

The following null hypotheses were stated:

1. There is no difference in level of moral development as measured by the Moral Development Inventory in the experimental group of first-graders read morality content stories and the control group read stories without morality content.

2. There is no difference in level of moral development measured by the Moral Development Inventory as a function of sex of subject.

3. There is no relationship between teacher rating of first-grade children's level of moral development and the children's rating on the Moral Development Inventory.

4. The subject's response to the multiple choice test is not related to any response set order preference.

5. There is no relationship between the children's level of language development as measured by the Comprehensive Tests of Basic Skills and their level of moral development as measured by the Moral Development Inventory.

Implications of the Study

The topic of the present study has implications for parents, educators, and child development specialists. Recent years have seen the accumulation of a body of findings relating moral development to parental practices. This study, which experimentally tested the relationship between the induction technique, a potential child-rearing variable, and moral development in first-grade children offers further guidance to parents along this dimension.

It was found that a positive relationship existed between the children's moral development and their mental health (Johnson & Medinnus, 1974; Mowrer, 1967). It would seem plausible that data which delineated methods for fostering moral development would also be advantageous to all persons who had contact with children and were concerned about their total development. The educator assumes considerable

responsibility for the development of children when children enter the first grade. Information concerning the relevance of morality content literature on the children's moral development would have value for teachers and school curriculum coordinators.

A dearth of information existed with respect to younger children and to the presence of guilt as an internalized moral controller. The development of an instrument for the measurement of guilt as an indicator of moral development would give educators and researchers the means to evaluate the moral effectiveness of various programs used with first-grade children.

Finally, the study had design implications for experimentally testing that which has previously been studied correlationally, e.g., child-rearing methods and moral development of children.

Limitations of the Study

The following limitations were considered pertinent to any interpretations of this study.

1. Generalizations are limited to subjects that are like those in the study--predominantly white, middle-class, first-grade children from Walton County, Florida. Demographic data from the Walton County school system was used to justify the middle-class designation. As such, the subjects were similar to those that formed the body of research on moral

development (Hall & Davis, 1975; Hurlock, 1964; Yarrow, Campbell, & Burton, 1968).

2. The subjects were limited to first-grade children. It was the intent of this study to determine whether guilt, an index of moral development, existed and could be measured in first-grade children. As such, this study may have contributed valuable data for answering questions concerning the age when the emotion guilt first appears in children. There was no attempt in this study to evaluate age-trend effects. The first-grade subjects used were all within nine months of the same age. At least three separate age levels must be studied in order to determine the relation of the phenomena in question to age (Anderson, 1956; Hicks, 1973; Winer, 1962). The present study then was cross-sectional and could only have developmental implication when viewed in the total context with other age-specific studies. Anderson (1956) noted that as early as the 1950's there were particular level studies, each adding to the total picture of development.

3. Only the subjects that were present for the five consecutive days of the experiment were included. In this sense, the study population may be seen as a self-selected group. The use of four different classrooms was expected to reduce the selectivity bias of absenteeism.

4. The Moral Development Inventory, adapted from the inventory used in the Hoffman and Saltzstein (1967) study, was validated by determining the relationship between the teacher ratings of the child's probable responses in a moral situation and the child's responses to the test. There was ample precedent for this method (Hoffman & Saltzstein, 1967; Miller & Swanson, 1960; Sears, 1961). It was not feasible in this study to compare the child's verbal responses with his overt behavior. This is definitely a limitation, the elimination of which justifies further research.

5. Subjecting the Moral Development Inventory to a test-retest reliability evaluation had the possible instrument effect error discussed by Turiel (1966) in discussing Kohlberg's instrument for measuring level of moral reasoning. The story content of the Moral Development Inventory may have contributed to the child's choosing the answer he remembered from the previous test administration.

6. The treatment effect was limited to five consecutive days. Although there were precedents for the effectiveness of brief treatments (Ambron & Irwin, 1975; Johnson & Medinnus, 1974), experimental designs have often been limited by conditions of accessibility to subjects. This researcher was convinced that the validation of the Moral Development Inventory required the close teacher-student familiarity that increased over time. Therefore it was considered necessary to run the experiment late in the

school year. Permission to intervene in the program of the public schools was thus limited to one week because of the inevitable rush of events that had to be accommodated in the latter weeks of school. It was also a limitation of this study that no provision was feasible for retesting terminal treatment effect.

7. Inherent in all research using constructs is the limitation of the research findings and interpretations to the specific definition of terms outlined in the study. Adherence to those definitions is essential for replication studies (Cronbach & Drenth, 1972).

Assumptions

The following assumptions were made concerning the present study:

1. Guilt was a measurable index for ascertaining level of moral development in first-grade children.
2. The children's responses to the Moral Development Inventory indicated identification with the protagonist in the stories.
3. The first-grade children chosen for participation in the study were representative of a larger general population of middle-class first-grade children.
4. The teachers' ratings for the children were reliably reported.

Definitions of Terms

The following definitions were established for this study:

Construct validity refers to how well the data generated by a test enables conclusions to be drawn about the nature of the particular construct with which the researcher is working (Cronbach, 1970).

Empathy is the ability to appreciate how someone else feels by putting yourself into his position and experiencing his feelings.

Guilt is a kind of negative emotion which occurs when an individual acknowledges that his behavior is at variance with a given moral value to which he feels obligated to conform (Ausubel, 1955). In the present study guilt is considered an index of moral development.

Identification is the gradual process by which individuals adopt as their own the characteristics, attitudes, values, or beliefs of others. Most frequently, identification will be with a parent (Johnson & Medinnus, 1974).

Induction refers to the process whereby the child has pointed out to him the painful consequences of his misdeeds on others (Hoffman & Saltzstein, 1967).

Internalization is the incorporation of the values and beliefs of others as one's own (Hurlock, 1964).

Moral development refers to the process whereby individuals adopt the standards of right and wrong laid down by their culture.

Socialization is a lifelong process by which individuals develop the values and beliefs, knowledge, awareness of social expectations, and appropriate role behavior (Watson & Lindgren, 1973).

CHAPTER II

REVIEW OF LITERATURE

The predominance of guilt as an index of moral development was reflected in the diversity of literature on the subject. Psychoanalytic, cognitive-developmental and social learning theorists have approached the study of moral development from different perspectives, varying radically on conclusions drawn. Major areas of concern included morality antecedents and prerequisites for moral internalization. Such variables as childrearing techniques, innate predispositions, and developmental age trends were studied (Sears, Maccoby, & Levin, 1957; Watson & Lindgren, 1973; Yarrow, Campbell, & Burton, 1968). Tenuous conclusions from these studies stressed the need for experimental research.

No attempt was made to survey the entire field of literature related to guilt and moral development. The literature reviewed included studies relating the inductive childrearing technique to the child's morality. The literature from the three major psychological theories of moral development was reviewed for background enrichment. Although the etiology of guilt includes sex guilt, hostile guilt, terminal guilt, maximum guilt, and morality-conscience guilt (Hoffman & Saltzstein, 1967; Mosher, 1966), only the latter two areas were reviewed in the present study.

The study of morality has slipped in and out of focus as a central area of interest in child development (Albert & Kluckhorn, 1959). Generally, the study of moral development has concentrated on how individuals adopt the standards of right and wrong that are laid down by their culture. The major concern has been with the internal aspects of socialization as a moralizing process and centered on the questions regarding the internalization of society's values and mores. In the descriptive sense, socialization is a class name for all the behaviors of the child that concern other people, directly or indirectly. If a child is a delinquent, for example, a fault in socialization is said to be present. If he is withdrawn or unpopular, he is considered to be inadequately socialized. If a little boy plays exclusively with little girls, his socialization is questioned. The list of behaviors and attributes included under socialization is as extensive as the possible range of social relations and characteristics that go into the product (Reese & Lipsitt, 1970).

According to psychoanalytic theory (Freud, 1949), the child is controlled by inborn primitive drives which he is forced to give up because of parental discipline. Immoral behavior was said to be motivated by unconscious needs for punishment and blame (Bijou, 1975). Cognitive-developmental theory (Kohlberg, 1963; Piaget, 1932) assumed that external rules became transformed into internal principles through cognitive processes. Accordingly, the child went through

the stages of heterony (rules are laid down by adults), equality (rewards and punishment are distributed equally and punishment is related to misdeeds), and equity (judgments are based on real situations) (Bijou, 1975). Social learning theory (Bandura & Walters, 1963; Hoffman & Saltzstein, 1967; Sears, Maccoby, & Levin, 1957) stressed child-rearing practices as morality antecedents.

The expectations are that punishment given by parents under certain conditions will bring about feelings of anxiety and will inhibit wrongdoing quite independent of whether, in the current situation, the child is punished. Avoidance and inhibition learned in the home should be generalized to any and every situation outside even when there is no supervision of the child. Thus the conscious build-up through reward and punishment will be taken into all situations. (Lovell, 1971, p. 102)

The development of conscience, as evidenced by guilt, has been presumed to reflect the internalization of moral standards; but the child-rearing antecedents of guilt have not been fully established. Henry and Short (1954) suggested that internalization of parental characteristics as evidenced in guilt was most likely to occur when the disciplining parent was also perceived as a source of affection. King and Henry (1955) found that males who reported their mothers as dominant in discipline showed a cardiovascular reaction under stress characteristic of guilt.

Social learning theory generated research which attempted to relate child-rearing practices to moral development (Hoffman, 1970). Several of these studies found correlations between guilt, one of the indices of morality, and

childrearing practices (Allinsmith, 1960; Allinsmith & Greening, 1955; MacKinnon, 1938; Sears, Maccoby, & Levin, 1957; Whiting, 1954). Hoffman (1975a) reviewed the research on child-rearing antecedents of guilt and found support for the theory that moral orientation, measured by high guilt, was associated with the mother's frequent use of inductive techniques. These techniques involved pointing out to the child the consequences of his behavior on others. Other researchers (Allinsmith, 1960; MacKinnon, 1938; Whiting, 1954) also found a relationship between high guilt in the child and his mother's use of induction, or reasoning. The use of induction techniques to foster moral development has not, however, been experimentally tested.

The Psychoanalytic Theory of Moral Development

According to psychoanalytic theory, morality is primarily motivational, emotional, and unconscious. "Moral standards are largely unconscious products of powerful irrational motives and are based on the need to keep antisocial impulses from becoming conscious awareness" (Hoffman, 1970, p. 261). The moral individual wrestled with the biological drives of the id, was denied biological fulfillment, and solved his problems by incorporating his view of parental attitudes as his own. The introjected parent was always, to a large extent, the fantasy product of the child who felt guilty about his sexual inclination toward that parent (Munroe, 1955). The conscience, formed as a resolution of the oedipal complex,

became strengthened by the admonitioning of the adult and formed the internal control of impulse necessary for adequate social function. The child who was deprived of an adequate parental figure would tend to remain at the narcissistic level. This lack of a person to identify with would produce a conscience-less individual who loved only himself.

In psychoanalytic theory the superego served as the vehicle for the conscience. It developed as an aftermath of the oedipus complex and dealt with the ego as a strict father would deal with his child. When the ego and superego conflicted, the superego served as pressure upon the ego, and the child felt guilty. Social stability was thus achieved, since the child controlled his sexual and aggressive impulses (Freud, 1949).

According to Freud (1949), girls did not resolve the oedipus complex quickly and dramatically, and therefore did not identify with parents as fully as boys did. This meant that girls had less internalized moral structures. This theory was contradicted by research studies which indicated that girls had more highly internalized moral structures than boys (Hoffman, 1975a).

Ausubel and Sullivan (1970) disagreed with Freud's emphasis on the superego. They believed that moral development was synonymous with ego development and that there was no necessity for considering a separate layer of the personality such as the Freudian superego.

The potential for guilt, according to psychoanalytic theory, originated with the beginning of human history, in the archaic structure of man's psyche (Puiynski, 1975). Bad deeds were then the result of the primitive existence of guilt. The natural moral order was passed down through antiquity, and moral maturity was a gradual transition from heterony to autonomy. Moral autonomy was the pinnacle of the integrated hierarchy of moral values. The child passed gradually through the phases of dependence on important adults and peers and became self-sufficient and independent. A positive emotional link between child and parent facilitated moral development. Rewarding through love and punishing through depriving the child of love was considered more important than material awards or physical punishment in his moral development. The socially conditioned dynamic development of the correct personality of the human being was a process of gradual transition from a system of external control to internal control as the child identified with the parent and accepted the parents' standards of conduct as his own, and as introjection (taking the thought or feeling of someone else as one's own) occurred (Izard, 1977).

The Cognitive-Developmental Theory of Moral Development

The cognitive-developmental approach to moral development is nativistic and assumes that the child is a moral philosopher (Kohlberg, 1963). External rules become transformed

into internal principles through cognitive processes which evolve developmentally in stages (Bijou, 1975). Research generated by the cognitive-developmental approach focused on the development of moral judgment, determined by analyzing children's responses to stories involving ambiguous moral situations. Moral judgment was emphasized, rather than moral behavior.

The cognitive-developmental approach was first postulated by John Dewey (1909). Three levels of moral development were outlined: pre-moral, conventional, and autonomous. The premoral individual was primarily a biological-impulse organism, the conventional level individual accepted the standards of the group, while the person at the autonomous level guided his conduct by thinking and judgment. Whereas Dewey's thinking about moral stages was theoretical, Jean Piaget made the first effort to define stages of moral reasoning in children through observation (Kohlberg, 1975). The essence of morality, according to Piaget, included both the individual's respect for the rules of social order and his sense of judgment (Hoffman, 1970). The sense of justice was investigated by telling stories about persons who committed transgressions and asking children to evaluate the act according to the degree of "wrongness." From these procedures, Piaget delineated two broad stages of moral development. In the first stage, moral realism, the child complied with rules and viewed behavior as totally wrong or totally right. He

believed in imminent justice and judged an act by the magnitude of its consequences. In the latter stage, autonomous morality, the child perceived rules as modifiable in response to human need. According to Piaget, both maturation and experience were essential in the transition from one stage to the next. Moral reasoning could be developed only by cooperation and genuine intellectual changes on the part of children. Changes in the child's behavior could come about only through time, although it was not time itself but rather the activities of the child in interaction with his environment that led to new structuring of the mental organization (Piaget, 1970). Piaget suggested that role-taking experiences which produced cognitive disequilibrium were major contributors to moral internalization. Glassco, Milgram, and Youniss (1970) found that children, exposed to an adult model who used more mature criteria of intentions, shifted their verbal responses toward that of the model. Although this exposure to the adult model who assigned greater weight to intentions may have produced cognitive disequilibrium which the subjects reduced by changing their views, the findings did not support the cognitive-development assumption of directional change. A model who stressed consequences over intentions, a lower-level cognition, would also be expected to produce cognitive disequilibrium (Hoffman, 1977).

While Piaget (1932) stressed the importance of peer interaction, it was also made clear that the parent's

child-rearing practices also played a role in moral development.

In order to remove all traces of moral realism, one must place oneself on the child's own level, and give him a feeling of equality by laying stress on one's own deficiencies. In the sphere of clumsiness and untidiness in general (putting away toys, personal cleanliness, etc.), in short in all the multifarious obligations that are so secondary for moral theory but so all-important in daily life (perhaps nine-tenths of the commands given to children relate to these material questions) it is quite easy to draw attention to one's own difficulties, even one's own blunders, and to point out their consequences, thus creating an atmosphere of mutual help and understanding. In this way the child will find himself in the presence, not of a system of commands requiring ritualistic and external obedience, but of a system of social relations such that everyone does his best to obey the same obligations, and does so out of mutual respect. (pp. 133-134)

Research findings on age trends in moral development provided considerable support for Piaget's theory that the child's morality level was related to his developmental level of reasoning (Grinder, 1964; Liu, 1950; Whiteman & Kosier, 1964). Havighurst and Neugarten (1955), however, in an extended study of several Indian tribes, found evidence against the universality of Piaget's stages. Bandura and McDonald (1963) attempted to demonstrate that Piaget's (1948) sequence of moral development changes was a function of reinforcement and imitative learning. Children were assigned to stages in terms of responses to paired story acts, one a well-intentioned act resulting in considerable damage and the other a maliciously motivated act which resulted in very little damage. The experimental subjects then observed

adult models who were reinforced when they expressed judgment opposite to the child's orientation. The child's own responses were reinforced when they ran counter to his dominant mode. The posttest used different stories, adding methodological strength to the study. The findings were that children were influenced to judge on the orientation mode opposite to their initial one. Bandura and McDonald concluded that this cast doubt on the validity of a developmental stage theory of morality. The weakness in this study was that the posttest was given immediately after the experimental treatment and that there was no follow-up test. Duration over time was considered by Piaget to distinguish between cognitive structures and superficially learned responses.

In recent years, Piaget's theories on moral development have been intensively scrutinized. Magowan and Lee (1970) investigated possible shortcomings in the story methodology. It was found that more immanent justice responses were made to familiar stories. Medinnus (1959) re-examined the problem of immanent justice in a study of 240 children aged six to twelve, and concluded that a child's expressed belief in immanent justice was a function of the meaningfulness of the story described, the presence or absence of rational alternative explanations, and the range of the child's experiences. Others (Dennis, 1943; Durkin, 1959; MacRae, 1954) had similar findings. The evidence strongly demonstrated the presence of situation-specific responses to Piagetian moral judgment

stories. Siegal (1975) offered the theory that these situation-specific responses represented problems in learning moral language, as the child often appeared to speak in self-contradictions.

In 1955 Kohlberg began to redefine and validate Piaget's stages of moral development. His aim was to retain the best of Piaget's schema and fit it into a more logically consistent framework. Kohlberg (1974) studied the same group of 75 boys from early adolescence through young manhood, and additionally, explored moral development in Great Britain, Canada, Taiwan, Mexico, and Turkey. The typology developed contained three levels of moral thinking, each level divided into two stages. The six stages, summarized, are as follows:

1. Obey rules to avoid punishment.
2. Conform to obtain rewards, have favors returned, and so forth.
3. Conform to avoid disapproval, dislike by other.
4. Conform to avoid censure by legitimate authorities and resultant guilt.
5. Conform to maintain the respect of the impartial spectator judging in terms of community welfare.
6. Conform to avoid self-condemnation. (Kohlberg, 1963, pp. 13-14)

A more complete definition of Kohlberg's moral stages is placed in Appendix A.

Kohlberg's (1963) stages of moral development implied that stages are organized systems of thought, individuals are consistent in level of moral judgment, the stages form an invariant sequence, and that while thinking at a higher level includes within it lower-stage thinking, there is a tendency for an individual to function at or prefer the highest stage available. Certain seemingly incompatible factors, such as low correlations found between maturity of moral judgment and IQ, were given less credence by Kohlberg than the philosophical logic that moral reasoning is clearly reasoning and advanced moral reasoning depends upon advanced logical reasoning (Stanton, 1976). Mature moral judgment was deemed to be a necessary but not sufficient condition for mature moral action. Krebs (1968) suggested that sixth-grade children at stages five and six resisted temptation to a greater extent than children at stages three and four. The conclusion was that the person who understood justice was more likely to practice it.

A major difference between Piaget's and Kohlberg's moral stages was the stronger emphasis by Kohlberg on the natural outgrowth of cognitive development (Hoffman, 1970). According to Kohlberg, participation in groups provided the individual with experience in taking alternate roles, thereby receiving a broader social perspective; but the particular authority to which he responds is determined by the child's comprehension of who is deemed to be the most legitimate

authority. This comprehension depends on the specific point of the child's cognitive development.

Kohlberg's stage theory was experimentally tested by Turiel (1966), using 47 seventh-grade boys. These subjects were assigned stage positions based on their responses to six of Kohlberg's stories about individuals in moral dilemmas. The subjects assigned to the experimental group were presented three different Kohlberg stories, instructed to take the role of the central figure in each story, and seek advice from the experimenter. The experimenter's advice consisted of arguments either one or two stages above or one stage below the subject's prior stage rating. When the subjects were retested a week later, it was found that although the experimental group did not shift to a significantly greater degree than the control group which had experienced no treatment, the net shift (experimental group minus control group) was greater for the +1 than the -1 group. The tentative conclusion was that Kohlberg's stages two, three, and four, the ones used in this study, were successive stages of cognitive levels and that children presented with statements at different moral levels tended to prefer those above rather than those below their own level. A weakness in the study design was that Kohlberg's moral judgment interview, which contained nine hypothetical conflict stories, was used as both pre- and post-test.

There is more agreement on Kohlberg's three levels of thought (preconventional, conventional, and post-conventional) than on the specific six stages (Hall & Davis, 1975). Kurtines and Greif (1974) pointed out that in Kohlberg's longitudinal study, there were few significant changes in moral judgment over time, and that there was no evidence that the students passed through the stages in a fixed order. College students often obtained lower scores than when they were in high school. It was also suggested that there were some anomalies in research relating Kohlberg's stages to overt behavior. Stage three characterized delinquents in one study and social conformists in another. Holstein (1976) found evidence against the expected stepwise progression. Kohlberg's research was only found to have demonstrated irreversibility and invariance in stages two, three, and four (Simpson, 1974). Another difficulty with Kohlberg's model was that it had not clarified whether the stages of moral development were strictly consecutive (a new mode of thought leaving behind the previous modes) or cumulative (the attainment of a new stage involving a continuation of the previous modes of thought) (Hall & Davis, 1975). Kohlberg's earlier views suggested that the stages were consecutive and mutually exclusive, but this was implied more than demonstrated. Williams and Williams (1970) held that the highest level of moral thought was one in which the individual had different and simultaneously existing modes of thought as

they related to any given situation; but objective, supportive research was lacking.

Kohlberg's criterion for assessment was attacked by Stanton, (1976, p. 614) who stated:

Kohlberg's typology and theoretical views are based on studies in which the moral dilemmas used were far removed from the reality of his subjects' experience as, for example, in questions relating to a story in which a man stole drugs to save the life of his wife (Kohlberg, 1963). This rather extreme dilemma, . . . would appear to place the emphasis on assessing intellectual ability required to interpret complex relationships rather than the quality of thinking related to moral judgments in real life.

In a lengthy criticism of Kohlberg's work and his claim for universality, Simpson (1974) suggested that his assumptions were ethnocentric and culturally biased. Research by Witkin (1974) and Hsu (1961) concurred that subculture may influence expectations and judgments in many ways. The greatest discrepancies in Kohlberg's theory were noted in the formulation of stages and phases beyond the onset of adolescence. Others, however, (Bull, 1969; Kay, 1968), found a very general agreement on the nature of the sequence of stages.

Kohlberg's theory has provided guidelines for new efforts in moral education. Young inmates in a New England reformatory participated in an experiment on ethics (Sanborn, 1971). The subjects were operating primarily at Stages one and two when the experiment began. Through open discussions, the inmates were exposed to more sophisticated moral concepts. As the experiment progressed, most of the subjects moved into stage four.

Although Kohlberg believed moral development to be a function of maturation facilitated by age-related experiences, leaving the occurrence of those age-related experiences to chance was not suggested. A general prescription for moral education that resembles programmed learning (Kohlberg, 1966) was advanced. Moral education, it was stated, should be directed toward bringing children through the stage of cognitive development step by step. Children should be given opportunities for more advancement in moral development (Kohlberg, 1969). Holstein (1970) offered support for this theory. Upper-middle-class parents at a principled level tended to have children who participated in moral discussion and were more advanced in moral development than those of parents who did not provide such opportunities for their children. However, in a review article by Kurtines and Greif (1974), much criticism was leveled against Kohlberg's theory and assessment of moral judgments. "In general, neither the cross-sectional, longitudinal nor sex difference data support the notion of an invariant developmental sequence" (p. 467). Holstein's (1972) three-year longitudinal study of moral development found no sequential progression from one stage to the next. Moreover, for the adults, as well as 16-year-olds, the most frequently occurring model response was stage three and four.

Kurtines and Greif (1974) pointed out measurement problems in scoring Kohlberg's dilemmas, and expressed concern

about the lack of data supporting a correlation between moral judgment and moral behavior at various levels of development.

They concluded by stating:

It is difficult to make a definitive statement about the ability of Kohlberg's cognitive-developmental model of moral development. After 15 years of research, the general lack of evidence for the model is suggestive. The possibility remains that the stages do reflect actual development and that the general lack of evidence reflects the inadequacy of the measuring device. . . . we can conclude that the value of the model remains to be demonstrated. (p. 469)

The Social Learning Theory of Moral Development

Social learning theorists (Bandura & Walters, 1963; Hoffman & Saltzstein, 1967; Sears, Maccoby, & Levin, 1957) stressed the motivational and emotional components of morality. As a theory, it differed from psychoanalysis in the assumptions made concerning the processes of internalization of moral values. Lovell (1971) summarized the social learning theory position:

The expectations are that punishment given by parents under certain conditions will bring about feelings of anxiety and will inhibit wrongdoing quite independent of whether, in the current situation, the child is punished. Avoidance and inhibition learned in the home should be generalizable to any and every situation outside even when there is no supervision of the child. Thus the conscious build-up through reward and punishment will be taken into all situations. (p. 102)

This position led to research which attempted to correlate parental practices with personality characteristics of children (Sears, Maccoby, & Levin, 1957) and other research which attempted to relate child-rearing categories (as power

assertion, love withdrawal, induction, and affection) to various indicators of conscience, such as guilt and resistance to temptation (Bijou, 1975). Most of these studies have been criticized because parental practices were determined by interviews and questionnaires (Yarrow, Campbell, & Burton, 1968), and because the laboratory procedures used operational rather than functional definition of concepts.

Behavioral research has concerned itself, not with internalized constructs that acted in the individual as mediators in self-discipline, but rather with environmental contingencies that altered behavior in the "desired" direction. According to Skinner (1953) self-controlling behavior in moral behavior comes into play when a response produces conflicting consequences.

The immediate and remote consequences of most acts are brought about by combinations of the physical environment. In moral self-controlling behavior, however, the long range contingencies are for the most part mediated by the social environment. (Bijou, 1975, p. 16)

Self-controlling behavior, then, was perceived as a chain of events that originated in the environment and was either strengthened or weakened by the environment (Skinner, 1969). Thoresen and Mahoney (1974) suggested teaching a child to engage in self-controlling behaviors by training him to respond to, observe, or discriminate his own behavior; to identify the conditions that precede and follow the behavior; and to engage in behavior that altered one or both of these

contextual conditions. Bijou (1975, p. 25) outlined a procedure for parents to use in teaching initial moral behavior.

The approach contained the following five steps:

- (1) Specify the goals of moral training.
- (2) Begin training at a realistic level.
- (3) Arrange conditions to expedite learning the target moral behavior.
- (4) Monitor progress and alter the procedures and goals as changes appear to be necessary, and
- (5) Maintain the acquired behavior.

Bijou's functional approach to the teaching of moral behaviors would account for situation specific moral behavior such as that noted in the classic Hartshorne and May (1930) study. It also would explain the appearance of mature moral behavior in children too young to satisfy that criterion through the cognitive-developmental approach. It was deficient in accounting for those behaviors judged universally moral, which present no evidence of having been taught.

One of the legacies of Freud, and the sociologist Durkheim as well, is the assumption now prevalent among social scientists that the individual does not go through life viewing society's central norms as externally and coercively imposed pressures to which he must submit. (Hoffman, 1970, p. 262)

The major concept in most moral development research has been the internalization of socially sanctioned standards of behavior, but theorists have disagreed on which socialization experiences are most likely to foster this internalization process. Several writers (Bandura & Walters, 1963;

Brim, 1968; Campbell, 1964) noted that internalization of control may merely reflect the norms of an absent reference group. Hoffman (1970, p. 264) stated that if

the behavior was based on the actor's positive attitude toward the reference figure . . . or concern about hurting or disappointing him--without regard for punitive consequences--then the behavior may be appropriately classed as internalized.

The individual who had internalized control, then, accepted society's standard as an obligation to himself. His motives were to avoid self-condemnation and guilt. This concept of internalization bypassed the cognitive-developmental theory of moral reasoning, but bridged psychoanalytic and learning theory on the variables of identification, modeling, and conditioning and the elevation of the role of the parent in the internalization process. It is especially similar to an individual's controlling behavior which Skinner (1953) believed came about when a response produced conflicting consequences. Certain parallels were drawn between the psychoanalytic concept "castration anxiety" and social learning theory's parental discipline technique of love-withdrawal, both of which were viewed as producing first fear in the child, then conformity behavior through identification, an aspect of moral internalization (Mowrer, 1952).

For several years, psychologists have viewed the kind of responses that an individual made to temptation as one of the indices of moral internalization (Fenichel, 1945; Redl

& Wineman, 1951; Sears, Maccoby & Levin, 1957). Social learning theorists have steered away from the construct moral internalization and centered on temptation behaviors, leaving the connection between the two variables to the theoreticians. Grinder's (1961) definition of temptation was based upon the elements of the temptation situation itself. It was stated that

. . . a state of temptation exists for an individual when the addition or presence of a rewarding stimulus (incentive) increases the probability of a class of responses which is forbidden because of its incompatibility (conflict) with some socially expected behavior. (p. 680).

Temptation, then, was a predecision-type conflict situation where an individual, free from danger of detection, must choose between a positive incentive and conformity to learned role behavior. Grinder designed a beanbag and a ray-gun game situation to assess temptation to transgress. Research by Burton, Maccoby, and AllinSmith (1961) indicated the procedures were successful. The findings in this study, however, were puzzling. Resistance to temptation was found to be negatively related to other interview measures of conscience, and it was suggested that resistance to temptation and guilt should be studied as distinct phenomena and not considered as alternate measures of internalized standards.

In a 1967 study, Hoffman and Saltzstein assessed seventh-grade children on several dimensions of moral development and compared them with parental discipline techniques coded

into three categories of power assertion, love withdrawal, and induction. This study challenged the hypothesis which stemmed from psychoanalytic and learning theories that anxiety over loss of love was the necessary motivational basis for moral development. It was found that the use of induction, a technique in which the parent pointed out the painful consequences of the child's act for the parent or other, was consistently associated with advanced moral development. One of the indices used as measurement of moral development was guilt.

The Induction Method and Moral Internalization

Research studies relating internalization of moral values and the capacity for guilt to parental practices characteristically divided parental discipline techniques into two categories: power assertion (physical punishment and material deprivation) and nonpower assertion (love-oriented discipline). Although the differential effects of these two methods were frequently discussed (Allinsmith & Greening, 1955; Hill, 1960; Sears, Maccoby, & Levin, 1957), Aronfreed's (1968) belief that the major differentiating variable was the information communicated by nonpower assertion techniques became a strong inducement to subdivide the nonpower assertion category and concentrate on this component. Hoffman and Saltzstein (1967) thus subdivided the nonpower assertion category into induction and love-withdrawal. In love-withdrawal, the parent gave direct but nonphysical expression

to his anger or disapproval of the child for engaging in the behavior. In induction, the parent pointed out the painful consequences of the child's act for the parent or for others. It was Hoffman's (1963) belief that this technique, supplying to the child knowledge that his actions caused pain to others, relied on empathy and thus provided a powerful emotional and cognitive support for development of moral controls. Martin (1975) described the inductive technique also as having a cognitive and an affective component. The cognitive component helped the child learn specific rules or limits for behavior, understand the reasons for such rules and anticipate the consequences for breaking such rules. Baumrind (1970), Parke (1969), and Pikas (1961) found that parents who used reason to legitimate their directives were more potent models and reinforcing agents than parents who did not encourage verbal exchange.

The affective component included an actual or implied threat of withdrawal of love, contingent upon the enactment of undesired behaviors. It made the child feel bad for behavior that hurt parents or others. It was hypothesized that if transgressions were followed in induction, the child would learn that the important part of transgressions consisted of the harm done to others (Hoffman & Saltzstein, 1967). It is possible that only inductions can arouse the need for love to an optimal degree, because the threat of love withdrawal implicit in inductions is relatively mild.

Inductions are less likely to disrupt the general affective relationship with the parent.

Yarrow and Waxler (1977) developed a research interest in how concern for others, or empathy, was born and bred. Their subjects were 128 children, ages 10 months to 7 years. The first study focused on the child's emerging sensitivities to affective events, such as a parent's or child's anger, pain, or joy. The second and third studies began with three-year-old children and investigated the development and relations of perspective-taking skills and prosocial behavior. They found that young children were often finely discriminating and responsive to others' needs, as evidenced by crying, holding hands over their ears, comforting a distraught parent or hitting the parent perceived the guilty one. The children's abilities to deal with another's perspective on the perceptual and cognitive tasks increased with age, the most substantial jump occurring between four and one-half and five years of age, but there was no overall relation between perspective-taking abilities and prosocial intervention. The researchers ventured the hypothesis that aggressions experienced may contribute to the development of sensitivity to feelings. In other words, direct learning experiences may be essential to the development of prosocial behaviors, whereas empathetic feeling may be more directly related to maturation.

Chandler (1973) demonstrated that certain empathetic perspective-taking skills could be taught, and suggested the utility of attempting to understand delinquent youth in terms of age-inappropriate egocentral thought. Other investigators (Feffer, 1970; Sarbin, 1954; Thompson, 1968) have found support for this view that prosocial behavior was related to the development of age-appropriate perspective-taking skills. Piaget and Inhelder (1956) and Looft (1972) found the social judgments of young children to be limited by egocentric thought.

Of the three types of parental discipline (love-withdrawal, power-assertion, and induction) frequently studied as a factor in the child's moral development (Hoffman & Saltzstein, 1967), induction seemed more capable of enlisting the child's natural proclivities for empathy in the struggle to control his impulses. Induction directs the child's attention to the other's pain, which should elicit an empathic response and also communicate to the child that he caused the pain. This combination of empathy and awareness of being the causal agent should produce a response having both cognitive and affective properties. "The child is thus gradually enabled to pick out on his own, without help from others, the effects of his behavior, and to react with an internally based sense of guilt" (Hoffman & Saltzstein, 1967, p. 55).

Previous arguments that the child below the age of six was not capable of empathic thought have been weakened by recent research. Piaget's famous three-mountain landscape task, on which children below six years of age typically made the egocentric error of attributing their own viewpoint to a doll situated in various locations around the landscape, was interpreted as suggesting severe limitations for the young child's empathetic or role-taking abilities (Hoffman, 1977). Ambron and Irwin (1975) studied 34 kindergarten children and 38 second-grade children on dimensions of role-taking and moral judgment. A significant correlation between role-taking and moral judgment was found. Seven-year-olds had significantly higher scores on role-taking than the five-year-olds. Other studies, however, have suggested that Piaget's original task, with its spatial-perceptual requirement, was a faulty method for measuring the role-taking competence of young subjects (Borke, 1975). Additionally, three-year-olds made very few errors when the display contained familiar toys and a verbal response was not required. Even two-year-olds were observed to turn a picture toward another person who asked to see it. One and two-day-old infants were found to cry in response to the sound of another infant's cry (Simner, 1971).

Empathy has been studied as unfolding through a series of developmental stages.

Empathy is seen as having an affective component that is given increasingly complex meaning as the child progresses through the four phases in the development of a cognitive sense of the other. . . . Thus the infant's empathetic response to another's distress includes no awareness of who is actually in distress. With "person permanence," this awareness exists but the other's inner states are unknown and assumed to be the same as one's own. With role-taking development, empathy becomes an increasingly veridical response to the other's inner states in the situation. (Hoffman, 1977, p. 300)

Empathy, thus, was viewed as a vital part of the process of moral internalization. The most recent theoretical account of how induction fosters internalization was summarized by Hoffman (1977):

. . . (a) Most techniques have some power-assertive, love-withdrawing, and inductive properties. (b) The first two comprise the motive arousal component which is necessary to get the child to stop what he is doing and attend. (c) Having attended, the child will often be influenced cognitively, and affectively (through empathy arousal) by the information contained in the inductive component and thus experience a reduced sense of opposition between his desires and external demands. (d) Too little arousal and the child may ignore the parent; too much and the resulting fear or resentment may prevent effective processing of the inductive content, thus perpetuating the felt opposition between desires and demands. Inductions ordinarily achieve the best balance and are therefore most effective in the discipline encounter. Furthermore, in keeping with recent developmental research on memory . . . the child may be expected over time to remember the ideas communicated in inductions but to forget that they originated with the parent. With no external agent to whom to attribute these ideas, he may then be expected to attribute them to himself. (p. 301)

The Phenomenon of Guilt

Recent years have seen the accumulation of a body of research utilizing capacity for guilt as a criterion for

measuring internalization of moral values (Hoffman & Saltzstein, 1967). The concept of the guilt feeling as an important psychological phenomenon was introduced into scientific scholarship by Freud and the followers of his psychoanalytic concepts (Freud, 1964). Over the years, the various schools of psychology have argued the question of the significance of the experience of guilt for the correct development of the personality of the human being. Some adherents to classical psychoanalysis defined guilt as a conflict between the ego and superego (Lewis, 1971; Piers & Singer, 1953), but psychologists representing the neo-Freudians or newer directions of dynamic psychiatry are of divided opinions. Ausubel (1958) defined the guilt feeling as a specific type of negative assessment of one's self in which shame, self-persecution, grief, and a lowered feeling of respect for one's self all have their part. Ausubel, then, considered guilt as specifically an ego problem.

Psychoanalysis sought the sources of guilt in the beginnings of human history, in the archaic structure of man's psyche. Existential philosophy which proceeded from the ontological perception of the genesis of guilt, posited that the experience of guilt appeared at the moment of mankind's birth and is connected with the individual's failure to reach his fullest potential (Puiynski, 1975). Yap (1965), elaborating on this concept, emphasized that guilt was not the result of bad deeds, but the precursor of bad deeds. Puiynski (1975) suggested that guilt had not

only an existential character but also a genetical character. The concept of guilt as learned, however, has enjoyed greater acceptance (Whiting & Child, 1953). In secular life, it was assumed that the correct process of socializing the individual was based on the gradual development of the personality, with attainment of maturity in the spheres of emotions, perception, and aspirations.

Puiynski (1975), in a theoretical treatise, summarized the Freudian explanation of the development of guilt. This treatise indicated that the young infant, totally dependent, received satisfaction of his basic needs from his parents. A condition of the satisfaction of the need for dependence was the adaptation of the child to the demands and prohibitions of the parents. This was obtained by means of rewarding through love and punishing by depriving the child of love. The basic condition of proper development was a mutual positive link between the child and his parents. Threat of loss of parental love caused the child to develop a feeling of guilt. Moral internalization, through identification with the parent, was more intensive the younger the child. This socially conditioned dynamic development of the correct personality was a process of gradual transition from a system of external control to a system of internal control. The feeling of guilt was integrally related to the feeling of responsibility. Without a feeling of responsibility, there would be no guilt. The absence of responsibility is generally considered as a manifestation of emotional immaturity.

Bronfenbrenner (1968) noted that Russian children, reared in a collective atmosphere, developed guilt feelings over any desire to rebel. This was accomplished by the love-withdrawal pattern of discipline. Great emphasis was given to the development of social morality: obedience and self-discipline. Since each child's status depended in part on the status of his collective, it was in each child's self-interest to assume responsibility for his neighbor. There was a pattern of continuity between the discipline used by the parent and that used by the collective. In essence, the system said that the collective--like Mother--would not love you unless you behaved as it required you to do.

In examining the etiology of the guilt feeling, attention must be drawn to the possible need for punishment. Although Horney (1939) separated the guilt feeling from a need for punishment, empirical observations rather confirmed the position of those researchers who treated both elements as component parts of one experience (Puiynski, 1975). The experience was said to occur early in childhood, resulting in the learned sequence of guilt, punishment, and forgiveness. The child learned that condemnation does not last forever, but may be redeemed through punishment. Therefore, a condition of the internalization of moral values was not only a fear of punishment but also a desire for forgiveness. Religions introduced synonymous terms: the immoral deed was equated with sin and the feeling of guilt was equated with

the feeling of sinfulness. The need for punishment was equated with the feeling of sinfulness. The need for punishment was equated with the need for penitence, which reduced the feeling of guilt and brought the sinner psychic relief (Benedict, 1946). Mowrer (1961) also maintained that discipline and punishment were essential in the development of guilt and that confession and expiation were essential in the maintenance of mental health when one had committed a misdeed. Thus guilt was looked upon as a positive sign of a person's socialization. Reese and Lipsett (1970) termed guilt as the most commonly agreed upon criterion behavior from which moral or conscience identification is inferred. Achenbach (1966) found that internalized guilt correlated with fewer social problems and better school performance in children. And the Psychology Encyclopedia (1973) stated that, "Failure to have guilt feelings is a mark of the antisocial person who has a weak conscience and who lives at odds with the customs and laws of society" (p. 110). Izard (1977) described guilt as ". . . the emotion most essential to the development of the affective-cognitive structures of conscience and the affective-cognitive-action patterns of moral behavior" (p. 421).

Not all psychologists agreed that guilt was a positive variable. Ellis (1962) accepted that there is such a thing as human wrongdoing and that some standard is necessary for social group living. A standard sustained by inculcating in people a sense of guilt was rejected.

It was believed that self-blame and a sense of worthlessness led to emotional disturbance. Ellis' rational psychotherapy helped the person from this syllogism:

- (a) Wrongdoing is self- or society-defeating;
- (b) I have made a mistake or committed a wrong act;
- (c) Therefore, I'd better stop being self-defeating by acknowledging my wrong doing and take considerable time and effort to work at not repeating it, so that eventually I'll become a less frequent wrongdoer. (1962, pp. 144-145)

Although Ellis did not accept guilt as a prerequisite to socialization as did others with psychoanalytic background, he agreed with their acceptance of a biologically based guilt.

Like the tendency to blame others, the propensity to blame oneself (or be guilty) may also in part be biologically based. This is not to deny that much or most of the intense guilt of men and women is acquired in the course of their early upbringing; for it would certainly seem to be. But here again we must suspect that if virtually all humans in all parts of the civilized and uncivilized world are intensely guilty or ashamed of many things they do, man must somehow be the kind of animal who, par excellence, is guilt-inducible. (1962, p. 398)

Miller and Swanson (1960) referred to this phenomenon as each generation's re-experiencing man's fall from grace; and in differential emotions theory guilt is considered a fundamental emotion which emerged like other fundamental emotions, through evolutionary-biological processes (Izard, 1977).

Eibl-Eibesfeldt (1971), in agreement with differential emotions theory (Izard, 1977) and the psychoanalytic school (Rickman, 1957), argued that there is a genetic basis for the development of a sense of personal responsibility and

guilt. Ausubel (1955) suggested that the biogenetic mechanism for guilt was so basically human that it would develop in all cultures if conditions were even minimally favorable. The single most important basis for the development of conscience and guilt, according to Ausubel, was the powerful need for parents and society to develop a sense of responsibility in growing children. Erikson, a neo-Freudian, explained the development of guilt as a failure to resolve the conflict of the Oedipal crisis.

In the phallic period the resolution of the Oedipal crisis leads to a heightening of conscience, and it is the time when the child needs to develop the prerequisites for masculine or feminine initiative or become prey to a deep and lasting sense of guilt. (Lidz, 1968, p. 81)

In Mowrer's (1961) view guilt developed essentially as a function of the learning process. Children, rewarded for doing "good" things and punished for doing "bad" things, developed their own sense of proper and improper behavior through the mechanisms of identification and imitation. Mowrer felt that a strict behavioristic analysis of guilt was inadequate to deal with the verbally abstract aspects of the phenomenon of guilt. He suggested that the development of guilt was facilitated when the learner was dependent upon another person who was the source of punishment or discipline. The necessity for punishment in the development of a sense of guilt was also suggested by Fraiberg (1959) and Anna Freud (1968). Both felt that parents should not ignore immoral behavior in their children, as the child

would not develop a sense of guilt and develop a conscience unless his misbehavior were punished. The punishment caused the child to incorporate the parents' values as part of his superego.

Several writers have drawn a close relationship between guilt and fear (Cheyne & Walter, 1970; Sarason, 1966; Switzer, 1968; Unger, 1962). Unger (1962) defined the affective component of guilt as a type of fear.

Unger made some cogent arguments for the proposition that psychological (love-oriented) disciplinary techniques are more potent in guilt training than physical or "materialistic" techniques. . . . He emphasized that psychological techniques of guilt training would work only if there was a background of parental nurturance and a continuing affectionate relationship between parent and child. Love cannot be withdrawn where none exists. (Izard, 1977, p. 43)

The use of punishment to produce guilt in the child is not without hazards. MacKinnon (1938) reported that college students who cheat on tests have been most often physically punished as children. In a study of child-rearing attitudes of subjects who varied in guilt, Schill, Evans, and McGovern (1976) found that high-guilt subjects had incorporated very high standards of proper behavior but were prone to prefer child-rearing practices that emphasized strictness, harsh discipline and rigid adherence to rules. Glueck and Glueck (1950) showed that physical punishment was used more frequently. From their research, Bandura, Ross, and Ross (1963) concluded that children who experienced power assertion

techniques were dependent upon external sanction and fear of punishment in the situation to control their behavior and tend to be more aggressive, generally because of the parents' modeling and reinforcement of such behavior.

Some nonhuman research has been conducted on the phenomenon of guilt. Solomon, Turner, and Lessac (1968) conducted a study with dogs, teaching them to resist the temptation of eating a desirable canned dog food and to eat instead a less desirable dry food. The length of delay between the dogs' yielding and time of punishment was varied. Generalizations from conditioning in dogs to socialization of children should be weighed carefully, but the implication from this study was that if a child has to wait a long time to get the expected punishment, he will have a longer time to experience guilt. Thus guilt would be increased by delay of punishment (Izard, 1977).

The degree of guilt as a function of the sex of the individual has only recently begun to be examined experimentally. Maccoby and Jacklin's (1975) Psychology of Sex Differences has become the major authority for assessing sex differences in behavior. It was found that the sexes did not differ in most aspects of behavior. Hoffman (1977), however, questioned the correctness of their conclusion that females were not more empathetic than males. In sixteen studies of this variable, few differences were significant, but since all differed in the same direction, Hoffman doubted

that the differences could have occurred by chance. Hoffman's (1977) findings contradicted Freud's theory that girls, not compelled to resolve the Oedipus complex quickly, have less internalized moral structures. In Hoffman's projection story-completion test, coded for guilt feelings and fear of punishment, it was found that in all of the child and adult samples, moral transgressions were more likely to be associated with guilt in females and fear in males. It was believed that this difference was due to the use of more induction child-rearing practices for girls and power-assertion practices for boys. This study suggested that females would behave more morally than males in the absence of external sanctions. Clearly, more research is needed in this area.

In discussing the phenomenology of guilt in this chapter the concern has been with moral or ethical guilt, not with guilt in the legal sense. Guilt has been considered as the emotion most essential to the development of the conscience and moral behavior and a necessary concomitant of personal and social responsibility. The intent has been to support the present researcher's belief that there is a need for more study in this area. The literature reviewed revealed a lack of theoretical cohesiveness concerning the antecedents of moral development, suggesting the need for experimental research. Guilt, however, was widely accepted as an index for measuring moral development. The induction method for producing guilt, and thus, moral internalization, is in its conceptual

infancy. No research was located in which this technique was experimentally tested. Correlational research, however, strongly suggested that there might be a link between this child-rearing method and high moral development as indicated by manifest guilt in the child.

Summary

A review of the literature pertaining to moral development and guilt, the index by which it is frequently measured, generated the following general conclusions:

1. According to psychoanalytic theory, the child's moral development is a function of superego strength, the aftermath of successful resolution of the Oedipal Complex. Superego and ego conflict produce guilt. The parent facilitates this process by rewarding the child with love and punishing him by depriving him of love. The child internalizes the parents' standards through identification.

2. The cognitive-developmental approach to moral development assumes that external rules become transformed into internal principles as the child's cognitive processes mature with age. Peer interaction and child-rearing practices provide the child with experiences, but his morality depends on the specific point of his cognitive development.

3. Social learning theory emphasizes teaching moral behavior through modeling and conditioning. The role of the parent in the internalization process is elevated.

4. The induction method of child rearing, in which the parent points out to the child the painful consequences of his act for the parent or for others, was found to be highly correlated with level of moral development in the child.

5. The effectiveness of the induction method was attributed to empathy and optimum arousal of the need for love. Previous arguments that the young child was not capable of empathetic thought have been weakened by recent research.

6. Capacity for guilt has been widely established as a criterion for measuring internalization of moral values.

7. There is a need for experimental research to test the effectiveness of the induction method in fostering moral development in the child.

8. Manifest guilt can be used as an index for measuring level of moral development.

CHAPTER III

METHODOLOGY

The purpose of the present study was two-fold. The primary purpose was to determine the effect of morality content stories on responses to a Guilt Inventory by first-grade children. Because of the lack of a research instrument to measure moral development in first-grade children, such an instrument was developed. It was the intent of the researcher to use a story completion test adapted from the methods of Hoffman and Saltzstein (1967). The methodology is presented in the following order:

1. Development of the Instrument

Procedures for Development of the Instrument

Pilot Study Related to the Development of the
Instrument

Instrument Procedural Discussion

Administration of the Instrument

Scoring of the Guilt Inventory

Analysis of the Instrument

2. The Experiment

Procedure

Selection of Subjects for the Experiment

Selection of Stories for the Experiment

Conducting the Experiment

Analysis of the Experimental Data

Development of the Instrument

The general guidelines for developing the research instrument for measuring moral development in first graders were those used by Hoffman and Saltzstein (1967) in their assessment of moral development in seventh graders. The general guidelines were as follows:

1. Predominantly middle-class boys and girls were used;
2. Test construction (the instrument) consisted of story completion items;
3. Moral development level was measured by assessment of the index guilt. The guilt ratings as adapted from Hoffman and Saltzstein (1967) are contained in Appendix B.

Procedures for Development of the Instrument

A multiple-choice story completion test was developed to assess the intensity and/or presence of morality guilt in first-grade children. The stories were constructed by the researcher in an attempt to adapt Hoffman and Saltzstein's (1967) guilt assessment instrument designed for seventh graders into an instrument that could be used to assess guilt level in first graders. The stories used will hereafter be referred to as the Inventory Stories. The instrument, comprised of the Inventory Stories, will hereafter be referred to as the Moral Development Inventory or M.D.I.

The Moral Development Inventory consisted of four Inventory Stories, adapted by the researcher from those used by

Hoffman and Saltzstein (1967). Each story contained a protagonist the age of the subject to be tested. Two of the stories contained female protagonists and two contained male protagonists. Each story met the criterion for guilt; that is, the wrong deed would be presented so that no external punishment can be involved with the protagonist's action. Each Inventory Story was followed by three multiple-choice endings which corresponded to high, medium, and low levels of guilt as described by Hoffman and Saltzstein (1967). A copy of the criteria is located in Appendix B. Each Inventory Story was read by two first-grade teachers, who were asked to make suggestions concerning the suitability of vocabulary content and comprehension level for first graders. The final draft of the Inventory Stories was the decision of the researcher.

A pilot study was conducted to determine whether the multiple choice endings chosen by the researcher were actual occurring responses in first-grade students and if guilt, as measured by the Moral Development Inventory, appeared in the first-graders' responses on the three levels as defined.

Pilot Study Related to Development of the Moral Development Inventory

Names of first-grade students in Okaloosa and Walton counties and parental permission were obtained through the efforts of Okaloosa-Walton Junior College students enrolled in Child Development or Psychology courses. Okaloosa-Walton

Junior College is located in the central population district, serving Okaloosa and Walton counties in Florida. The college serves a student body of approximately 9,000 students each year. Of this number, approximately 200 students enrolled in the aforementioned courses were asked to locate, within their resident localities, the names and telephone numbers of parents who had first-grade children enrolled in Okaloosa or Walton county schools and who would be willing to be contacted by the researcher for possible participation in the pilot study.

Three hundred sixteen first-grade students (168 males, 148 females) were selected for the pilot study. This constituted approximately a 15% sample of the first-grade students enrolled in Okaloosa and Walton counties and included students enrolled in 82 different first-grade classes. The male-female distribution of the sample was approximately that of the larger population.

All subjects were individually tested by the researcher. The four stories comprising the Moral Development Inventory (see Appendix C) were read, without the multiple-choice endings, to each of the first-grade subjects. Each subject was asked to complete the story by including (a) what he thought happened next in the story and (b) how he thought the protagonist felt about what happened. The interview sessions were tape recorded and then transcribed. From the subject's responses, the researcher compiled a list of all

story endings, eliminating those of obvious duplication and those unrelated to the stories. The researcher expected high similarities in response, as prior studies had suggested that the emotion guilt was not highly differentiated or complex in young children (Izard, 1977; Miller & Swanson, 1960). The responses for each Inventory Story were then arranged in random order. Two professionals in the mental health field ranked each response as (1) low, (2) medium, or (3) high indication of guilt, applying the criteria for each guilt level used by Hoffman and Saltzstein (1967) (Appendix B). The raters were instructed to adhere closely to the Hoffman and Saltzstein (1967) criteria for guilt level ratings. When interrater reliability was satisfied, the responses were inspected by the researcher to ascertain that all three levels of guilt, as defined, were represented and that previously structured multiple-choice responses were actual occurring responses, in agreement with the ratings.

After the multiple-choice M.D.I. met the above criteria, it was further adapted for first-grade use. The Inventory Stories were illustrated in cartoon fashion, with each multiple-choice item a separate cartoon block (see Appendix C), as using a combination of auditory (reading the story to the child) and visual (showing the cartoon sequences) stimuli was believed to increase the child's attentiveness to the test (Watson & Lindgren, 1973).

Instrument Procedural Discussion

The development of the instrument took into consideration the following factors:

1. There may be a tendency for an individual to respond to a multiple-choice test in a preset way, to distort responses in a particular direction more or less regardless of the content of the stimulus (Cronbach, 1970). Lanyon and Goodstein (1971), however, have suggested that the distorting influence has been overemphasized. To determine whether the first-grade child's responses indicated a set-style, the response frequency was analyzed by Chi-square. The null hypothesis was that the subject's response to the multiple-choice test would not be related to any set order preference. Rejection requirement was the .05 level of significance.

2. Can guilt be quantified? Several studies have quantified guilt (Hoffman & Saltzstein, 1967; Miller & Swanson, 1960). Mosher (1966) defended the psychometric foundation of measures of guilt as a necessary condition to success in establishing validity of the construct guilt in research investigations.

3. What testing limitation might one expect when working with the first-grade child? Miller and Swanson (1960) found that since young children do not realize the significance of a testing situation, they do not disguise their thoughts as much as adults. This finding suggested that a young child's responses to the Moral Development Inventory

would probably indicate his true response as opposed to the desire to give the response expected of him. Is the first-grade child cognitively capable of choosing from three variables? Bruner (1973) found symbolic representation highly developed toward the end of the preschool period. The ability of first-grade children to remember three choices and make a decision based on memory has been substantiated by Weir (1962), Reese (1965), and Odom (1967).

Administration of the Instrument

The Moral Development Inventory was administered individually to each subject in four first-grade classes in Walton County, Florida, assigned to this researcher by the Project Director of the Walton County Schools, Florida. The total population included 73 children (33 females, 40 males). The age range was 6 years 1 month to 6 years 11 months. The average age was 6 years 3 months. The subjects were predominantly white, middle-class. Only six children were from minority groups. Students who were repeating the first grade were not included in the population.

To acclimate the subjects to the presence of the researcher and thereby reduce the anxiety they might have toward a stranger, the researcher visited each classroom on two days prior to the administration of the M.D.I., remaining one hour each day. The teacher was requested to introduce the researcher as a college student who is "going to be with us for a few days so that she can learn more about what

first graders are studying these days." The researcher also ate in the cafeteria with the subjects on those days.

On the third day, the teacher was requested to say, "Ms. _____, who has been visiting with us, is writing some stories for first graders. She would like to talk to each one of you individually about these stories. She will be in the next room waiting to talk to you. When I call your name, I would like for you to go into the next room and talk to Ms. _____. When she has finished talking to you, return to this room and another person will be called. When you come back to this room, please do not talk to any of your friends about the stories until Ms. _____ has had a chance to talk to each one of you." The teacher was instructed to continue with the daily classroom activities, interrupted only briefly to instruct the next child to go into the next room. The teacher was advised that no child was to be pressured to leave the room if he felt reluctant to do so.

The Inventory Stories were presented in random order each time the M.D.I. was administered. The three multiple-choice endings were also presented in random order for each story.

The subjects were given the following directions:

I am going to read you a story that is not complete. It needs an ending. Please listen carefully to the story so that you can help me choose the best ending for the story.

The tester simultaneously read the Inventory Story and placed each separate segment of the illustrated M.D.I. before the subject. Before the three multiple-choice endings were presented the subject was told:

I am going to tell you three ways that this story can end. I want you to pick the ending that you think is the best ending for the story. Wait until you have heard all three endings before you choose the best ending.

After the subject had indicated his choice, the tester asked the subject to explain what that picture told about the story to ascertain that the child was aware of the ending he had chosen. The entire procedure was repeated if necessary. There was no time limit requirement placed on the testing procedures.

The tester marked the subject's response on the Scoring Sheet for the M.D.I., a copy of which is contained in Appendix D. After the child had indicated his response, the tester said:

That is fine. Thank you. Now let us go to the next story.

When all four Inventory Stories had been administered, the tester said to the subject:

Thank you for helping me decide the best endings for the stories. Your choices were very good. You may now return to your (classroom, etc.).

The subject's responses to each of the Inventory Stories were weighted and summed. The summed score will be referred to as the M.D.I. Score. The possible numerical range on the M.D.I. was 4-12.

Scoring of the Guilt Inventory

A weighted score of 1 for coded low guilt response, 2 for coded medium guilt response, and 3 for coded high guilt response was given for each multiple-choice ending to the Inventory Story. The scores for the four Inventory Stories were summed. The resulting score was the subject's M.D.I. Score. The possible numerical range of scores was 4-12 inclusive. For labeling and interpretation, the following intervals were used:

- a. scores between 4-6 inclusive = low guilt;
- b. scores between 7-9 inclusive = medium guilt;
- c. scores between 10-12 inclusive = high guilt.

Rating of Students' Overt Reaction to Transgression as Perceived by the Teacher

The teacher was requested to complete a form, a copy of which is in Appendix E. This form, a measure of the child's overt reaction to transgression as perceived by the teacher, was a replication from the Hoffman and Saltzstein (1967) study. The answers were indicators of low, medium, and high levels of guilt, with the following quantifications:

low = 1

medium = 2

high = 3

The numerical range possible score for each subject on the teacher form was 3-9 inclusive.

Analysis of the Instrument

In discussing the problems of construct validity, Cronbach (1972) suggested using operationally different indicators thought to measure the same trait. The validity of a construct is increased if such indicators agree. The subject's score on the M.D.I. was correlated with the respective rating of the subject's overt reaction to transgression as perceived by the teacher. The Pearson product-moment correlation coefficient analysis was used and the table of critical values for this statistic consulted to determine whether the relationship was significant at the .05 level. A similar procedure was used to analyze the relationship between the student's score on the M.D.I. and his language score on the Comprehensive Tests of Basic Skills. No significant relationship between the student's M.D.I. score and his language score would strongly suggest that the M.D.I. was measuring something other than language skills (Cronbach, 1970). Hoffman (1975b) noted that guilt level did not appear to be related to IQ or language abilities. Miller and Swanson (1960) also found language skills unrelated to indices of moral development. The null hypothesis was thus assumed: there will be no significant relationship between subject's score on the Moral Development Inventory and his score on the language portion of the Comprehensive Tests of Basic Skills.

The Comprehensive Tests of Basic Skills tests are given to all first-grade students in the Okaloosa-Walton county

schools. These scores were obtained from the students' cumulative folders. The complete reference for the Comprehensive Tests of Basic Skills is contained in the Bibliography.

The M.D.I. was subjected to two tests of reliability: split-half and test-retest. The Pearson product moment correlation coefficient was used for analyzing the relationships. The optimum criterion expected was a minimum r value of .80. The M.D.I. was additionally analyzed for homogeneity of variance by the F test. The null hypothesis was that there is no difference between the individual stories that comprise the M.D.I. in guilt-level score. Rejection requirement was the .05 significance level. The results were discussed.

The Experiment

Problem: Can level of moral development, as measured by the developed Moral Development Inventory, be increased in first graders through exposure to morality content stories?

Four first-grade classes in Walton County were assigned to the researcher by the Project Director of the Walton County, Florida, school system. These first-grade classes were not the same classes used to analyze the M.D.I. Two of the classes were randomly assigned to the experimental conditions and two to the control conditions (Campbell & Stanley, 1963). The experimental group was read morality content stories, one each day for five consecutive days. The control group was read five stories without morality content in like

manner. The subjects in the two groups were individually administered the Moral Development Inventory on the fifth day after the reading of the fifth story. The difference in the responses to the Moral Development Inventory between the experimental and control groups was analyzed.

Procedures

Selection of subjects for the experiment. The subjects for the experiment were demographically similar to the subjects the researcher used to validate the instrument. The original number of subjects was 82 (43 males, 39 females). The number of subjects used in the final experiment data analysis was 54 (25 males, 29 females). The subject elimination reasons will be discussed. The age range was 6 years 6 months to 7 years 3 months. The average age was 6 years 10 months.

First-grade subjects were used in an attempt to adapt for first grade use the Guilt Inventory used by Hoffman and Saltzstein (1967) with seventh-grade students. The intent in the present study was to choose subjects from an area similar in socio-economic background to those used in the Hoffman and Saltzstein (1967) study. The subjects in the Hoffman and Saltzstein (1967) study were predominantly middle-class. The profile of the Walton counties is also middle-class as determined by the 1970 census.

Selection of stories for the experiment. All stories used were chosen from the Okaloosa and Walton county school

libraries. The researcher chose five stories that illustrated the morality induction process as defined by Hoffman and Saltzstein (1967) and five stories that were without the morality induction process. The ten stories were read by two professionals in the mental health field and one librarian to determine whether they agreed that this criterion had been met. A deliberate attempt was made by the researcher to choose stories comparable in length and of general interest. Additionally, two first-grade teachers read the stories to determine if they were suitable for use in the first grade. A list of the stories used has been placed in Appendix F. These stories will be referred to as the Treatment Stories.

Conducting the Experiment

The classes that comprised the experimental group were read one morality story each day for five consecutive days. The classes that comprised the control group were read one nonmorality story each day for five consecutive days. The order of the stories read each day was random. The following standard procedures were followed:

1. The researcher asked the classroom teacher to introduce her to the students as a college student who wanted to receive some practice in reading stories to first-grade students.

2. The teacher was asked to remain in the room while the stories were being read. If there had been a discipline problem that the researcher believed to be disturbing to the

rest of the class, the researcher would have nodded to the teacher, who would then quietly remove that student from the class. No students were removed for discipline reasons. Other discipline problems, not deemed by the researcher to be sufficiently disturbing to require removal of the student, were ignored.

3. The researcher gave the following instructions to the students:

I am going to read you a story. I want you to listen very carefully, because I am going to ask you some questions about the story when I have finished reading it to you. Please do not interrupt or ask any questions.

4. The story was then read.

5. A short true-false test was given to all students to determine whether they had been listening. A copy of the test for each story and the type of answer sheet used are contained in Appendix G. Any student who missed more than one question on any one story was eliminated from the population to be analyzed.

6. The teacher was asked to help give out and take up the test answer sheets, making sure the students' first names were on each sheet.

7. The teacher was requested not to discuss the stories after the researcher left the room and to discourage the students from talking about the stories.

8. After the reading of the first story and the testing for listening, the researcher situated herself in a room or

hallway near the classroom. The teacher was requested to send out, one at a time, approximately one-fourth of the students in the class. The order was determined by the teacher. By the end of the fourth day, the researcher had talked individually to each student in the classroom.

9. The researcher asked each student three questions:
 - a. How long have you lived in this town?
 - b. What is your favorite color?
 - c. What do you want to do when you grow up?

The researcher wrote this down, but the information was irrelevant to the experiment. The purpose of this procedure was to acclimate the student to the researcher in a simulated testing situation.

10. After the reading of the story on the fifth day, the researcher administered the Moral Development Inventory individually to each student. Both the Inventory Stories and the multiple-choice answers were presented in random order. Detailed procedures for administering the Moral Development Inventory are included in the first part of the Methodology chapter. No student was pressured into participating in the experiment. Students who missed more than one answer on the daily tests were not notified of this. All students were administered the Moral Development Inventory, but only those who met the listening criteria were included in the analysis of the data.

11. Procedures were identical for experimental and control group classes.

Analysis of the Experimental Data

To determine whether story treatment (morality content stories and nonmorality content stories) had an effect on the moral development level of the subject as measured by the Moral Development Inventory, a 3-way ANOVA was conducted. The three factors were:

- a. type of story at two levels
- b. sex at two levels
- c. part of the instrument at four levels

Of interest were main treatment effect and whether the effect of the type of treatment story depended upon sex, group, or upon the part of the instrument. The .05 level of significance was chosen as appropriate for rejection of the null hypothesis. Treatment effect and the various interactions were discussed.

CHAPTER IV

RESULTS

The results from the analyses of data are reported in two parts. Part one deals with the development of the instrument, the Moral Development Inventory (M.D.I.), which was purported to measure level of moral development in first-grade children. The results of the univariate 3-factor analysis of variance, detailing the main effects of the experiment and the interactions, are presented in the second part.

The Instrument

A pictorial multiple-choice story completion test was developed to assess the intensity and/or presence of morality guilt in first-grade children. This instrument, hereafter referred to as the Moral Development Inventory (or M.D.I.), contained four Inventory Stories. Following each Inventory Story three multiple-choice endings were given. These endings corresponded to high, medium, and low levels of guilt as described by Hoffman and Saltzstein (1967). The M.D.I. was read independently by two first-grade teachers in Walton County, Florida. Both teachers rated the stories suitable in comprehension level and vocabulary for use with first-grade children. To confirm that the multiple-choice endings chosen

by the researcher were actual occurring responses in first-grade children, a pilot study was run.

Pilot Study

Three hundred sixteen first-grade students (168 males, 148 females) responded to the four Inventory Stories. The stories were presented without the multiple-choice endings so that the subjects could complete the stories with open endings. Of the 168 males and 148 females who were administered the M.D.I., the responses of 92 males and 78 females were eliminated by the researcher because they did not relate to the stories. The responses from the remaining 76 males and 70 females were compiled, eliminating duplications, and arranged in random order for each of the four Inventory Stories.

The compilations and eliminations of responses resulted in 7 possible responses for Story 1, 5 for Story 2, 5 for Story 3, and 6 for Story 4. The children's responses were inspected by the researcher and it was noted that all three levels of guilt, as defined, were represented and that the previously structured multiple-choice responses were actual occurring responses.

A psychiatric social worker and a clinical psychologist independently rated the responses to each Inventory Story as low (1), medium (2), or high (3) indication of guilt, using the criteria for guilt level suggested by Hoffman and Saltzstein (1967) (see Appendix B). The response distribution by level of guilt was as follows: 22% high, 52% medium,

and 26% low. An interrater reliability of $r = .85$ was obtained, using the Pearson product-moment correlation method. The criterion previously selected for interrater reliability was thus satisfied.

The pilot study confirmed that the multiple-choice endings to the Inventory Stories were actual occurring responses among first graders and that the level of guilt the researcher assigned each response correlated highly with that of the two raters. The two raters disagreed on only two of the multiple-choice guilt ratings previously chosen by the researcher. On Inventory Story 3, rater I marked choice B as medium and choice C as high. Rater II marked B as high and C as medium. The disagreement was later resolved by the two raters consulting Hoffman and Saltzstein's (1967) criteria for guilt level and they agreed that B was medium and C was considered high level of guilt. The pilot study offered no justification for changes in the instrument.

Analysis of the Instrument

The M.D.I. was administered to 73 first graders (33 females and 40 males) in the Walton County, Florida, schools. None of these subjects had been in the Pilot Study. A score for each subject was obtained. The possible numerical range in score was 4 (low) = 12 (high) inclusive. The actual occurring range was 4 = 11. For a summary of the responses to the M.D.I. see Table 1.

Table 1
Summary Table of the Ss Total Scores
on the M.D.I.

	Total Ss	Male Ss	Female Ss
Number	73	40	33
Range	4 - 11	5 - 11	4 - 11
Mean	7.75	7.78	7.63

Validity

If a subject's responses to a test were a function of testing set, validity of the test score was adversely affected. To determine whether there was a tendency for the first graders to respond to a multiple-choice test in a preset way, the response order frequency was analyzed by Chi-square. The Chi-square data are presented in Tables 2 and 3.

The null hypothesis was accepted. There was no significant relationship between subject choice and order of presentation of three choices. The subjects demonstrated no preference for first-, second-, or third-choice endings presented.

The M.D.I. proposes to measure level of guilt in first-grade children as an indicator of their level of moral development. In order to increase validity of this measurement of moral development, the subjects' scores on the M.D.I. were correlated with a respective rating of the same subjects' overt reaction to transgression as perceived by the teacher. A copy of the Teacher Rating Form (adapted from Hoffman & Saltzstein, 1967) is shown in Appendix E. The possible score range on the Teacher Rating Form (T.R.F.) was 3 (low) = 9 (high) inclusive. The actual occurring score range was 3-9. The Pearson product-moment correlation coefficient yielded an r value of .37, $p < .001$. The null hypothesis of no relationship between teacher rating of subjects' moral development and subjects' M.D.I. score was rejected.

Table 2

Number of Subjects Who Responded to First, Second,
or Third Choice Order on Each of the
Four Inventory Stories

Inventory Stories	Total Ss Who Responded to First Choice	Total Ss Who Responded to Second Choice	Total Ss Who Responded to Third Choice
1	28	28	17
2	31	22	20
3	18	32	23
4	22	29	22

N = 73

Table 3

Chi-square Values for Order of Subjects' Responses
to First, Second, and Third Choices on Each
of Four Inventory Stories

Inventory Stories	x^2 Values for Ss Responding to First Choice	x^2 Values for Ss Responding to Second Choice	x^2 Values for Ss Responding to Third Choice
1	.42	.002	.59
2	1.57	1.19	.01
3	1.84	.65	.30
4	.30	.06	.11

$x^2 (2, 3) = 7.04$

not significant

The subjects' scores on the M.D.I. were correlated with the subjects' scores on the language portion of the Comprehensive Tests of Basic Skills to determine whether the M.D.I. measured something other than language skills. The Pearson product-moment correlation yielded an r value of .32, $p < .01$. The null hypothesis that there would be no significant relationship between the subjects' scores on the M.D.I. and the language portion of the Comprehensive Tests of Basic Skills was rejected. The implications of this significant relationship are discussed in the next chapter.

Reliability

The usefulness of any test is directly proportional to its ability to measure a stated variable or variables with consistency. The correlation between scores on the odd and even sets is a split-half correlation, from which a reliability coefficient for the entire test may be estimated. The Pearson product-moment correlation yielded a coefficient of $r = .46$, $p < .001$. In order to compute the correlation, the test was divided in half. A correction value for the whole test was computed, yielding $r = .63$. This correlation was below the desired criterion value of $r = .80$. The F max test for homogeneity of variances indicated that there was no significant difference between the Inventory Story with the largest variance and the Inventory Story with the smallest variance (F max = 1.36, $df = 73$, $k = 4$, $p > .05$). After considering the brevity factor of the test (only four

Inventory Stories) and the nonsignificant F_{\max} Statistic, the researcher decided to accept the split-half correlation coefficient value of $r = .63$ as reasonable and no changes were made in the instrument.

To determine whether the M.D.I. yielded the same or similar results if administered again, the test was readministered to 20 of the subjects one week after the first administration. The Pearson product-moment correlation yielded a value of $r = .82$ in this test re-test evaluation of reliability. This correlation was a strong indication of test stability.

Experiment

Eighty-two first-grade students, enrolled in four different classrooms, were the subjects in this experiment. Methodological controls resulted in the loss of 28 subjects. The subjects used in the experimental analysis totaled 54. Of these, 27 (12 males, 15 females) were in the experimental treatment group and 27 (13 males, 14 females) were in the control treatment group. For a summary of subject information, see Table 4.

To determine the effect of sex, group assignment, and each Inventory Story on the children's M.D.I. scores, a three-way (2 sex x 2 group x 4 stories) univariate analysis of variance was performed on the subjects' scores on each Inventory Story (range = 1-3) (Table 5). This analysis indicated that there was a significant effect for sex

Table 4

Number of Subjects by Classroom, Treatment Condition,
Sex, and Removal Method

	Classroom I	Classroom II	Classroom III	Classroom IV	Totals
	Experimental Group	Experimental Group	Control Group	Control Group	
Male Ss	11	13	9	11	44
Female Ss	9	10	9	10	38
Male: Absentee Ss Removed	3	2	2	2	9
Female: Absentee Ss Removed	2	1	2	2	7
Male: Listening Criteria, Ss Removed	2	5	1	2	10
Female: Listening Criteria, Ss Removed	0	1	1	0	2
Total Male: Ss Remaining	6	6	6	7	25
Total Female: Ss Remaining	7	8	6	8	29

(a between-subjects factor), $F(1, 50) = 4.59, p < .04$; group (a between-subjects factor), $F(1, 50) = 17.28, p < .0001$; and the sex by group interaction, $F(1, 50) = 8.17, p < .006$ (see Table 5).

This analysis also indicated no significant difference for the presentation of stories (a within-subjects factor), $F(3, 150) = 2.30, p < .08$. There was also no significant story by sex interaction, $F(3, 150) = .54, p < .66$; nor was there a significant story by group interaction, $F(3, 150) = 1.64, p < .18$. There was a significant three-way (sex x troupe x story) interaction, $F(3, 150) = 2.73, p < .05$.

The experimental subjects' total M.D.I. mean scores ($M = 8.91$) were significantly higher than the total M.D.I. scores of the control subjects ($M = 7.29$), indicating a strong treatment effect. Further analysis of the data by summing the total M.D.I. scores by sex across both the experimental and control groups indicated that females ($M = 8.55$) had significantly higher morality scores than males ($M = 7.64$). The possible total score range on the M.D.I. was 4-12.

The average morality scores on each separate Inventory Story within the M.D.I. (score range = 1-3) were: first story ($M = 2.03$); second story ($M = 2.02$); third story ($M = 1.87$); and fourth story ($M = 2.20$). There was no significant difference in the four Inventory Story means (Table 5). For a summary of the total M.D.I. mean scores and Inventory Story mean scores, see Table 6.

Table 5

Univariate Analysis of Variance on the Effects
of Sex, Group Assignment, and Stories
on Average Morality Score

Source	df	MS	F
Sex	1	2.4	4.59*
Group	1	9.06	17.28*
Sex x Group	1	4.28	8.17*
S (Group x Sex)	50	.52	
Story	3	1.00	2.30
Story x Sex	3	.24	.54
Story x Group	3	.72	1.64
Story x Group x Sex	3	1.19	2.73*
Story x S (Group x Sex)	150	.44	

* $\underline{p} < .05$

Table 6

Summary Table of Total M.D.I. Means (R = 4-12)
and Inventory Story Means (R = 1-3)

	Total Ss	Experimental Male Ss	Control Male Ss	Experimental Female Ss	Control Female Ss
Number	54	12	13	15	14
M.D.I. Means	8.11	9.08	6.31	8.80	8.28
Inventory Story 1 Means	2.03	2.00	1.77	2.13	2.21
Inventory Story 2 Means	2.01	2.25	1.77	2.33	1.71
Inventory Story 3 Means	1.87	2.25	1.23	2.07	1.93
Inventory Story 4 Means	2.20	2.58	1.54	2.27	2.43

Presented in Table 7 are the results of Newman-Keuls post hoc comparisons on the triple interaction which indicated that Experimental Males had significantly higher morality scores on Story 4 than on Story 1, Story 2, or Story 3. There was no difference among the Experimental Males' scores on Story 1, Story 2, or Story 3.

Results of Newman-Keuls post hoc comparisons on the triple interaction indicated that Control Males had significantly lower morality scores on Story 3 than on any of the other stories. There were no differences among the Control Males' morality scores on Story 1, Story 2, or Story 4. These results are presented in Table 8.

Results of Newman-Keuls post hoc comparisons on the triple interaction indicated that there were no significant differences among the experimental female subjects' morality scores on any of the four stories (Table 9).

Results of Newman-Keuls post hoc comparisons on the triple interaction indicated that Control Females had significantly lower morality scores on Story 2 and Story 3 than they did on Story 4. They also had significantly lower morality scores on Story 2 than on Story 1. There were no differences between their scores on Story 2 and Story 3, Story 1 and Story 3 or Story 1 and Story 4 (Table 10).

Results of Newman-Keuls post hoc comparisons on the triple interaction indicated that on Story 1, males assigned to the control group had significantly lower morality scores

Table 7

Summary Table for Newman-Keuls Post Hoc Comparisons on
Average Morality Score for Experimental Male Subjects
for the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Story 4	Story 3	Story 2	Story 1	r	C.V. at
	2.58	2.25	2.25	2.0		.05
2.58		.33*	.33*	.58*	4	.32
2.25			-	.25	3	.29
2.25				.25	2	.26
2.0						

* $p < .05$

Table 8

Summary Table for Newman-Keuls Post Hoc Comparisons on
Average Morality Score for Control Male Subjects
for the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Story 1	Story 2	Story 4	Story 3	r	C.V. at
	1.77	1.77	1.54	1.23		.05
1.77	-	-	.23	.54*	4	.32
1.77		-	.23	.54*	3	.29
1.54				.31*	2	.26
1.23						

* $p < .05$

Table 9

Summary Table for Newman-Keuls Post Hoc Comparisons on
Average Morality Score for Experimental Female Subjects
for the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Story 2	Story 4	Story 1	Story 3	r	C.V. at
	2.33	2.27	2.13	2.07		.05
2.33	-	.06	.20	.26	4	.32
2.27		-	.14	.20	3	.29
2.13			-	.06	2	.26
2.07						

Table 10

Summary Table for Newman-Keuls Post Hoc Comparisons on
Average Morality Score for Control Female Subjects
for the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Story 4	Story 1	Story 3	Story 2	r	C.V. at
	2.43	2.21	1.93	1.71		.05
2.43	-	.22	.50*	.72*	4	.32
2.21		-	.28	.50*	3	.29
1.93			-	.22	2	.26
1.71				-		

* $p < .05$

than female subjects assigned to either the control or experimental group. Otherwise there were no significant differences between average morality scores on Story 1 (Table 11).

Results of Newman-Keuls post hoc comparisons on the triple interaction indicated that on Story 2, Control Males and Control Females had significantly lower morality scores than Experimental Males or Experimental Females. There was no significant difference between the Control Males and Control Females or between the Experimental Males and Experimental Females in terms of morality scores on Story 2 (Table 12).

Results of Newman-Keuls post hoc comparisons on the triple interaction indicated that on Story 3, males assigned to the control group had significantly lower morality scores than all other subjects. In addition, females assigned to the control group had significantly lower scores on Story 3 than males assigned to the experimental group. There was no difference between morality scores of Experimental Males and Experimental Females or between Experimental Females and Control Females on Story 3 (Table 13).

Results of Newman-Keuls post hoc comparisons on the triple interaction indicated that on Story 4, Control Males had lower average morality scores than Control Females, Experimental Females, and Experimental Males. Experimental Females had lower morality scores on Story 4 than Experimental Males. There was no difference between the scores

Table 11

Summary Table for Newman-Keuls Post Hoc Comparisons
on Average Morality Score on Story 1 for
the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Control Females	Experimental Females	Experimental Males	Control Males	r	C.V. at
	2.21	2.13	2.00	1.77		.05
2.21	-	.08	.21	.44*	2	.26
2.13		-	.13	.46*	2	.26
2.00			-	.23	2	.26
1.77				-		

* $p < .05$

Table 12

Summary Table for Newman-Keuls Post Hoc Comparisons
on Average Morality Score on Story 2 for
the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Experimental Females	Experimental Males	Control Males	Control Females	r	C.V. at
	2.33	2.25	1.77	1.71		.05
2.33	-	.08	.56*	.62*	2	.26
2.25		-	.48*	.54*	2	.26
1.77			-	.06	2	.26
1.71						

* $p < .05$

Table 13

Summary Table for Newman-Keuls Post Hoc Comparisons
on Average Morality Score on Story 3 for
the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Experimental Males	Experimental Females	Control Females	Control Males	r	C.V. at
	2.25	2.06	1.93	1.23		.05
2.25	-	.19	.32*	1.02*	2	.26
2.06		-	.13	.83*	2	.26
1.93			-	.70*	2	.26
1.23				-		

* $p < .05$

of the Experimental and Control Females or between the Experimental Males and Control Females on Story 4 (Table 14).

Results of Newman-Keuls post hoc comparisons on the double interaction (sex x group) (Table 15) indicated that across all four stories, Control Male Subjects had lower morality scores than all other subjects. There were no differences in morality scores among the Experimental Females, Experimental Males, and Control Female subjects. The average morality score for Control Males was 1.58; for Experimental Males, 2.27; Control Females, 2.07; and Experimental Females, 2.20.

Table 14

Summary Table for Newman-Keuls Post Hoc Comparisons
on Average Morality Score on Story 4 for
the Group x Sex x Story Interaction

<u>Average Morality Score</u>						
	Experimental Males	Control Females	Experimental Females	Control Males	r	C.V. at
	2.58	2.43	2.26	1.54		.05
2.58 v	-	.15	.32*	1.04*	2	.26
2.43		-	.17	.89*	2	.26
2.26			-	.72*	2	.26
1.54				-		

* $p < .05$

Table 15

Summary Table for Newman-Keuls Post Hoc Comparisons
on Average Morality Score for the
Group x Sex Interaction

<u>Average Morality Score</u>						
	Experimental Males	Experimental Females	Control Females	Control Males	r	C.V. at
	2.27	2.20	2.07	1.58		.05
2.27	-	.07	.20	.69*	4	.34
2.20		-	.13	.62*	3	.31
2.07			-	.49*	2	.26
1.58				-		

* $p < .05$

CHAPTER V
DISCUSSION

The Moral Development Inventory (M.D.I.) was developed by the researcher in an attempt to adapt Hoffman and Saltzstein's (1967) Guilt Inventory (suitable for seventh graders) into an instrument that would measure level of moral development in first graders. The format of the resultant M.D.I. differed considerably from the Hoffman and Saltzstein (1967) instrument, which consisted of an open-ended story completion test scored by only Drs. Hoffman and Saltzstein. Their scoring process was contaminated by examiner bias, as not only were item statements scored, but also "tone" of the subject's story. Hoffman and Saltzstein (1967) achieved high interrater reliability only through repeated conferences. The intent of those researchers, however, was not to develop an instrument that would be used routinely to measure guilt, but rather to determine, through their close and dedicated work, whether students who were judged high in moral development (high guilt score) were the products of a particular child-rearing technique. Their research confirmed that there was a positive correlation between high guilt score and the induction method of child discipline.

The present experimenter accepted certain basic assumptions from the Hoffman and Saltzstein (1967) research:

1. The levels of guilt as defined by Hoffman and Saltzstein (1967) were reasonable descriptions of low, medium, and high levels of guilt.
2. Their methods of assigning child-rearing technique labels to parents were reliable and valid.

Therefore, in the development of the M.D.I., there was a deliberate attempt to reconcile story-completion multiple-choice answers with the three levels of guilt defined by Hoffman and Saltzstein (1967) and to score expressions of guilt in a manner similar to theirs. A deliberate attempt was made to model the M.D.I. stories after those used by Hoffman and Saltzstein (1967), i.e., stories concerned with situations of conscience-guilt, stories in which the protagonist was the age of the subject, and story situations that could definitely occur in the life experiences of the subjects being tested. The Hoffman and Saltzstein (1967) instrument allowed for open-ended story completion, whereas the M.D.I. was structured with three multiple-choice endings. The M.D.I. was further adapted for use with first graders by simplifying the vocabulary and illustrating the stories in cartoon fashion.

The decision to structure the answers into a multiple-choice test by the present researcher was based on the belief that an open-ended test for first graders would yield much

useless data, and that reliability in scoring would be difficult to achieve. The pilot study more than confirmed this.

Pilot Study

Of the 168 males and 148 females who were administered the M.D.I. with open endings, the responses of 92 males and 78 females were eliminated by the researcher because they did not relate to the stories. Responses that were useless for the purpose of this study fell into several categories:

1. Subjects who would not talk. No pressure was placed upon any subject. Such subjects who would not talk were thanked for listening and dismissed. There were only five such subjects.

2. Subjects who said "I don't know." It was the researcher's belief that most of these subjects, given no guidelines for story endings, were reluctant to venture an ending that might be wrong.

3. Subjects who were eager to talk--about everything except the story endings.

4. Subjects who centered on some action or event in the story other than the completion of the story. Two examples of such responses were:

- a. to Story 2--"You should always buy Coca-Cola in the cans."
- b. to Story 4--"If you eat too many potatoes you'll throw up."

5. Subjects whose responses indicated that "your sins will find you out." Although each story was structured so that no external punishment could be involved with the protagonist's actions--he could always escape detection--many subjects ignored this entirely and recommended immediate punishment. One male subject was particularly vehement about Story 3: "That's cheating! His mother should beat the hell out of him!"

The open-ended responses which could be coded into levels of guilt were highly similar, yielding a total of 16 different story endings: 7 possible responses for Story 1, 5 for Story 2, 5 for Story 3, and 6 for Story 4. This suggested that guilt may not be a highly differentiated emotion in first-grade children, but since all three levels of guilt, as defined by Hoffman and Saltzstein (1967) were represented in the responses, the Pilot Study served its purpose. The interrater reliability ($r = .85$) between the two mental health professionals who rated all the responses into high, medium, or low levels of guilt suggested that the guilt levels descriptions (Appendix B) allowed discrete guilt level decisions to be made with considerable agreement. This finding added strength to the researcher's decision that those levels of guilt were reasonable descriptions of high, medium, and low levels of guilt. Because the results from the pilot study did not warrant any changes in the M.D.I. it was considered ready to use.

Validity of the Instrument

Test technology is seldom refined enough to meet all demands. One of the demands is for assessment of performance. In a multiple-choice test, such as the M.D.I., it was necessary to determine whether the choice was a function of decision or set. "Response sets dilute a test with factors not intended to form part of the test content, and to reduce its logical validity" (Cronbach, 1950, p. 145). Three types of response set were considered. The response set, to work for speed rather than accuracy, was controlled by the testing procedure. No time limit was stipulated and the subject was allowed to "ponder" at leisure. The subject was not promised any reward for finishing the test; therefore, he did not rush toward finishing in order to receive certain benefits. The response set of responding just to respond was controlled by the experimenter's asking each subject to tell her what the story ending was after the subject pointed toward one of the three pictorial choices. If the subject could not tell the experimenter the content of the pictorial choice, the entire test item was repeated. No subject needed the story repeated more than once. The Chi-square value for order of subject's responses to first, second, and third choice, $\chi^2 (2, 3) = 7.04, p > .05$, indicated that there was no significant relationship between subject's choice and order of presentation of three choices. According to Cronbach (1950) the only major form of fixed alternative test

which has been found relatively free from response sets is the multiple-choice item. The present study also confirmed this. The experimenter was satisfied that response set did not act as a significant source of error in measurement on the M.D.I.

To establish construct validity, the subjects' scores on the M.D.I. were correlated with respective ratings of the same subjects' overt reaction to transgression as perceived by the teacher. The Pearson product-moment correlation yielded an r value of .37, $p < .001$. Although this relationship was highly significant, it was not a strong relationship for predictive purposes. Low correlations between different measures of a construct are the rule rather than the exception (Mischel, 1968). Further measures of the construct moral development would be advisable. Suggestions include parent ratings, peer ratings, and direct behavioral assessment of the subject in a moral-dilemma situation. Tests of natural observations often yield additional data; however, more than any other procedure, they represent a direct assessment of criterion behavior. In the final analysis, what a person does is more important than what he says he will do. The value of a paper and pen test is its economical time utility. If such a test can closely approximate the actual behavior, it serves the utility purpose. Assuming the subjects being tested identified with the protagonist, this test is similar to a check-list which contains both behaviors and feelings.

The present level of validity suggested that while the M.D.I. may not be sufficiently sensitive to use for individual diagnosis, it may be considered useful for assessing group directional change. This will be discussed further in relation to the experiment.

The positive significant correlation between the M.D.I. scores and the subjects' language grade level scores from the Comprehensive Tests of Basic Skills suggested that the test had a language component. The test is a verbal instrument, requiring language comprehension. Although the children's responses required only that they point to item A, B, or C of the pictorial M.D.I., they had to understand the verbal directions and the story theme. Modest correlations with ability to communicate (.24-.40) are to be expected with tests that require social insight (Lanyon & Goodstein, 1971). Nevertheless, the data considered in total suggested that something was being measured other than just language comprehension. Several subjects who scored high on the M.D.I. had very low language grade-level scores. The highly significant difference in scores on the M.D.I. which resulted from a one-week treatment effect was unlikely to be related to a sudden increase in language grade level score or language comprehension improvement, as the control subjects were exposed to an equal language treatment.

Reliability of the Instrument

The test-retest reliability coefficient, $r = .82$, was a strong indication of the M.D.I.'s stability. The time interval was one week. There is a possibility that memory of the individual Inventory Stories was a factor in the high test-retest score. Turiel (1966) discussed the testing problems of using story situations and suggested that the subjects might be responding to the memory of their prior responses. The test-retest coefficient, however, was higher than the majority of reported similar test-retest coefficients in the more than 300 measures of child behavior and development reported in Tests and Measurement in Child Development: A Handbook (Johnson & Bonmarito, 1971). Further information could have been provided by retesting another sample of the original 73 subjects after two weeks. This was not possible, because of the end of the school term.

A Pearson product-moment correlation of $\underline{r} = .63$ on a split-half reliability test is quite high considering that there were only 4 items on the M.D.I. Reliability increases with the length of a test, and a reliability coefficient of $\underline{r} = .70$ is considered a good indication of reliability (Cronbach, 1972). The most obvious solution to increase the reliability of the M.D.I. would be to add more items. The attention span of the first grader, however, is a definite factor to consider. The experimenter felt that the average time of 20 minutes to administer the M.D.I. individually was

already straining the limits of the first grader's attention span. Dividing the testing session into two or more testing periods would reduce internal validity as the researcher would have little or no control over such variables as historical events or maturation between testing sessions.

Strength was added to the split-half reliability measurement by considering it in conjunction with the nonsignificant F max test for homogeneity of variances; however, whatever aspect of personality functioning one is measuring, it is important to include an adequate sample of relevant situations (Goldfried & Sprafkin, 1974). It is questionable whether four item-situations were an adequate sample. This matter will be discussed more fully under Experiment.

The following conclusions were drawn from the data on the development of the M.D.I.:

1. The null hypothesis that the subject's response to the multiple-choice test would not be related to any set order preference was accepted.

2. The subject's M.D.I. score was significantly related to the respective rating of the subject's overt reaction to transgression as perceived by the teacher.

3. The subject's M.D.I. score was also positively related to his grade-level score on the language portion of the Comprehensive Tests of Basic Skills.

4. The test-retest reliability measure surpassed the previously stated criterion for acceptance ($r = .82$).

5. The split-half reliability measure, $r = .63$, was lower than preferred, but was accepted as reasonable considering that the test only had 4 items and that the F max test for homogeneity of variance was nonsignificant.

Experiment

The present study looked at issues related to moral development and attempted to determine whether level of moral development, as measured by the M.D.I., could be increased in first-grade subjects through exposure to morality content stories. A three-way univariate analysis of variance was used to analyze the data. The factorial design allowed for studying of the effects on the dependent variable (the M.D.I. score) of three independent variables (sex, group, and story). The design also made it possible to determine whether the three independent variables had an interaction effect on the children's responses to the stories. Additionally, the design permitted the researcher further analysis of the M.D.I. by determining whether there was a differential treatment effect on each of the four Inventory Stories that comprised the M.D.I.

Although the original number of subjects in the experiment was 82, various controls to strengthen the effects of the independent variable (type of treatment) on the dependent variable (M.D.I. score) resulted in 27 subjects (13 males, 14 females) in the experimental treatment group and 27 subjects (12 males, 15 females) in the control treatment group.

The null hypothesis, that there would be no difference in level of moral development as measured by the M.D.I. in the experimental group of first graders read morality content stories and the control group read stories without morality content, was rejected, $F(1, 50) = 17.28, p < .0001$. This indicated a very large between-subjects treatment effect, which was somewhat surprising. The immediacy effect may not have remained stable, and further research is necessary to determine whether the treatment effect was stable over time, but the implications are promising. Hoffman and Saltzstein's (1967) position that the induction method of child-rearing was positively related to moral development in the child would appear strengthened.

The strong significant treatment effect may have been related to lack of complete randomness in the design. The research took place in a classroom situation in which the independent variable was manipulated under as carefully controlled conditions as the situation would permit. However, the individual classes were randomly assigned to experimental and control conditions, but the individual subjects were not. Additionally, there was no pre-test to determine whether the subjects differed in level of moral development before the treatment. This weakness in design, however, was counter-balanced by the magnitude of the treatment effect.

The ANOVA also indicated no significant difference for the presentation of stories, $F(3, 150) = 2.30, p < .08$.

There was no differential treatment effect among the four Inventory Stories, which meant that each story was equally effective in measuring the treatment effect. This was in agreement with the reliability data from the development of the M.D.I., i.e., high split-half reliability ($r = .63$), and homogeneity of variances for the 4 Inventory Stories, $F_{\max} = 1.36$ ($df = 73, k = 4$), $p > .05$.

The significant effect for sex resulted in rejection of the null hypothesis that there would be no difference in level of moral development as a function of sex of subject, $F(1, 50) = 4.59$, $p < .04$. The sex by group interaction was also significant, $F(1, 50) = 8.17$, $p < .006$. The difference in the M.D.I. mean scores of the control males ($M = 6.3$) and the control females ($M = 8.3$) is in agreement with Hoffman's (1977) finding that females had more highly internalized moral structures. Hoffman attributed the difference to the use of more induction child-rearing practices for girls and power-assertion for boys. If his assumption is correct, it may offer an explanation for the tremendous difference in scores between the control males as measured by the M.D.I. and the experimental males ($M = 6.30 - M = 9.08$) and the lesser difference in scores of control and experimental females ($M = 8.28 - M = 8.80$). It is certainly suggestive of differential child-rearing practices for males and females, but supportive of the effectiveness of the induction method for both. Although the treatment effect was significant

for males and not for females, the experimental females' mean scores were higher than the control female mean scores.

While the lower mean scores of the control males was anticipated, the failure to find a significant treatment effect for females was puzzling. Either the treatment (five morality-content stories) did not work for females, or the M.D.I. was not sufficiently sensitive to detect differences that the treatment caused. It was expected that the story treatment, with its strong language component, would be more effective with females. Maccoby and Jacklin (1974) noted that although females have greater verbal ability than males, these differences do not emerge until about age eleven. Perhaps the language component would create greater differences in effect with older children. It is more probable that the females were already acclimated to the induction process (Hoffman, 1977), thus failing to show a significant change in moral development level with a five-day treatment. The initial superior effect of the treatment on males may have leveled under a longer treatment condition.

The difference in the male-female scores could also have been related to the specific sample used. In the development of the M.D.I., 73 subjects (33 females, 40 males) from four different classrooms were assessed. The nontreatment scores for males and females were very similar (females, $\bar{M} = 7.63$; males, $\bar{M} = 7.78$). In the experiment, using four different classrooms, the M.D.I. nontreatment mean score for females was $\bar{M} = 8.28$ and for males, $\bar{M} = 6.31$.

In the present study, the significant main effects are qualified by a significant three-way interaction. This interaction indicated that there were, in fact, no general consistencies. The interaction in the ANOVA, using Independent Story score as the dependent variable, suggested that an accurate picture could only be obtained by considering the interaction of sex x group x stories. Newman-Keuls post hoc comparisons of means tests scores detected several significant differences.

The Newman-Keuls tests scores were examined for possible trends related to the differential treatment effect by sex. Experimental males had significantly higher morality scores on Story 4 than on Story 1, 2, or 3. There were no significant differences among the experimental female subjects' M.D.I. scores on any of the four stories. Story 4, then, was more sensitive in measuring treatment effect for males, but the four stories measured treatment effect equally in females.

Since two of the Inventory Stories contained male protagonists and two contained female protagonists, the data were examined to determine whether the subjects scored higher on Inventory Stories with like-sex protagonists. Maccoby and Jacklin (1974) noted that children choose same-sex models and use these models more than opposite-sex models for patterning their own behavior. The present study did not find that subjects scored higher on the Inventory Stories with same-sex protagonists. There was no significant difference

in experimental females scores on Stories 1 and 3 (male protagonists) and Stories 2 and 4 (female protagonist). The experimental males scored highest on Story 4 (female protagonist), with no significant difference in the scores on Stories 1 (male), 2 (female), and 3 (male).

The scores were also examined to identify any trends related to story theme. At least a portion of the measuring strength of the M.D.I. can be attributed to its item situation variability. The Inventory Stories deal with morality in four different behavioral situations. Story 1 concerns the subject's willingness to allow an animal (a pet) to be punished instead of himself. Story 2 concerns peer group pressure versus self-initiated personal responsibility. Story 3 deals with school dishonesty, and Story 4 with deceiving a parental figure. The similarity of the experimental female subjects' scores on each test could indicate that their experience and training had provided opportunities for moral behavior to have generalized to several different situations (Endler & Hunt, 1969). The significantly higher score on Story 4 for males could indicate cultural conditioning. Although the primary morality issue in Story 4 was deception, it was the only story that concerned a work task and the responsibility of finishing a job. There is some evidence that differential parental treatment of the sexes affects this variable as early as preschool (Hurlock, 1964).

Experimenter reliability was not investigated. This researcher individually tested each subject and also performed the treatment conditions to both the control and experimental groups. Whereas the latter condition had built-in control (the difference between the two groups was not the result of different researcher personality), it is essential to note that different scores may have been obtained by different testers. The experimenter personal attributes effect has been studied by many (Hyman, 1954; Johnson, 1976, Masling, 1965), indicating that experimenters differing in such characteristics as sex, age, ethnic identity, prestige, anxiety, friendliness, dominance, and warmth at times obtained divergent results when testing similar subjects. The scores on the M.D.I. in the present study could have been affected by the fact that this researcher has had many years' experience working with young children, the children (with only 6 exceptions) were white, as was the experimenter, the first graders all had been receiving instructions from female teachers, there were no disruptive discipline problems, and all children readily agreed to being tested. It is suggested that further research utilize experimenters of different sexes, races, and ages, and experimenters who do not know the experimental hypothesis.

In summary, the analyses of the experimental data provided the following information:

1. The treatment effect (reading of morality stories for five consecutive days) was highly significant, but only

for the male subjects. The control males had significantly lower scores than all other subjects, but the treatment effect was successful in bringing those scores up to the experimental female scores.

2. The individual Inventory Stories were equally sensitive in measuring treatment effect for female subjects. Story 4 was more effective in measuring treatment effect for males.

3. Subjects did not score significantly higher on Inventory Stories with like-sex models.

Implications of the study and suggestions for further research are discussed in Chapter VI.

CHAPTER VI
SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary

The purposes of the study were two-fold: (a) to develop an instrument for measuring level of moral development in first-grade children and (b) to experimentally test the hypothesis that moral development could be fostered in first-grade children by reading them morality stories selected for their similarity to the induction child-rearing technique described by Hoffman and Saltzstein (1967). The induction technique was defined as instructing the child concerning the consequences of his wrong deeds on others. Hoffman and Saltzstein (1967) found a positive correlation between the induction child-rearing technique and high moral development in children. A later review of the research on child-rearing (Hoffman, 1975a) also supported the generalization that moral orientation was associated with the mother's frequent use of inductive discipline techniques.

The study of morality has at various times been a central issue in child development (Albert & Kluckhorn, 1959), but theory concerning moral development has not been conclusive. Psychoanalytic theory, prevalent in the 1950's, stressed the importance of the parent in shaping the moral character of the child (Hoffman, 1970). The parent was

encouraged to punish the child by depriving him of love. To avoid losing parental love, the child accepted the parents' standards of conduct as his own, thus developing a system of internalized control (Izard, 1977). Later studies (Allin-smith & Greening, 1955; Hill, 1960; Hoffman & Saltzstein, 1967) did not find the parental discipline technique of love-withdrawal related to moral development.

The cognitive-developmental theoretical approach to moral development assumed that rules became transformed into internal principles through cognitive processes which evolved in stages (Kohlberg, 1963; Piaget, 1932). Although the theory stressed intellectual development over time as the major precursor to moral development, peer interaction, child-rearing practices, and even direct instruction were considered important environment facilitators (Kohlberg, 1966; Piaget, 1932; Stanton, 1976). The weakness in the cognitive development theory centered predominantly on the issue that there was not a strong correlation between high moral reasoning and moral behavior (Stanton, 1976). The theory also indicated that moral development always increased, that there was no regression (Kohlberg, 1975). Research refuting that position is prevalent (Hsu, 1961; Simpson, 1974; Witkin, 1974).

The social learning theory position that reward and punishment in the home were major variables in the development of morality (Lovell, 1971) led to research which attempted

to locate those specific parental practices that correlated with morality characteristics in children (Sears, Maccoby, & Levin, 1967). Behavior research was interested in the environmental contingencies that altered behavior in the desired way. Skinner (1969) explained moral behavior as learned through practice and reinforcement. Some of the most profound and discriminating teaching of morality probably takes place through the study of literature. McClelland (1961) found that the materials the school selects for study influence the children's values, attitudes, and behavior. The social learning theory of moral development was the basis for the present experimental study on fostering moral development.

Although theory on how a person becomes moral was not conclusive, there was much agreement among theorists that guilt could be used as a measurable index for the presence of moral development (Ausubel, 1958; Fenichel, 1955; Freud, 1964; Hoffman & Saltzstein, 1967; Puiynski, 1975; Redl & Wineman, 1951). The first purpose of the present study, to develop an instrument for measuring level of moral development in first-grade children, accepted the assumption that guilt was a measurable index of moral development and used it to determine level of moral development in first graders.

A multiple-choice story completion test was developed to assess the intensity and/or presence of morality guilt in first-grade children. The instrument was modeled after the instrument used by Hoffman and Saltzstein (1967) to

measure guilt level in seventh graders. The instrument, hereafter referred to as the Moral Development Inventory, or M.D.I., consisted of Four Inventory Stories. Each story contained a protagonist the age of the subject to be tested. Each story presented a situation in which the protagonist committed a wrong deed. The wrong deed was known only to the protagonist and no external punishment was involved. Each Inventory Story was followed by three multiple-choice endings, constructed by the researcher to correspond to high, medium, and low levels of guilt as described by Hoffman and Saltzstein (1967). The Moral Development Inventory was subjected to a pilot study for verification that the story theme and vocabulary were suitable for use with first graders, and that the multiple-choice endings were representative of the endings given by first graders. Three hundred first-grade students (168 males, 148 females) were administered the M.D.I. without the multiple-choice endings so that the subjects could complete the stories with open endings. Analysis of the responses from the pilot study offered no justification for changing the structure of the Moral Development Inventory.

The Moral Development Inventory was then administered to 73 first graders (33 females, 40 males) in Walton County, Florida. A numerical score was obtained for each subject (see Appendix D for scoring procedures). Each subject's score was correlated with a respective rating from the subject's teacher (see Teacher Form in Appendix E). The

Pearson product-moment correlation coefficient, $r = .37$, $p < .001$, resulted in the rejection of the null hypothesis that there would be no significant relationship between the subjects' scores on the M.D.I. and the teacher rating of subject's moral development. The correlation between the subject's score on the M.D.I. and his score on the language portion of the Comprehensive Tests of Basic Skills, $r = .32$, $p < .01$, also resulted in the rejection of the null hypothesis that there would be no significant relationship between the subjects' scores on the M.D.I. and the scores on the language portion of the Comprehensive Tests of Basic Skills. The possibility that the subjects' response was a function of response set was ruled out by Chi-square analyses, $\chi^2 (2, 30 = 7.04, p > .05$. The Moral Development Inventory was analyzed for split-half reliability by the Pearson product-moment correlation coefficient, $r = .63, p < .001$. The Pearson product-moment correlation coefficient was also used to determine test-retest reliability, $r = .82$. The computed F max statistic for homogeneity of variances was 1.36, $df = 73, k = 4, p > .05$, indicating internal consistency in the four Inventory Stories. The Moral Development Inventory was accepted as satisfactory, and was used to assess the subjects in the experiment.

The subjects in the experiment were 27 first graders assigned to experimental treatment conditions and 27 first graders assigned to control treatment conditions. The

experimental subjects were read one morality story each day for five consecutive days. The morality stories were chosen to illustrate the induction technique as defined by Hoffman and Saltzstein (1967). The control subjects were read one nonmorality story each day for five consecutive days. The experimental effects were assessed by a three-way ANOVA, which determined that there was a significant main effect for group, a significant effect for sex, and a significant effect for group by sex. The analysis indicated no significant differences in treatment effect on the individual Inventory Stories, but there was a significant three-way (sex x group x story) interaction. Newman-Keuls post hoc comparisons on the means scores revealed that the differences were primarily related to the very low scores of the control males, and the lack of significant treatment effect on the females (see Tables 6-14).

Conclusions

When analyzing the results of the study the following conclusions were drawn.

1. It was possible to foster moral development in the male first-grade children by reading them morality content stories chosen to duplicate the inductive method described by Hoffman and Saltzstein (1967). The results of the study did not show a significant treatment effect, as measured by the M.D.I., for female subjects.

2. It was possible to develop an instrument for assessing moral development in first-grade children, using the index guilt for measuring moral development level. The reliability and validity analyses of the Moral Development Inventory indicated that more research was needed before it could be used as an instrument for individual prediction, but it functioned well in measuring treatment effect.

3. Three levels of guilt, as described by Hoffman and Saltzstein (1967), did appear in the first-grade subjects.

4. Male and female subjects did differ in level of moral development as measured by the M.D.I. in both the control and experimental groups.

5. The induction technique for fostering moral development would seem to be an effective child-rearing procedure for producing moral children, especially males.

Recommendations

Based upon the procedures, findings, and conclusions of the study, the following recommendations for future research are suggested:

1. The effectiveness of the induction child-rearing method should be studied experimentally. The method in the present study of using morality stories that illustrated the induction technique infers, but does not prove, that the induction child-rearing method produces moral children. Parents could be taught the induction method, perhaps by a simulated situation technique, and the M.D.I. could be used to assess the effectiveness of that procedure.

2. The type of literature that is used in the schools should be considered in terms of its effect on the moral development of the students. Parents have shown great concern about the removal of certain books from school libraries and classrooms, especially books dealing with sex and violence. A more positive approach is warranted. Parents should show more concern about the addition of books to the libraries and classrooms that illustrate the induction process, i.e., instruct the child concerning the consequences of his bad deeds on other persons.

3. One of the aims of the present research was to determine whether three levels of guilt existed in first-grade children. Younger children need to be tested to determine the age when the emotion guilt first appears and at what ages it begins to increase in intensity.

4. The M.D.I. should be further evaluated.

a. The test was designed for first-grade children. Its utility for younger and older children should be explored.

b. The samples in the present study were predominantly white, middle-class, Southerners. Samples differing in demographic complexion are needed so that more generalizations from the M.D.I. scores could be made. Children in private church schools could be compared with public school children. The M.D.I. scores should be higher, as the church school children are usually exposed to more morality-content stories.

c. The M.D.I. should be administered under conditions of greater experimenter control. These conditions include reliability checks on the test administrator, investigation of the effect of tester variables such as race, age, and sex, using a test administrator who is blind to the hypotheses, and determining how much training is necessary for the test administrator.

d. Retest after a longer period than one week, or develop alternate forms. A test instrument that utilizes story situations may be retesting the subject's memory rather than the test construct. Alternate forms could reduce the instrument effect in a pre- or posttest experiment, thus increasing validity.

e. The M.D.I. scores were correlated with language grade scores and with teacher ratings of the subjects' morality. Additional measures for correlational analysis are peer ratings, physiological measurements on the subject while he is tested, parent ratings, and observations in behavioral situations. The M.D.I. should also be correlated with IQ scores. Miller and Swanson (1960) did not find moral development related to IQ. The positive relationship between M.D.I scores in the present study and the language scores of the subjects indicates that more research is needed in this area to substantiate that the M.D.I. is measuring something other than intelligence.

f. Follow up studies could determine whether first graders who scored high on the M.D.I. showed evidence of higher morality later in life. This would strengthen the predictive validity of the M.D.I.

g. Correlate the M.D.I. scores with scores on a test of empathy. Empathy appears to be the cornerstone for the effectiveness of the induction method. If so, there should be a high correlation between empathy and M.D.I. scores.

h. Administer the test with and without the pictures. If the results do not vary, the test might be easier to administer without the pictures.

i. If different samples yield the same treatment effect with males and not with females, as in the present study, more work should be done on the M.D.I. Perhaps the test could be lengthened, or a broader range of situations could be included in the Inventory Stories. The treatment conditions should also be altered to determine whether the insignificant treatment effect with the females was a function of the treatment or the instrument. A longer treatment should be considered, and retest measurements taken periodically to determine the stability of the treatment effect over time.

j. Directly test the sex of the protagonist effect by using the same situation-story but changing the sex of the protagonist. Male and female scores on this variable could be compared.

k. The researcher has a responsibility of integrity when working with human subjects in the area of morality, but especially so when working with young children. The researcher must show sensible regard for the social and moral codes of the community in which he works and must avoid any exercise of excessive and unwarranted control over the personal destinies of children. Sugarman (1970) discussed the responsibility that research has to the public in truthful reporting of results and avoiding of excess claims of results or unwarranted generalizations from specific pieces of evidence. In the present study, the researcher has given special consideration to the matter of ethics in dealing with children in the treatment procedures, in the testing situation, and in the handling of the data.

BIBLIOGRAPHY

- Achenbach, T. M. The classification of children's psychiatric symptoms: A factor analytic study. Psychological Monographs, 1966, 80 (6).
- Albert, E. M., & Kluckhohn, C. Selected bibliography on values ethics, and esthetics in the behavioral sciences and philosophy, 1920-1958. Glencoe, Illinois: Free Press, 1959.
- Allinsmith, W. The learning of moral standards. In D. R. Miller & G. E. Swanson, Inner conflict and defense. New York: Holt, 1960.
- Allinsmith, W., & Greening, T. C. Guilt over anger as predicted from parental discipline: A study of superego development. American Psychologist, 1955, 10, 320. (Abstract)
- Ambron, S. R. Child development (2nd ed.). New York: Holt, Rinehart, & Winston, 1978.
- Ambron, S. R., & Irwin, D. M. Role taking and moral judgments in five- and seven-year-olds. Developmental Psychology, 1975, 11 (1), 102.
- Anderson, J. Child development: An historical perspective. Child Development, 1956, 27 (2), 181-196.
- Aronfreed, J. Conduct and conscience. New York: Academic Press, 1968.
- Ausubel, D. P. Relationships between shame and guilt in the socializing process. Psychological Review, 1955, 62, 378-390.
- Ausubel, D. P. Theory and problems of adolescent development. New York: Grune and Stratton, 1958.
- Ausubel, D. P., & Sullivan, E. V. Theory and problems of child development (2nd ed.). New York: Grune & Stratton, 1970.
- Bandura, A., & McDonald, F. J. Influence of social reinforcement and the behavior of models in shaping children's moral judgments. Journal of Abnormal and Social Psychology, 1963, 67, 274-281.

- Bandura, R., Ross, D., & Ross, S. A. Imitation of film-mediated aggressive models. Journal of Abnormal and Social Psychology, 1963, 66, 3-11.
- Bandura, A., & Walters, R. H. Social learning and personality development. New York: Holt, Rinehart & Winston, 1963.
- Baumrind, D. Socialization and instrumental competence in young children. Young Children, 1970, 26 (2), 104-119.
- Bell, R. Q. Stimulus control of parent or caretaker behavior by offspring. Developmental Psychology, 1971, 4, 63-72.
- Benedict, R. The chrysanthemum and the sword. Cambridge: Riverside Press, 1946.
- Bijou, S. W. Moral development in the preschool years: A functional analysis. Revista Mexicana de Analisis de la Conducta, 1975, 1 (1), 11-29.
- Borke, H. Piaget's mountain revisited: Changes in the ego-centric landscape. Developmental Psychology, 1975, 11, 240-243.
- Brim, O. G., Jr. Adult socialization. In J. A. Clausen (Ed.), Socialization and society. Boston: Little Brown, 1968.
- Brody, G. F. Relationship between maternal attitudes and behavior. Journal of Personal and Social Psychology, 1965, 2, 317-323.
- Bronfenbrenner, U. (Participant). The acquisition and development of values: Perspectives on research. Report of a conference, May 15-17, 1968, Washington, D. C. Bethesda, Maryland: National Institute of Child Health and Human Development.
- Brown, C. C. (Ed.). Methods of psychophysiology. New York: Williams & Wilkins, 1967.
- Bruner, J. S. Beyond the information given: Studies in the psychology of knowing. New York: Norton, 1973.
- Bryk, A. S., & Weisberg, H. I. Value-added analysis: A dynamic approach to the estimation of treatment effects. Journal of Educational Statistics, 1976, 1 (2), 127-155.
- Bull, N. J. Moral judgment from childhood to adolescence. London: Routledge and Kegan Paul, 1969.

- Burton, R. V., Maccoby, E. E., & Allinsmith, W. Antecedents of resistance to temptation in four-year-old children. Child Development, 1961, 32, 689-710.
- Campbell, J. D. Peer relations in childhood. In M. L. Hoffman & L. W. Hoffman (Eds.), Review of child development research. New York: Russell Sage Foundation, 1964.
- Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental design for research. Chicago: Rand McNally, 1963.
- Chandler, M. J. Egocentrism and antisocial behavior: The assessment and training of social perspective-taking skills. Developmental Psychology, 1973, 9, 326-332.
- Cheyne, J. A., & Walters, R. H. Punishment and prohibition: Some origins of self-control. In T. M. Newcomb (Ed.), New directions in psychology, IV. New York: Holt, Rinehart & Winston, 1970.
- Comprehensive Tests of Basic Skills, Level B, Form S, Examiner's Manual. Monterey, California: CTB/McGraw-Hill, 1974.
- Cronbach, L. J. Further evidence on response sets and test design. Educational and Psychological Measurement, 1950, 10, 3-31.
- Cronbach, L. J. Essentials of psychological testing (3rd ed.). New York: Harper & Row, 1970.
- Cronbach, L. J., & Drenth, P. J. D. (Eds.). Mental tests and cultural adaptation. The Hague, Netherlands: Mouton Publishers, 1972.
- Dennis, W. Animism and related tendencies in Hopi children. Journal of Abnormal and Social Psychology, 1943, 38, 21-37.
- Dewey, J. Moral principles in education. Boston: Houghton-Mifflin, 1909.
- Dolto, F. French and American children as seen by a French child analyst. In M. Head & M. Wolfenstein (Eds.), Childhood in contemporary cultures. Chicago: University of Chicago Press, 1955.
- Durkin, D. Children's concepts of justice: A comparison with the Piaget data. Child Development, 1959, 30, 59-67.

- Eibl-Eibesfeldt, I. Love and hate: The natural history of behavior patterns. New York: Holt, Rinehart & Winston, 1971.
- Ellis, A. Reason and emotion in psychotherapy. New York: Lyle Stuart, 1962.
- Endler, N. S., & Hunt, J. McV. Generalizability of contributions from sources of variance with R-R Inventories of Anxiousness. Journal of Personality, 1969, 37, 1-24.
- Feffer, M. H. A developmental analysis of interpersonal behavior. Psychological Review, 1970, 77, 197-214.
- Fenichel, O. The psychoanalytic theory of neurosis. New York: Norton, 1945.
- Fishbein, H. D., Lewis, S., & Keiffer, K. Children's understanding of spatial relations: Coordination of perspectives. Developmental Psychology, 1972, 7, 21-33.
- Fraiberg, S. H. The magic years. New York: Scribner's, 1959.
- Freud, A. The writings of Anna Freud (Vol. 4). New York: International Universities Press, 1968.
- Freud, S. An outline of psychoanalysis. New York: Norton, 1949.
- Freud, S. New introductory lecture: The standard edition. London: Hogarth Press, 1964.
- Glassco, J., Milgram, N. A., & Youniss, J. The stability of training effects on intentionality of moral judgments in children. Journal of Personal and Social Psychology, 1970, 14, 360-365.
- Glueck, S., & Glueck, E. Unraveling juvenile delinquency. Cambridge, Mass.: Howard University Press, 1950.
- Goldfried, M. R., & Sprafkin, J. N. Behavioral personality assessment. New York: University Programs Modular Studies, 1974.
- Grinder, R. E. New techniques for research in children's temptation behavior. Child Development, 1961, 32, 679-688.
- Grinder, R. E. Parental childrearing practices, conscience, and resistance to temptation of sixth-grade children. Child Development, 1962, 33, 803-820.

- Grinder, R. E. Relations between behavior and cognitive dimensions of conscience in middle childhood. Child Development, 1964, 35, 881-891.
- Hall, R. T., & Davis, J. U. Moral education in theory and practice. Buffalo, New York: Prometheus Books, 1975.
- Harper, I. The young as a source of stimuli controlling caretaker behavior. Developmental Psychology, 1971, 4, 73-88.
- Hartshorne, H., May, M. A., & Shuttleworth, F. K. Studies in the organization of character. New York: MacMillan, 1930.
- Havighurst, R. J., & Neugarten, B. L. American Indian and white children. Chicago: University of Chicago Press, 1955.
- Henry, A. F., & Short, J. F. Suicide and homicide. Glencoe, Illinois: Free Press, 1954.
- Hersh, R. H., & Paolitto, D. P. Moral development: Implication for pedagogy. Education Digest, 1977, XIII (5), 13-16.
- Hicks, C. R. Fundamental concepts in the design of experiments. New York: Holt, Rinehart & Winston, 1973.
- Hill, W. F. Learning theory and the acquisition of values. Psychological Bulletin, 1960, 67, 317-331.
- Hoffman, M. L. Personality, family structure, and social class as antecedents of parental power assertion. Child Development, 1963, 34, 869-880.
- Hoffman, M. L. Moral development. In P. H. Mussen (Ed.), Carmichael's manual of child psychology (Vol. 2, 3rd ed.). New York: John Wiley, 1970.
- Hoffman, M. L. Moral internalization, parental power, and the nature of parent-child interaction. Developmental Psychology, 1975, 11 (2), 228-239. (a)
- Hoffman, M. L. Sex differences in moral internalization and values. Journal of Personality and Social Psychology, 1975, 32 (4), 720-729. (b)
- Hoffman, M. L. Personality and social development. Annual Review of Psychology, 1977, 28, 295-321.

- Hoffman, M. L., & Saltzstein, H. D. Parent discipline and the child's moral development. Journal of Personality and Social Psychology, 1967, 5 (1), 45-57.
- Holstein, C. B. Parental consensus and interaction in relation to the child's moral judgment. (Doctoral dissertation, University of California, Berkeley, 1969). Dissertation Abstracts International, 1970, 31, 1888A-18889A (University Microfilms No. 70-17, 579).
- Holstein, C. B. The relation of children's moral judgment level to that of their parents and to communications patterns in the family. In R. C. Smart & M. S. Smart (Eds.), Readings in child development and relationships. New York: MacMillan, 1972.
- Holstein, C. B. Irreversible, step-wise sequence in the development of moral judgments: A longitudinal study of males and females. Child Development, 1976, 47, 51-61.
- Horney, K. New ways in psychoanalysis. New York: Norton, 1939.
- Hsu, R. L. K. Culture pattern and adolescent behaviour. International Journal of Social Psychiatry, 1961, 7, 33-53.
- Hurlock, E. B. Child development (4th ed.). New York: McGraw-Hill, 1964.
- Hyman, H. H. Interviewing in social research. Chicago: University of Chicago Press, 1954.
- Izard, C. E. Human emotions. New York: Plenum Press, 1977.
- Janis, I. L. (Ed.). Personality: Dynamics, development, and assessment. New York: Harcourt, Brace and World, 1969.
- Johnson, O. G., & Bommarito, J. W. Tests and measurements in child development: A handbook. San Francisco: Jossey-Bass, Inc., 1971.
- Johnson, R. C., & Medinuis, G. R. Child psychology: Behavior and development. New York: John Wiley, 1974.
- Johnson, R. W. Subject performance as affected by experimenter expectancy, sex of experimenter, and verbal reinforcement. Canadian Journal of Behavioral Science, 1970, 2, 60-66.

- Kay, W. Moral development. London: George Allen and Unwin, 1968.
- King, S., & Henry, A. Aggression and cardiovascular reactions related to parental control over behavior. Journal of Abnormal and Social Psychology, 1955, 50, 206-210.
- Kohlberg, L. The development of children's orientation toward a moral order. Vita Humana, 1963, 6, 11-33.
- Kohlberg, L. Moral education in the schools: A developmental view. The School Review, 1966, 74, 1-30.
- Kohlberg, L. Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally, 1969.
- Kohlberg, L. Education, moral development and faith. Journal of Moral Education, 1974, 4, 5-16.
- Kohlberg, L., & Turiel, E. Moral development and moral education. In G. Lesser (Ed.), Psychology and educational practice. Glenview, Illinois: Scott, Foresman, 1971.
- Krebs, R. L. Some relationships between moral judgment, attention and resistance to temptation. Unpublished doctoral dissertation, University of Chicago, 1968.
- Kurtines, W., & Greif, E. B. The development of moral thought: Review and evaluation of Kohlberg's approach. Psychological Bulletin, 1974, 81, 453-470.
- Lanyon, R. I., & Goodstein, L. D. Personality assessment. New York: John Wiley, 1971.
- Lewis, H. B. Shame and guilt in neurosis. New York: International Universities Press, 1971.
- Lidz, T. The person: His development through the life cycle. New York: Basic Books, 1968.
- Liu, C. The influence of cultural background on the moral judgment of children. Unpublished doctoral dissertation, Columbia University, 1950.
- Looft, W. R. Egocentrism and social interaction across the life span. Psychological Bulletin, 1972, 78, 73-92.
- Lovell, K. An introduction to human development. Glenview, Illinois: Scott, Foresman, 1971.

- Maccoby, E. E., & Jacklin, C. N. Psychology of sex differences. California: Stanford University Press, 1974.
- MacKinnon, D. W. Violation of prohibitions. In H. W. Murray (Ed.), Explorations in personality. New York: Oxford University Press, 1938.
- MacRae, R., Jr. A test of Piaget's theories of moral development. Journal of Abnormal and Social Psychology, 1954, 49, 14-18.
- Magowan, S. A., & Lee, T. Some sources of error in the use of the projective method for the measurement of moral development. British Journal of Psychology, 1970, 61, 535-543.
- Martin, B. Parent-child relations. In F. D. Horowitz (Ed.), Review of child development research (Vol. 4). Chicago: University of Chicago Press, 1975.
- Masling, J. Differential indoctrination of examiners and Rorschach responses. Journal of Consulting Psychology, 1965, 29, 198-201.
- McCandless, B. R., & Evans, E. D. Children and youth: Psychosocial development. Hinsdale, Illinois: Dryden Press, 1973.
- McClelland, D. C. The achieving society. Princeton: Van Nostrand, 1961.
- Medinnus, C. R. Immanent justice in children: A review of the literature and additional data. Journal of Genetical Psychology, 1959, 90, 253-262.
- Miller, D. R., & Swanson, G. E. Inner conflict and defense. New York: Henry Holt, 1960.
- Mills, I. Moral decision-making, religious reinforcement and some educational implications. Journal of Moral Education, 1977, 6 (3), 166-169.
- Mischel, W. Personality and assessment. New York: John Wiley & Sons, 1968.
- Mosher, D. L. The development and multitrait-multimethod matrix analysis of three measures of three aspects of guilt. Journal of Consulting Psychology, 1966, 30 (1), 25-29.

- Mowrer, O. H. Identification: A link between learning theory and psychotherapy. In O. H. Mowrer, Learning theory and personality dynamics. New York: Ronald Press, 1952.
- Mowrer, O. H. The crisis in psychiatry and religion. Princeton, N. J.: Van Nostrand, 1961.
- Mowrer, O. H. (Ed.). Morality and mental health. Chicago: Rand McNally, 1967.
- Munroe, R. L. Schools of psychoanalytic thought. New York: Henry Holt, 1955.
- Night of terror. Time, July 25, 1977, pp. 12-22.
- Odom, R. D. Problem-solving strategies as a function of age and socioeconomic level. Child Development, 1967, 38, 747-752.
- Parke, R. D. Some effects of punishment on children's behavior. Young Children, 1969, 24, 224-240.
- Piaget, J. Moral judgment of the child. New York: Harcourt, Brace, 1932.
- Piaget, J. The moral judgment of children. Glencoe, Illinois: Free Press, 1948.
- Piaget, J. Science of education and the psychology of the child. New York: Orion Press, 1970.
- Piaget, J., & Inhelder, B. The child's conception of space. London: Routledge and Kegan Paul, 1956.
- Piers, G., & Singer, M. B. Shame and guilt: A psychoanalytic and a cultural study. Springfield, Ill.: Thomas, 1953.
- Pikas, A. Children's attitudes toward rational versus inhibiting parental authority. Journal of Abnormal and Social Psychology, 1961, 62, 315-321.
- Psychology Encyclopedia 73/74. Guilford, Connecticut: Dushkin Publishing Group, 1973.
- Puiynski, S. Psychologia poczucia winy. Psychiatria Polska, 1975, 9 (4), 443-456.
- Ray, W. S. An introduction to experimental design. New York: MacMillan, 1960.

- Rheingold, H. L. The social and socializing infant. In D. A. Goslin (Ed.), Handbook of socializing theory and research. Chicago: Rand McNally, 1969.
- Redl, F., & Wineman, D. Children who hate. New York: Free Press, 1951.
- Reese, H. W. Discrimination learning set and perceptual set in young children. Child Development, 1965, 36, 153-161.
- Reese, H. W., & Lipsitt, L. P. Experimental child psychology. New York: Academic Press, 1970.
- Rickman, J. (Ed.). A general selection from the works of Sigmund Freud. New York: Doubleday Anchor Books, 1957.
- Sanborn, S. Means and ends: Moral development and moral education. Harvard Graduate School of Education Bulletin, Fall, 1971.
- Sarason, S. B. The measurement of anxiety in children: Some questions and problems. In C. D. Spielberger (Ed.), Anxiety and behavior. New York: Academic Press, 1966.
- Sarbin, T. R. Role theory. In G. Lindzey (Ed.), Handbook of social psychology. Cambridge, Mass.: Addison-Wesley, 1954.
- Schill, T., Evans, R., & McGovern, T. Child-rearing attitudes of subjects who vary in guilt. Journal of Clinical Psychology, 1976, 32 (2), 203-284.
- Sears, R. R. Relation of early socialization experiences to aggression in middle childhood. Journal of Abnormal and Social Psychology, 1961, 63 (3), 466-492.
- Sears, R. R., Maccoby, E., & Levin, H. Patterns of child rearing. Evanston, Illinois: Row Peterson, 1957.
- Sears, R. R., Rau, L., & Alpert, R. Identification and child rearing. Stanford, California: Stanford University Press, 1965.
- Siegal, Michael. Spontaneous development of moral concepts. Human Development, 1975, 18, 370-383.
- Simner, M. L. Newborn's response to the cry of another infant. Developmental Psychology, 1971, 5, 136-150.
- Simpson, E. L. Moral development research: A case study of scientific cultural bias. Human Development, 1974, 17, 81-106.

- Skinner, B. F. Science and human behavior. New York: MacMillan, 1953.
- Skinner, B. F. Contingencies of reinforcement: A theoretical analysis. New York: Appleton-Century-Crofts, 1969.
- Solomon, R. L., Turner, L. H., & Lessac, M. S. Some effects of delay of punishment on resistance to temptation in dogs. Journal of Personality and Social Psychology, 1968, 8 (3), 233-238.
- Stanton, M. The assessment of moral judgments: Cultural and cognitive considerations. Religious Education, 1976, LXXI (6), 610-621.
- Sugarman, J. M. Research, evaluation, and public policy: An invited editorial. Child Development, 1970, 41(2), 263-266.
- Switzer, D. K. A psychodynamic analysis of grief in the context of an interpersonal theory of self. Dissertation Abstracts, 1968, 29 (13), 381.
- Thompson, L. A. Role-playing ability and social adjustment in children. (Doctoral dissertation, University of Southern California, 1968). Dissertation Abstracts International, 1968, 29, 3499B. (University Microfilms No. 69-4547)
- Thoresen, C. E., & Mahoney, M. J. Behavior self-control. New York: Holt, Rinehart & Winston, 1974.
- Turiel, E. An experimental test of the sequentiality of developmental stages in the child's moral judgments. Journal of Personality and Social Psychology, 1966, 6, 611-618.
- Unger, J. M. On the development of guilt reactivity in the child. Unpublished doctoral dissertation, Cornell University, 1962.
- Watson, R. I., & Lindgren, H. C. Psychology of the child (3rd ed.). New York: John Wiley & Sons, 1973.
- Weir, M. W. Effects of age and instruction on children's probability learning. Child Development, 1962, 33, 729-735.
- Whiteman, P. H., & Kosier, P. Development of children's moralistic judgments: Age, sex, I.Q. and certain personal experiential variables. Child Development, 1964, 35, 843-851.

- Whiting, J. W. M. (Fourth presentation). In J. M. Tanner & B. Inhelder (Eds.), Discussions on child development: II. London: Tavistock Publications, 1954.
- Whiting, J. W. M., & Child, I. L. Child training and personality. New Haven: Yale University Press, 1953.
- Williams, N., & Williams, S. The moral development of children. London: MacMillan, 1970.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill Book Company, 1962.
- Witkin, H. A. Social conformity and psychological differentiation. International Journal of Psychology, 1974, 9, 11-29.
- Yap, P. M. Phenomenology of affective disorder in Chinese and other cultures. In A. V. S. de Reuck & R. Porter (Eds.), Transcultural psychiatry. London: Churchill, 1965.
- Yarrow, M. R., Campbell, J. D., & Burton, R. V. Child rearing. San Francisco: Jossey-Bass, 1968.
- Yarrow, M. R., & Waxler, C. Z. The emergence and functions of prosocial behaviors in young children. In E. M. Hetherington & R. D. Parke (Eds.), Contemporary readings in child psychology. New York: McGraw-Hill, 1977.
- The youth crime plague. Time, July 11, 1977, pp. 18-28.

Appendix A

Kohlberg's Definition of Moral Stages

I. Preconventional level

At this level, the child is responsive to cultural rules and labels of good and bad, right or wrong, but interprets these labels either in terms of the physical or the hedonistic consequences of action (punishment, reward, exchange of favors) or in terms of the physical power of those who enunciate the rules and labels. The level is divided into the following two stages:

Stage 1: The punishment-and-obedience orientation. The physical consequences of action determine its goodness or badness, regardless of the human meaning or value of these consequences. Avoidance of punishment and unquestioning deference to power are valued in their own right, not in terms of respect for an underlying moral order supported by punishment and authority (the latter being Stage 4).

Stage 2: The instrumental-relativist orientation. Right action consists of that which instrumentally satisfies one's own needs and occasionally the needs of others. Human relations are viewed in terms like those of the marketplace. Elements of fairness, of reciprocity, and of equal sharing are present, but they are always interpreted in a physical, pragmatic way. Reciprocity is a matter of "you scratch my back and I'll scratch yours," not of loyalty, gratitude, or justice.

II. Conventional level

At his level, maintaining the expectations of the individual's family, group, or nation is perceived as valuable in its own right, regardless of immediate and obvious consequences. The attitude is not only one of conformity to personal expectations and social order, but of loyalty to it, of actively maintaining, supporting, and justifying the order, and of identifying with the persons or group involved in it. At this level, there are the following two stages:

Stage 3: The interpersonal concordance or "good boy--nice girl" orientation. Good behavior is that which pleases or helps others and is approved by them. There is much conformity to stereotypical images of what is majority or "natural" behavior. Behavior is frequently judged by intention--"he means well". becomes important for the first time. One earns approval by being nice.

Stage 4: The "law and order" orientation. There is orientation toward authority, fixed rules, and the maintenance of the social order. Right behavior consists of doing one's duty, showing respect for authority, and maintaining the given social order for its own sake.

III. Postconventional, autonomous, or principled level

At this level, there is a clear effort to define moral values and principles that have validity and application apart from the authority of the groups or persons holding these principles and apart from the individual's own identification with these groups. This level also has two stages:

Stage 5: The social-contract, legalistic orientation, generally with utilitarian overtones. Right action tends to be defined in terms of general individual rights and standards which have been critically examined and agreed upon by the whole society. There is a clear awareness of the relativism of personal values and opinions and a corresponding emphasis upon procedural rules for reaching consensus. Aside from what is constitutionally and democratically agreed upon, the right is a matter of personal "values" and "opinion." The result is an emphasis upon the "legal point of view," but with an emphasis upon the possibility of changing law in terms of rational consideration of social utility (rather than freezing it in terms of Stage 4 "law and order"). Outside the legal realm, free agreement and contract are the binding element of obligation. This is the official morality of the American government and constitution.

Stage 6: The universal-ethical-principle orientation. Right is defined by the decision of conscience in accord with self-chosen ethical principles appealing to logical comprehensiveness, universality, and consistency. These principles are abstract and ethical (the Golden Rule, the categorical imperative); they are not concrete moral rules like the Ten Commandments. At heart, these are universal principles of justice, of the reciprocity and equality of human rights and of respect for the dignity of human beings as individual persons.

--"From Is to Ought," pp. 164-165.
Reprinted from The Journal
of Philosophy, October 25,
1973

Appendix B

Story Completion Guilt Ratings as Adapted
from Hoffman and Saltzstein (1967)

Rating	Weight	Description
low	1	Story ending contains no explicit guilt. Guilt is externalized or denied. Mild cognitive terms are used rather than feeling terms, e.g., "He knew he was cheating."
medium	2	Story ending contains moderate though clear-cut guilt feelings. Usually indicated by use of mild feeling terms, e.g., "after thinking about it, he gave the prize to the real winner."
high	3	At this level, the guilt varies from a very intense, severe reaction that may not last very long, to one that is not so intense but lasts a day or longer. Examples are as follows: (1) They got little sleep worrying about the accident they caused, (2) He threw the prize down the sewer.

Appendix C

Moral Inventory for First-Grade Children

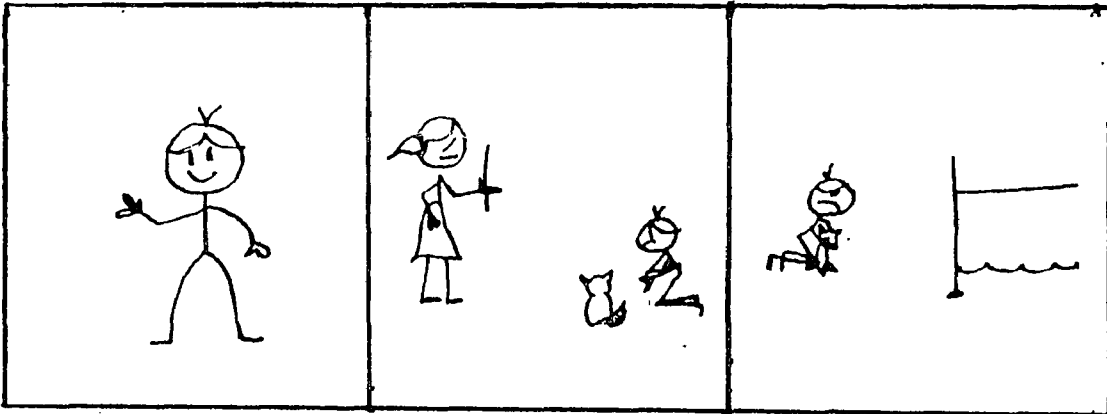
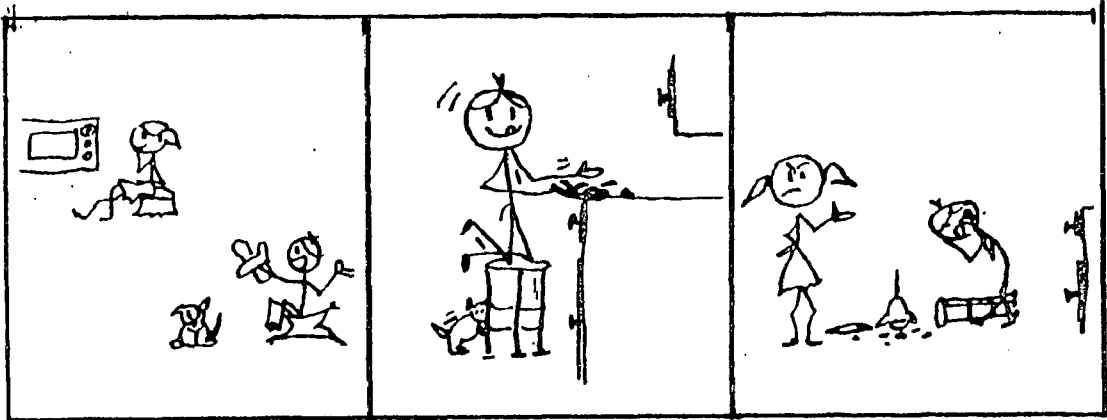
Story 1:

Johnny's mother went to the store and left Johnny with a babysitter. While the babysitter was watching television, Johnny sneaked into the kitchen to get some cookies his mother had told him not to eat before his supper. Johnny accidentally knocked the cookie plate on the floor; and immediately Taffy, the dog, ran over and started eating the spilled cookies. The babysitter walked in, looked at Taffy eating the cookies, and said, "Oh, that bad dog to get into the cookies. I will have to spank him." The babysitter then started to get a switch to spank Taffy.

Which of these endings to the story do you think is the best ending? Wait until you have heard all three endings before you choose the best ending.

- a. Johnny was happy because no one knew that he had sneaked the cookies.
- b. Johnny felt bad after Taffy got a spanking. He told the babysitter what had really happened.
- c. He couldn't watch when the babysitter whipped Taffy. Later he hugged Taffy and cried and cried.

Moral Development Inventory: Story 1



A

B

C

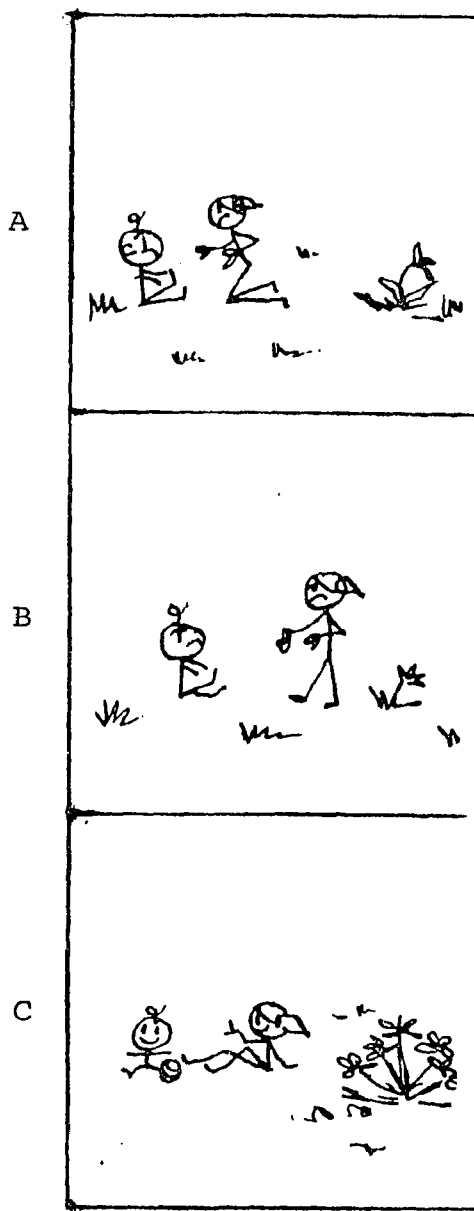
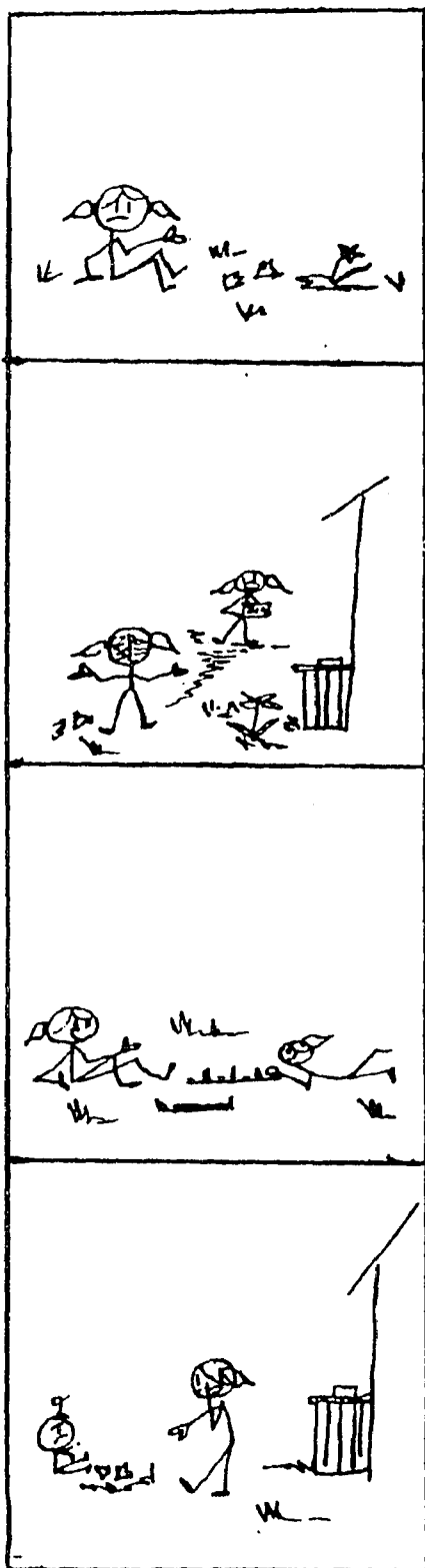
Story 2:

Jenny was playing in the yard by herself when she noticed a broken Coca-Cola bottle on the ground near the porch. Just as she started to pick it up carefully and place it in the trash can, her friend Susan called to her, "Jenny, come on over to my house and see my new game." Jenny immediately forgot about the dangerous broken bottle and ran over to Susan's house. The two girls played with the game most of the afternoon. When Jenny went home she found out that her baby brother had fallen and cut his foot very badly on that broken Coca-Cola bottle.

Which of these endings to the story do you think is the best ending? Wait until you have heard all three endings before you choose the best ending.

- a. Jenny told her brother to always wear shoes when he played in the yard. It was his fault that he cut his foot.
- b. Jenny was sad because she had not picked up the bottle. She told her brother she was sorry about his cut foot.
- c. Everytime Jenny looked at her brother's cut foot, she felt sad all over again that she had not picked up the bottle. She stayed home all week and played with her brother.

Moral Development Inventory: Story 2



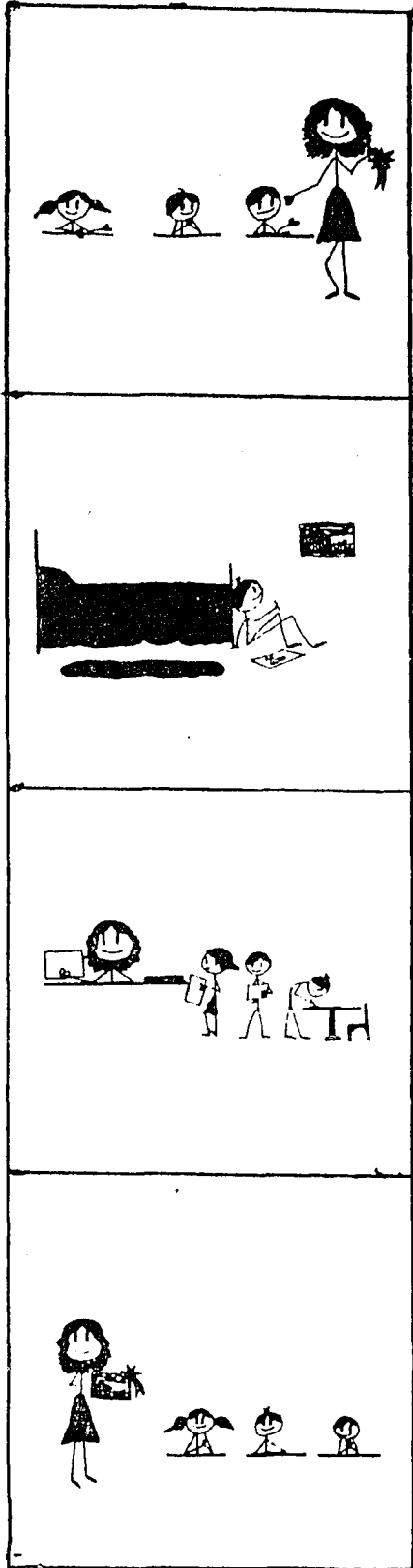
Story 3:

David's first-grade teacher said that she would give a prize to the boy or girl who drew the best picture of a farm. David could not draw very well, but he really wanted the prize. Just as David was beginning to work on his picture, he remembered that he had a very nice picture of a farm in his desk. It was one his older brother had drawn for him a few days ago. When he was sure that no one was looking, David took out his brother's picture and put his own name on it before turning it in to the teacher. When the teacher judged the pictures, she chose David's picture as the best. No one knew that David's brother had drawn the picture. David received the prize, and all of the children clapped to show David how happy they were that he could draw such a fine picture.

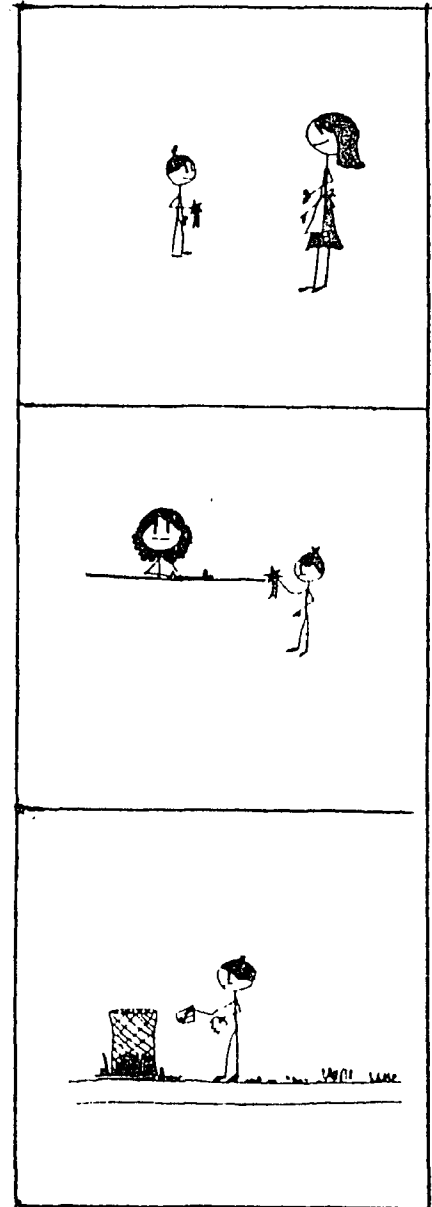
Which of these endings to the story do you think is the best ending? Wait until you have heard all three endings before you choose the best ending.

- a. David was happy because he had the prize. He took the prize home and showed it to his mother.
- b. He felt sad because it wasn't really his drawing. He gave the prize back to the teacher.
- c. All the way home from school, he felt terrible about signing his name on his brother's picture. He tore the picture up and threw it away.

Moral Development Inventory: Story 3



A



B

C

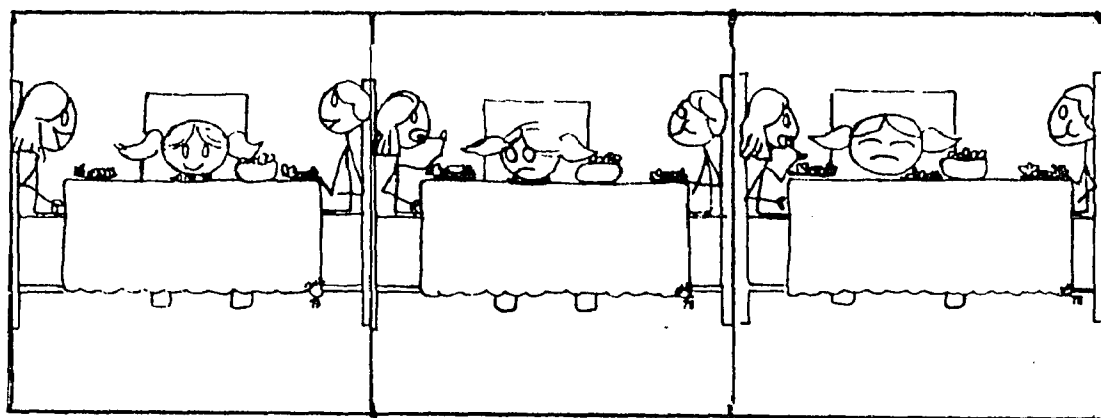
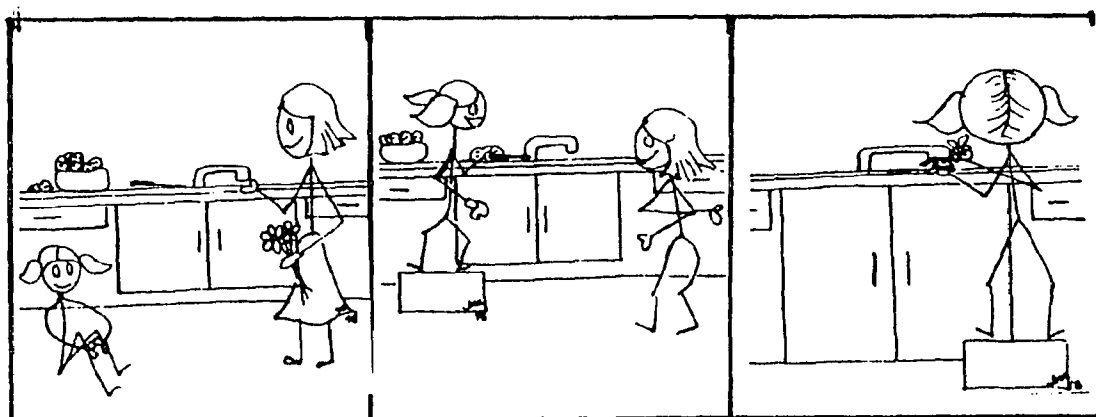
Story 4:

Betty's mother went to visit a friend who was in bed sick. Before she left, Betty's mother told her to finish peeling all of the potatoes and cut them up so that they would be ready to cook for dinner. Betty had peeled all but two of the potatoes when several of her friends came by and invited her to go to the park with them. Betty took the two potatoes she had not peeled and threw them down the garbage disposal along with the potato peelings and went with her friends.

Which of these endings to the story do you think is the best ending? Wait until you have heard all three endings before you choose the best ending.

- a. Betty never told her mother about the two potatoes she had thrown away because everyone had enough potatoes to eat.
- b. At dinner time Betty kept looking at her mother to see if her mother had noticed that some potatoes were missing.
- c. Betty could not eat the potatoes at dinner time, because she felt so bad about disobeying her mother.

Moral Development Inventory: Story 4



A

B

C

Appendix D

Scoring Sheet for Moral Development Inventory

Directions: The story endings are to be presented in random order, using the following combinations: a b c, a c b, c b a, b c a, b a c. Circle the letter of the response given. On the line beside each letter, list the order of presentation.

First name of child _____

Group _____

Sex _____

Responses:

Story 1:

a. _____

b. _____

c. _____

Story 2:

a. _____

b. _____

c. _____

Story 3:

a. _____

b. _____

c. _____

Story 4:

a. _____

b. _____

c. _____

Appendix E

Form for Teacher to Fill Out
One on Each Child
--adapted from Hoffman and Saltzstein (1967)

- I. If _____ (first name of child) found a quarter in the classroom and no one saw him/her, do you think he/she would (mark only one)
1. bring the quarter to you
 2. keep the quarter
 3. keep the quarter but show some sign of concern over doing this
- II. In your opinion, how does _____ (same child) typically react when "caught doing something wrong"? (mark two)
1. looks for someone else to blame
 2. where possible tries on his own initiative to rectify the situation
 3. cries, looks sad, seems to feel bad
 4. makes excuses
 5. accepts responsibility for what he has done
 6. denies he did it

Code for scoring:

- I. 1 - 3 points
2 - 1 point
3 - 2 points
- II. 1 - 1 point
2 - 3 points
3 - 2 points
4 - 2 points
5 - 3 points
6 - 1 point

Appendix F

Control Treatment Stories

- Austin, M. Poppet. In Childcraft: The how and why library (Vol. 2). Chicago: Field Enterprises Educational Corporation, 1976.
- Brown, M. W. Willie's pockets. In J. Hale & M. Johnstone (Eds.), Through golden windows: Good times together. Eau Claire, Wisconsin: E. M. Hale, 1958.
- Bryon, D. There was Tammie. In J. Hale & M. Johnstone (Eds.), Through golden windows: Good times together. Eau Claire, Wisconsin: E. M. Hale, 1958.
- Clark, M. The white goat. In J. Hale & M. Johnstone (Eds.), Through golden windows: Good times together. Eau Claire, Wisconsin: E. M. Hale, 1958.
- Haywood, C. Circuses are fun. In J. Hale and M. Johnstone (Eds.), Through golden windows: Good times together. Eau Claire, Wisconsin: E. M. Hale, 1958.

Experimental Treatment Stories

- Broken window. In S. W. Lindberg, The annotated McGuffey. New York: Van Nostrand Rheinhold, 1976.
- Golden rule. In H. H. Vail (Ed.), McGuffey's fourth eclectic reader (Rev. ed.). New York: American Book Company, 1920.
- Halvar's house. In Childcraft: The how and why library. Chicago: Field Enterprises Educational Corporation, 1976.
- No crown for me. In H. H. Vail (Ed.), McGuffey's third eclectic reader (Rev. ed.). New York: American Book Company, 1920.
- Thoughtless boys. In S. W. Lindberg, The annotated McGuffey. New York: Van Nostrand Rheinhold, 1976.

Appendix G

True-False Quiz for Control Treatment Stories

Story: "Poppet"

1. Poppet was a dog. (false)
2. Poppet wanted to hunt for a bear. (true)
3. Poppet went hunting with a bag. (true)

Story: "Circuses are Fun"

1. Betsy wanted to be a bareback rider when she grew up. (true)
2. Miss Grey's first-grade class decided to give a circus performance. (true)
3. The organ-grinder lost his dog. (false)

Story: "There Was Tammie"

1. Tammie was a cat. (false)
2. Tammie kept trying to follow the automobile. (true)
3. Tammie got lost and never was found. (false)

Story: "Willie's Pockets"

1. Willie's new suit had 20 pockets. (false)
2. Willie put a handkerchief in one of his pockets. (true)
3. At the end of the day, Willie had all of his pockets full. (true)

Story: "The White Goat"

1. Andrewshek liked to swing on the gate. (true)
2. Andrewshek's aunt's name was Sarah. (false)
3. The white goat climbed on top of the roof. (true)

Appendix G

True-False Quiz for Experimental Treatment Stories

Story: "No Crown for Me"

1. The person who told Susan to finish her work before she played was her father. (false)
2. The crown was made of violet flowers. (true)
2. What Susan did wrong was to deceive her grandmother. (true)

Story: "The Broken Window"

1. The boy who broke the window was named Samson. (false)
2. The window was broken with a snowball. (true)
3. George decided to pay for the window with his dollar. (true)

Story: "The Thoughtless Boys"

1. The first person who fell down was a farmer's boy who was not hurt. (true)
2. The man who hurt his ankle was running to the store to buy bread. (false)
3. Edward and William learned that playing tricks on people can bring much harm to them and to others. (true)

Story: "Halvar's House"

1. Halvar was a very small man. (false)
2. Halvar gave the stranger 7 fat goats for one skinny cow. (true)
3. Halvar's house was filled with happiness because Halvar was good to other people. (true)

Story: "The Golden Rule"

1. The Golden Rule says that you should always buy groceries on Saturday. (false)
2. Susan received too much money from farmer Thompson, but she kept it herself. (false)
3. Susan was happy because she had obeyed the Golden Rule. (true)

True-False Test for Treatment Stories

Instructions to be read to students: I am going to read you a sentence about the story you have just heard. If the sentence is right, or true, mark the circle next to the smiling face. If the sentence is false, or not true, mark the circle next to the frowning face.

1.



2.



3.

