

Spring 1986

Theory and research in social education 14/02

National Council for the Social Studies. College and University Faculty Assembly

Follow this and additional works at: https://digitalcommons.usf.edu/coedu_pub



Part of the [Education Commons](#)

Scholar Commons Citation

National Council for the Social Studies. College and University Faculty Assembly, "Theory and research in social education 14/02 " (1986). *College of Education Publications*. 49.
https://digitalcommons.usf.edu/coedu_pub/49

This Article is brought to you for free and open access by the College of Education at Digital Commons @ University of South Florida. It has been accepted for inclusion in College of Education Publications by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.

THEORY AND RESEARCH

in Social Education

Vol. XIV

No. 2

Spring 1986

**Beem and
Brugman**

The Effects of Values Development Lessons

Kelly

**Discussing Controversial Issues: Four
Perspectives on the Teacher's Role**

**Wilen and
Clegg**

**Effective Questions and Questioning: A
Research Review**

Diem

**Computers in a School Environment:
Social Consequences**

Leming

**Rethinking Social Studies Research and the
Goals of Social Education**

VanSickle

**Toward More Adequate Quantitative
Instructional Research**

Theory and Research in Social Education

Volume XIV

Number 2

Spring 1986



**College and University
Faculty Assembly**

TRSE is the official journal of the College and University Faculty Assembly of the National Council for the Social Studies. Published quarterly, it is a general review open to all social studies educators, social scientists, historians and philosophers. A general statement of purpose, and submission, subscription and advertising information may be found at the end of the journal. © 1986 by the College and University Faculty Assembly of the National Council for the Social Studies. All rights reserved.

Editor:

A. Guy Larkins
University of Georgia

Associate Editors:

Charles Berryman
Judith P. Goetz
University of Georgia

Book Review Editor:

C. Warren McKinney
University of Southern Mississippi

Editorial Board:

Mary Jo Brown
University of Georgia

Charles K. Curtis
University of British Columbia

Janet Eyler
Vanderbilt University

John Haas
University of Colorado

Richard C. Hunnicutt
Georgia Pacific Corporation

Sandra J. Lesourd
Washburn University

Jack L. Nelson
Rutgers University

James P. Shaver
Utah State University

William Stanley
Louisiana State University

**The College and University Faculty
Assembly Executive Committee 1985-86**

Chair: Millard Clements
New York University

Secretary: Suzanne M. Shafer
Arizona State University

Treasurer: Samuel R. Bell
Bradley University

1986 Program Co-Chairs:
Peter Martorella
North Carolina State University

Jack Nelson
Rutgers University

Members:
Susan Adler
Rockhurst College

Millard Clements
New York University

Gloria Contreras
University of Texas

Thomas Popkewitz
University of Wisconsin

Suzanne M. Shafer
Arizona State University

S. Samuel Shermis
Purdue University

William Stanley
Louisiana State University

Mary Kay Tetreault
Lewis and Clark College

Jane White
University of Maryland

**The National Council for the Social Studies
Officers 1985-86**

President: Paul Shires
Glendale-Nicolet High School
Milwaukee, WI

President-Elect: Jan Tucker
Florida International University

Vice President: Donald O. Schneider
University of Georgia

Executive Director: Frances Haley

Reviewers for this Issue of TRSE

The editors wish to express special appreciation to the following scholars who served as referees of manuscripts submitted.

Sally Oldham Buss, Mt. Clemens, Michigan
Cleo Cherryholmes, Michigan State University
Allison Gilmore, Mercer University
Jesse Goodman, Indiana University
John Haas, University of Colorado
James Hantula, University of Northern Iowa
Michael Hawkins, University of Georgia
Richard Hunnicutt, Georgia Pacific Corporation
Gary McKenzie, University of Texas
Warren McKinney, University of Southern Mississippi
William Pulliam, University of Delaware
James P. Shaver, Utah State University
Mark Shug, University of Wisconsin-Milwaukee
William Stanley, Louisiana State University
Elmer Williams, University of Georgia

Theory & Research in Social Education, ISSN 0093-3104, is published quarterly by the College & University Faculty Assembly of the National Council for the Social Studies. Membership is \$20/year, \$15 of the dues are allocated for subscription to Theory & Research in Social Education. Second class postage paid at Washington, DC and additional offices. Postmaster: send address changes to: Theory & Research in Social Education, 3501 Newark St. NW, Washington, DC 20016.

Volume XIV
Number 2
Spring, 1986

CONTENTS

A. L. Beem and D. Brugman	The Effects of Values Development Lessons on Pupils' Well-Being, Pleasure in School, Mutual Relationships and on Pupils' Valuational Behavior during Classroom Dialogues	97
Thomas E. Kelly	Discussing Controversial Issues: Four Perspectives on the Teacher's Role	113
James S. Leming	Rethinking Social Studies Research and the Goals of Social Education	139
William W. Wilen and Ambrose A. Clegg, Jr.	Effective Questions and Questioning: A Research Review	153
Richard A. Diem	Computers in a School Environment: Preliminary Report of the Social Consequences	163
Ronald L. VanSickle	Toward More Adequate Quantitative Instructional Research	171

The Effects of Values Development Lessons on Pupils' Well-Being, Pleasure in School, Mutual Relationships, and on Pupils' Valuational Behavior during Classroom Dialogues

A. L. Beem and D. Brugman
Leiden State University

Abstract

Values development was conceptualized as the systematic execution of four valuation processes by pupils. These processes were stimulated during especially designed lessons. Effects of values development lessons were investigated on pupils' self-report scores by a quasi-experimental pretest-posttest untreated control group design, and on pupils' valuation behaviors during values development dialogues in the experimental group. Experimental and control groups contained 27 and 21 classes respectively, with pupils' ages ranging from 10 to 15 years. Effects could be found only on behavioral measures. In the experimental group the predictability of scores from classroom behavior was also investigated. Pleasure in school and inner well-being in the classroom were predictable. It is suggested that effects on self-report measures were lacking because the negative aspects of classroom behavior outweigh positive aspects. Increasing the frequency of the valuation processes is recommended, for which a thorough teacher training program is needed.

Several approaches to values education have been introduced in the past two decades (cf., Hersh, Miller, & Fielding, 1980). The approach presented here, values development, aims at the improvement for four valuation processes. The purpose of the present study is to investigate the effects of a values development curriculum designed to stimulate the execution of the valuation processes, on (1) self-report measures of inner well-being, pleasure in school, and relations between pupils; (2) pupils' execution of the valuation processes during the lessons as measured by observations of their verbal behavior. Furthermore, the self-report measures and observations will be related.

Values Development

Values development is an educational application for classroom practice of valuation theory (Hermans, 1976). This theory describes the organization

and development of persons' experiential worlds. The basic concept of the theory is *value area*, which is defined as anything persons find of importance in their situation. The theory assumes that people, as organizers of their experiences, strive after congruence between experience, symbolization and behavior.

The development of value areas (hereafter to be called *values*, because of the familiarity of that term) results from a process of inquiry, which involves the systematic execution of the following valuation subprocesses (Van der Plas, 1981): (a) Expliciting one's values and feelings, e.g., "I hate having my biweekly reports controlled by my parents", (b) Expliciting values and feelings of other persons and institutions, e.g. "My parents trust my efforts at school." This process of cognitive and affective perspective-taking may heighten the educational relevancy of the valuation theory (Kohlberg, 1973, 1975; Rokeach, 1975), (c1) Relating values and feelings within one's own value system, e.g. "I cannot feel myself responsible, if I'm not trusted", (c2) Relating values and feelings of different value systems to each other, one of which may be one's own, e.g. "I don't understand how the school wants pupils to feel responsible. The school controls efforts of the pupils by asking their parents to sign the biweekly reports", (d) Validating values and feelings by acting on them, e.g. a proposal by pupils to protest against the biweekly reports and actually asking their parents to subscribe to this protest.

Values development is operationalized as the repeated execution of these processes in different social systems, followed by a more frequent and integrated execution. The aim of values development is the development and maintenance of a flexible system of values, a system that satisfies four epistemological criteria to evaluate a person's organization of his experiential world: comprehensiveness, parsimony, cohesiveness, and verifiability. A flexible system is supposed to be more open to incorporating new learning experiences, and to be accompanied by a positive inner well-being. Hence, in values development, as in the values clarification approach (Raths, Harmin, & Simon, 1966; Simon, Howe, & Kirschenbaum, 1972), the cognitive and the dynamic nature of valuing is stressed. However, there are also important differences between approaches (Beem & Brugman, 1983).

Lockwood (1978) and Leming (1981) have reviewed research in classrooms on the effects of values clarification and moral development. Values clarification does not in general produce effects on self-report measures (like self-esteem, self-concept, intraclass relations), but it does appear to affect positively some behavioral measures during values clarification lessons, and affects teachers' reports about pupils' behaviors.

The apparent inconsistency between effects on self-report and behavioral measures prompts the question whether the pertinent behaviors are related to self-report measures, which in general cover a wider range of situations. Hence, an important purpose of the present study is to investigate whether classroom behaviors during values development lessons, and especially

behaviors related to the valuation processes, are related to the self-report measures.

In addition to Lockwood's critical methodological comments, it should also be mentioned that these studies employed rather small samples, considering the fact that traditional analyses of classroom data are considered inadequate at the present moment. For example, the largest number of pupils participating in a study was 282, and individual raw scores were probably used in the analyses. This may of course reduce the reliability of the conclusions (Burstein, 1980). Therefore, a considerably larger sample was obtained for this study.

The results to be presented were obtained in the experimental phase of a research project phased according to a descriptive-correlation-experimental loop (Rosenshine & Furst, 1973). Results of the earlier phases are discussed in Van der Plas (1981).

Research Questions, Hypotheses, and Definitions

Three hypotheses will be tested in this study:

1. Values development lessons will increase valuation behaviors of pupils in classroom dialogues during these lessons.
2. Values development lessons will increase pupils' inner well-being in the classroom, at home, and with friends, their pleasure in school, and will improve the relations between pupils. Hereafter these measures will be called self-report measures.
3. The valuation behavior of pupils, and teachers' and pupils' related behavior will be related to the outcomes on pupils' self-report measures.

The second hypothesis will be tested by comparing an experimental and a control group in a quasi-experimental design. The first and third hypotheses are tested in the experimental group only. The valuation behaviors as measured are verbal expressions of pupils referring to the valuation processes. These observations were made only in the experimental group, i.e. classes receiving the values development lessons. The observation instrument is described in the Appendix; categories 14-17 refer to the valuation processes. Following Casteel & Stahl (1975), we assume that a verbal expression of a valuation is an indication of the occurrence of a process. The first hypothesis will be tested by comparing differences between mean scores of the experimental group on these categories in the first and last months of the experiment.

To test the second hypothesis, self-report measures were obtained. *Pleasure in school* refers to the attitude of pupils towards the school as an institution. *Relations between pupils* refers to perceptions of mutual friendship, acceptance, and willingness to help each other. *Inner well-being in the classroom, at home and with friends* refers to the extent to which personal experiences in these situations are positively evaluated. These measures were obtained as pretests and posttests in experimental and control group.

These variables were seen as operationalizations of the general well-being of persons (Van der Plas, 1981), assumed to be dependent on the flexibility of the values system. As noted above, the valuation processes aim at stimulating the flexibility of the values system, and hence the general well-being of pupils.

Like values clarification, it is assumed that values development needs a classroom climate of openness and an empathic teacher. Therefore, scores were obtained as pretests for the classroom climate variables *empathy of teacher*, *openness/diversity in the classroom*, and *friction in the classroom* to control for their possible influence. Openness/diversity in the classroom refers to the extent to which the opportunity exists for novel and unusual activities. Friction in the classroom refers to the existence of quarrels and conflicts. Empathy of teacher refers to the acceptance and understanding of pupils by the teacher. The second hypothesis was tested by comparing measures of experimental and control group in an analysis of covariance.

To test the third hypothesis, correlations were computed between the categories of the observation instrument and the posttest self-report measures. It was predicted that the categories 14–17 (see appendix) would have positive correlations with the self-report measures. In Beem and Brugman (1985) the structure of classroom behavior during the lessons was investigated to establish the relations between the valuation processes and other behavioral categories. From these results, it was predicted that the categories 5, 6, 7, 11, 12, 21, 22 and 1, 4 would have negative and positive correlations, respectively, with the self-report measures. Furthermore, multiple regressions were computed to select the most important predictors from the observation categories.

Before results pertaining to these hypotheses are presented, the research design and methods of data analysis will be discussed.

Method

Sample

At the start, the sample consisted of 33 classes (15 from elementary and 18 from secondary school) in the experimental group, and 21 classes (11 from elementary and 10 from secondary school) in the control group. Teachers of the classes in both groups were volunteers. The sample was not drawn at random nor was the assignment to experimental and control group randomized. Each class contained a majority of pupils of about the same age. In both groups the age of pupils from 10–15 years, with mean age 12. The total number of pupils was 850.

In the end, 27 (12 from elementary and 15 from secondary school) and 25 (12 from elementary and 13 from secondary school) classes were left of the experimental group, for which sufficient data were available to test hypothesis 2 and hypotheses 1, 3, respectively. In the control group all classes completed the experiment.

All schools were situated in or around Leiden.

Procedure

The design was a quasi-experimental pretest-posttest untreated non-equivalent control group design. The experimental group received a treatment; the control group received no treatment. Two lessons in October, administered in random order, were the same for all classes of the experimental group. The same applied to two lessons in March.

These lessons were videotaped and classroom dialogues during the lessons were subsequently scored by trained observers, using the observation instrument and a scoring apparatus to be described below.

Pupils in experimental and control groups answered the self-report and classroom climate measures in the following order: in September, the classroom climate variables empathy of teacher, openness/diversity in the classroom, and friction in the classroom; in the first week of October and at the end of March, inner well-being in the classroom, at home, with friends, pleasure at school and relations between pupils. The measurements in October and March are hereafter called pretests and posttests, respectively.

Treatment

The treatment consisted of values development lessons. In principle, these lessons were given bi-weekly from the beginning of October till March, but at least ten lessons were given during this period. A lesson lasted about one hour. Each time, the pupils together chose 1 of the 15 lessons from the workbook. The topics cover a substantial part of the experiential world of the pupils. The workbook contains so-called value sheets to stimulate values development.

The teacher's task was described in written instructional aids, based on theoretical considerations and results of the correlational phase of the project. The aids gave information about the goals of values development and suggestions for a sincere and lively performance during the lessons. Important skills (values and feelings in the classroom, redirecting, accepting, active listening and asking open questions, etc.) can be practiced with four exercise units.

The Observation Instrument

The observation instrument used to register classroom behavior consists of 22 categories, whose content and shorthand descriptions used in the text are described in the Appendix. The categories are mutually exclusive and exhaustive. Teacher categories are generally in line with the Flanders Interactions Analysis Categories (Flanders, 1970).

Category 14 (expressed personal values) refers to the first valuation process. Category 17 (refers to values of others) refers to the second valuation process. These are affective and social aspects of values development. Category 15 (personal behavior description) and 16 (gives opinion) refer to behavioral and cognitive aspects of these processes. Alternating occurrence of 14 and 17 refers to the third valuation process. The instrument is a final version developed from more detailed instruments, and does not contain a

category for validating values and feelings because this behavior was never scored in former observational studies. Category 15 may be relevant to validating values and feelings.

Scoring Apparatus

Observers scored classroom behavior in 22 categories by pressing buttons, numbered 1-9, on a keyboard. A category number was automatically linked to time in seconds from the start of the lesson. The observers could control and correct their scoring quite easily, because the category number selected was visible on the screen.

Reliability of the Observation Instrument

For each observer, the percentage of time during classroom dialogue that a class spent in a particular category of the observation instrument was computed. These proportions p_i were transformed by arcsine transformation: $2 \sin^{-1} p_i$. The mean was then computed over observers, both for the first two and for the last two lessons. (These averaged scores will be referred to as first and last month respectively).

To estimate reliability of the scores, a generalizability study was conducted (Cronbach, Gleser, Nanda & Rajaratnam, 1972). A sample of 10 classes, 3 observers and 3 lessons was randomly drawn to estimate variance components from a completely crossed design for the finite populations of 27 classes, 7 observers and 3 lessons. The Appendix contains the estimated reliabilities for mean scores over the same two lessons for each class, obtained from two observers, which were randomly allocated anew to each lesson (for more details see Beem & Brugman, 1985).

Elffers and Tavecchio (1979) suggest that the reliability should be at least .65, a value deduced from probabilities of wrong classifications. With the exception of categories 1, 17 and 21, the reliabilities exceed this value. In general, the reliabilities are quite satisfactory.

The self-report and classroom climate measures

The pupil self-report measures contain four-point Likert-type items, mostly ranging from *never* to *always*. Except for the inner well-being measures, the classroom climate and self-report measures were adaptations of instruments used in Great Britain and the United States (Finlayson, 1970; Moos & Trickett, 1974; Novotny & Overman, 1975; Walberg & Anderson, 1968). Construct validity was investigated during the correlational phases of the project using target matching (Harman, 1976, ch. 15). Congruence coefficients were .78 (relations between pupils), .72 (pleasure at school), .74 (empathy of teachers), .59 (friction) and .56 (openness/diversity). These values were deemed moderate, but sufficient for our purposes. Similar coefficients for the inner well-being measures were not obtained.

Social desirability was also measured to control for its influence if necessary, using a selection of items from a social desirability scale frequently used in the Netherlands.

In Table 1, one item for each variable and total scale reliabilities are presented. Scores were optimally scaled using the computer programs HOMALS (Gifi, 1981; Young, 1981).

Analysis

The mean score for classrooms was used as the unit of analysis, for theoretical and statistical reasons. We assume that the better the valuation processes proceed during classroom dialogues, the better these will proceed at the individual level, and vice versa. Therefore it is possible to speak of values development of a group. From this argument it follows that individual scores are not independent, which would make statistical analysis at the individual level complicated. Moreover, classes were sampled as intact groups. It also follows that mean scores are meaningful. If classroom scores were to be properly computed, however, the contribution of individual pupils should probably be weighted differentially.

Distributional assumptions were tested using tables from D'Agostino and Tietjen (1971, 1973) for the distribution of skewness and kurtosis in small

Table 1
Questionnaires with example of items, and pretest and posttest reliabilities of mean classroom scores.

Questionnaires	Example of items	Reliabilities	
		Pretest	Posttest
1. Relations between pupils (9 items).	My classmates are not nice to me.	.84	.85
2. Pleasure in school (8 items).	I think time passes quickly at school.	.86	.86
3. Inner well-being	I feel happy (in the classroom, at home, with friends).	.85	.83
a. in the classroom (10 items)		.73	.82
b. at home (10 items)		.85	.86
c. with friends (10 items)			
4. Empathy of teachers (10 items).	The teacher frankly admits if he does not know something.	.90	—
5. Openness/diversity (7 items).	We think of new things to do at school.	.84	—
6. Friction (8 items)	Some of the pupils quarrel with each other.	.94	—
7. Social desirability (7 items).	I never boast.	.74	.64

samples. A table in Hawkins (1980, p. 117) was used to test for the presence of outliers. The data proved to be reasonably well-behaved.

Results

The first hypothesis was tested with a *t*-test for paired observations. The results are presented in Table 2. The hypothesis was confirmed except for category 15. It can also be concluded from the values of *omega squared* (Hays, 1973, p. 417) that the changes are quite substantial.

The second hypothesis was tested by analysis of covariance. Except for inner well-being in the classroom, no other covariates had to be included for predicting a posttest than its own pretest. Openness/diversity in the classroom was substantially correlated with most posttests, but in the presence of the pretests its regression weight did not differ significantly from zero.

The null hypothesis of no difference between experimental and control group could not be rejected. The largest difference, for relations between pupils, was in favor of the control group; the descriptive significance level and the corresponding *omega squared* value (Keren & Lewis, 1979) were .09 and .06, respectively. Other significance levels and *omega squared* values were less than .56 and .01, respectively. Details are reported in Beem and Brugman (1983).

The third set of hypotheses concerns the relations between the posttests and category scores in the last month. The correlations are presented in Table 3.

The predictions are in general confirmed for pleasure in school, and are partly confirmed for inner well-being in the classroom. The predictions are not confirmed for relations between pupils, inner well-being at home and with friends. For the last two variables the correlations even suggest an inverse relationship.

The results of a stepwise regression to select important predictors of the posttests are presented in Table 4. The selection was made for both teachers' and pupils' categories. To reduce the influence of chance capitalization, a new variable, created with a uniform random number generator, was added

Table 2
Category means for two months, and omega squared.

Categories	First month	Second month	P-value*	Omega squared
14. Expresses personal values	.30	.40	0	.21
15. Personal behavior	.65	.61	.39	0
16. Gives opinion	.22	.30	0	.22
17. Refers to values of others	.17	.28	0	.32

*Descriptive significance level of *t*-test for paired observations.

Table 3
Correlations between Posttests and Behavioral Categories in Last Month

Categories	Pleasure in school	Relations between pupils	Inner well-being in the classroom	Inner well-being at home	Inner well-being with friends
1. Asks open questions	33	18	07	-33	-37
2. Asks closed questions	-09	11	-44	-20	-37
3. Probes	13	01	16	-24	30
4. Redirects	41	11	02	-39	-20
5. Gives direction	-25	11	06	31	21
6. Lectures	-14	14	26	09	-23
7. Being personal	-06	0	-01	-11	-28
8. Summarizes	-29	-15	-18	27	30
9. Accepts	17	06	-04	-21	-19
10. Praises	-16	-05	-22	33	21
11. Corrective feedback	-20	-07	-08	24	-20
12. Sharp negative feedback	-47	-22	-17	13	0
13. Irrelevant behavior	04	-04	02	-06	04
14. Expresses personal values	53	12	20	-06	04
15. Personal behavior	55	14	37	-25	01
16. Gives opinion	33	12	21	-25	-21
17. Refers to values of others	28	-16	-05	-02	11
18. Gives information	17	03	-01	10	35
19. Asks questions	13	41	-04	35	0
20. Structured working	-15	02	-26	22	-10
21. Unordered behavior	-40	-05	-33	20	-17
22. Silence	-37	04	-55	25	01

Note. Decimal point omitted.

Table 4
Categories stepwise selected for the Prediction of Posttest from Observation Categories in Last Month

Teachers' categories								
Questionnaires	Pleasure at school			Inner well-being in the classroom				
	$R^2 = .55; AR^{2b} = .43, p < .01$			$R^2 = .52; AR^2 = .36; p < .03.$				
Selected categories	B	T	P	Selected categories	B	T	P	
2. Asks closed questions	-1.80	-2.10	.05	2. Asks closed questions	-1.61	-3.37	.00	
4. Redirects	3.09	3.01	.01	3. Probes	.54	1.29	.21	
8. Summarizes	-1.34	-2.23	.04	6. Lectures	.75	2.30	.09	
12. Sharp negative feedback	-1.15	-2.32	.03	7. Being personal	-.41	-1.91	.07	
13. Irrelevant behavior	1.26	1.49	.15	8. Summarizes	-.59	-1.69	.11	
				12. Sharp negative feedback	-.47	-1.44	.17	
Pupils' categories								
	$R^2 = .48; AR^2 = .38; p < .01$			$R^2 = .46; AR^2 = .37; p < .01.$				
Selected categories	B	T	P	Selected categories	B	T	P	
14. Expresses personal values	1.02	1.58	.13	20. Structured working	-.49	-1.83	.08	
15. Personal behavior	.40	.78	.45	21. Unordered behavior	-.43	-1.78	.09	
16. Gives opinion	1.07	1.56	.13	22. Silence	-.72	-2.92	.01	
19. Asks questions	3.21	2.01	.06					

Note. ^aSquared multiple correlation

^bSquared multiple correlation adjusted for shrinkage.

to the set of independent variables for each dependent variable separately. A category was not included in the regression equation if it entered the equation after the random variable and if it did not raise the amount of explained variance by more than five percent. Results for relations between pupils, inner well-being at home and with friends are not presented, because these variables could in general not be predicted adequately. The squared (adjusted) multiple correlations with inner well-being in the classroom and pleasure in school are quite satisfactory. The signs of the regression weights also conform to the expectation, except for category 6.

Pleasure in school is the only variable which has among its most important positive predictors pupils' categories related to values development (categories 14, 15, and 16). This also expresses itself in the positive weight of category 4, which refers to teacher behavior related to values development categories. The majority of the other predictors have a negative regression weight.

The regression equations partially explain the results of the analyses of covariance. It appears that for some categories with negative regression weights (categories 2, 8, 12, and 22) the frequency of occurrence increases from the first to the last month (Beem & Brugman, 1985). These effects counterbalance positive influences.

Summary and Discussion

In this paper effects of values development lessons were assessed. Three hypotheses were formulated: (1) the lessons would result in more frequent verbal behavior referring to three valuation processes; (2) the experimental group, which received as treatment values development lessons during several months, would have higher posttest scores than an untreated control group on self-report measures of pleasure in school, relations between pupils, inner well-being in the classroom, at home and with friends; (3) values development related behavior would be positively related to the self-report measures.

Verbal behavior by pupils and teachers was registered with an observation instrument of which four observational categories referred to the valuation processes. Except for one of these categories, the frequency increased significantly from the first to the last month of the experiment. Hence the first hypothesis was largely confirmed. The second hypothesis could not be confirmed. The largest difference, for relations between pupils, was in fact in favor of the control group, although it was not statistically significant.

Relationships between self-report scores and classroom behavior during the lessons in the last month of the experiment were also investigated. Correlations between pleasure in school and observation categories were generally as expected and substantial. The expectations were not confirmed for relations between pupils, inner well-being at home and with friends, and were confirmed to a lesser extent for inner well-being in the classroom since several correlations were either in an unexpected direction or not substantial.

Using multiple regression techniques, the question was also investigated whether the self-report measures were predictable from the verbal behavior of pupils and teachers separately. Only pleasure in school and inner well-being in the classroom appeared to be predictable. However, only pleasure in school had among its most important predictors observation categories which refer to the valuation processes.

Of course, differences between posttest means in a quasi-experimental design may be attributed to the nonrandom assignment to experimental and control groups. Results of a multivariate analysis of variance showed, however, that experimental and control group mean scores on self-report pretests and classroom climate variables did not differ significantly ($p < .43$).

As noted above, it is not uncommon that effects can be demonstrated on behavioral variables but not on self-report measures (Leming, 1981; Lockwood, 1978). The results of the correlation and regression analyses indicate that for three of the self-report measures this can be explained by the fact that these measures are not substantially related to the execution of the valuation processes. Thus, effects of values development lessons do not appear to generalize to situations outside the classroom.

Inner well-being in the classroom is predictable from behavioral categories which are primarily negatively related to the valuation processes (Beem & Brugman, 1985). Although these categories indicate that pupil initiative is not encouraged, the behaviors they refer to may well be conducive to the acquisition of cognitive abilities during regular lessons (cf., Gage, 1984). This suggests that the predictability results largely from other lessons than the values development lessons.

Hence it appears that pleasure in school may be the only variable which is influenced by the execution of the valuation processes. Since a significant increase in frequency of the execution of the processes does not result in significant differences between experimental and control groups, the increase in frequency may simply be too low. It should be noted that the affective and social aspects of the processes (categories 14 and 17) have an especially low frequency as compared to other behaviors. Pleasure in school is actually the only variable for which the adjusted posttest mean is slightly higher in the experimental group than in the control group. Therefore, possible influences may be negligible as compared to that of other classroom behaviors. However, another explanation of the relation between this variable and the processes cannot be ruled out. Pleasure in school may simply capture aspects of classroom atmosphere which facilitate the execution of the processes.

Whether the first interpretation is correct can only be tested in an experiment in which the treatment is more powerful. This may be accomplished by selecting exercises to increase the frequency of the valuation process during values development and/or content-centered lessons. Attention should be especially paid to the process of validating values and feelings. Also,

thorough training of teachers may increase the frequency of the processes. If a more powerful treatment can be accomplished, the influence of other aspects of the classroom atmosphere might be comparatively reduced.

We conclude that only one self-report measure appears to be related to the valuation processes, and evidence for the relationship is weak. Perhaps the ambitious and abstract ultimate aims of values education are not measurable effects of short-term programs. Hence, the current research is still at a stage of exploring measurable intermediate effects which may be consistent with those ultimate aims.

References

- Beem, A. L., & Brugman, D. (1983). The effects of values development lessons on classroom behavior and pupils' outcomes as measured by questionnaires, and the relations between classroom behavior and those outcomes. *Internal Report*. Leiden: LICOR.
- Beem, A. L., & Brugman, D. (1985). An exploration of the structure of classroom behavior during values development lessons. *Studies in Educational Evaluation*, 11, 339-357.
- Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. In Berliner, D. (Ed.), *Review of research in education: Vol. 8*. Washington, DC: American Educational Research Association.
- Casteel, J. D., & Stahl, R. J. (1975). *Value clarification in the classroom: A primer*. Santa Monica, CA: Goodyear.
- Cronbach, L. J., Glaser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- D'Agostino, R., & Tietjen, G. L. (1971). Simulation probability points of b for small samples. *Biometrika*, 58, 669-672.
- D'Agostino, R., & Tietjen, G. L. (1973). Approaches to the null distribution of b. *Biometrika*, 60, 169-173.
- Elffers, H., & Tavecchio, L. W. C. (1979). *Variance components in test generalizability research: Which, when, why?* V. O. R.-publikatie Vereniging voor Onderwijsresearch.
- Finlayson, D. S. (1970). *School climate index*. The Mere, Upton Park, Slough: NFER.
- Flanders, N. A. (1970). *Analyzing teaching behavior*. Reading, MA: Addison-Wesley.
- Gage, N. L. (1984). What do we know about teaching effectiveness? *Phi Delta Kappan*, 66, 87-93.
- Gifi, A. (1981). *Non-linear multivariate analysis*. (Prelim. ed.). Leiden, The Netherlands: Department of Data theory, University of Leiden.

- Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press.
- Hays, W. L. (1973). *Statistics for the social sciences* (2nd ed.). London: Holt, Rinehart & Winston.
- Hawkins, D. M. (1980). *Identification of outliers*. London: Chapman & Hall.
- Hermans, H. J. M. (1976). *Value areas and their development*. Lisse, The Netherlands: Swets & Zeitlinger.
- Hersh, R. H., Miller, J. P., & Fielding, G. D. (1980). *Models of moral education*. New York: Longman.
- Keren, G., & Lewis, C. (1979). Partial omega squared for anova designs. *Educational and Psychological Measurement*, 39, 119-129.
- Kohlberg, L. (1973). The child as a moral philosopher. In H. Kirschenbaum & S. Simon (Eds.), *Readings in values clarification*. Minneapolis: Winston Press.
- Kohlberg, L. (1975). The cognitive-developmental approach to moral education. *Phi Delta Kappan*, 56, 670-677.
- Leming, J. S. (1981). Curricular effectiveness in moral/values education: A review of research. *Journal of Moral Education*, 10, 147-164.
- Lockwood, A. L. (1978). The effects of values clarification and moral development curricula on school-age subjects: A critical review of recent research. *Review of Educational Research*, 48, 325-364.
- Moos, R. H., & Trickett, E. J. (1974). *Manual: Classroom environment scale*. Palo Alto, CA: Consulting Psychologists Press, 1974.
- Novotny, J. M., & Overman, B. C. (1975, April). *Measuring classroom affect*. Paper presented at the meeting of American Educational Research Association, Washington, DC.
- Raths, L. E., Harmin, M., & Simon, S. B. (1966). *Values and teaching*. Columbus, OH: Charles E. Merrill.
- Rokeach, M. (1975). Toward a philosophy of value education. In J. R. Meier, B. Burnham, & J. Cholvat (Eds.), *Values education*. Waterloo, Canada: Wilfrid Laurier University Press.
- Rosenshine, B., & Furst, N. (1973). The use of direct observation to study teaching. In R. M. W. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally.
- Simon, S., Howe, L., & Kirschenbaum, H. (1972). *Values clarification, a handbook of practical strategies*. New York: Hart.
- Van der Plas, P. (1981). *Waardenontwikkeling in het onderwijs* (Values development in education) 's Gravenhage, The Netherlands: Stichting voor Onderzoek van het Onderwijs S. V. O. (With a summary in English).
- Walberg, H. J., & Anderson, G. J. (1968). Classroom climate and individual learning. *Journal of Educational Psychology*, 59, 414-419.

Appendix

The Values Development Interaction Category System

Reliability Teacher talk categories

- .55 1. *Asks open questions*: asks relatively open-ended questions which call for unpredictable, person-specific response, about a new topic, subject-matter.
- .84 2. *Asks closed questions*: asks questions requiring predictable response, short reply or yes-or-no answer, about a new topic. Response can often be judged as right or wrong.
- .91 3. *Probes*: Asks new, often more specific question about the same topic, mostly of the same pupil.
- .73 4. *Redirects*: asks same question about same topic, of another pupil.
- .79 5. *Gives direction*: all directions, commands, procedures or orders which a pupil is expected to comply with.
- .82 6. *Lectures*: factual information to clarify a topic, procedure.
- .98 7. *Being personal*: personal values and feelings, situations, opinions (cf. categories 14, 15, 16).
- .67 8. *Summarizes*: tries to state the crux of the whole matter based on one or more pupil statements. If the teacher brings his own opinions or judgments into play, then shift to the feedback categories 10, 11, 12 or category 7.
- .71 9. *Accepts*: accepting response (hm, hm; yes) repeating pupil statements.
- .83 10. *Praises or encourages*: positive, enthusiastic response. Direct motivating and supporting response.
- .78 11. *Gives corrective feedback*: correcting response about behavior of a pupil with the intention to change pupils' behavior from non-acceptable to acceptable by reminding the pupil of accepted rules or creating a new rule, agreement.
- .91 12. *Gives sharp negative feedback*: correcting response about pupil or pupils behavior with the intention to change pupils behavior from non acceptable to acceptable. Note the intonation.
- .73 13. *Shows irrelevant behavior*: other teacher talk, all teacher talk that does not fit in one of the aforementioned categories.

Pupil talk categories

- .92 14. *Expresses personal values and feelings*: utterance in which a pupil explicitly expresses a feeling, an appraisal.
- .94 15. *Describes personal behavior and situation*: talking about own behavior, situations the pupil was confronted with, without referring to feelings, appraisals, opinions.
- .73 16. *Gives opinion*: own opinion, judgment without making feelings verbally explicit.
- .49 17. *Refers to values and feelings, behaviors, situation descriptions, opinions of others*: repeats, summarizes, reflects values and feelings of others.

Reliability Teacher talk categories (continued)

- .84 18. Gives information: see category 6.
- .86 19. Asks questions: any question posed to teacher or other pupils.
- .69 20. Works, other pupil talk: all task oriented behavior that does not fit into one of the aforementioned categories.
- .58 21. Disrupts: considerable noise of some or all pupils which disrupt planned activities.
- .71 22. Silence during classroom dialogues.
-

NOTE: In the text, the categories are referred to by the underlined parts of the descriptions.

Discussing Controversial Issues: Four Perspectives on the Teacher's Role

Thomas E. Kelly
John Carroll University

Abstract

Much debate exists over the role teachers should assume in the discussion of controversial issues. This paper presents and critiques four perspectives on that debate. The perspectives are characterized as exclusive neutrality, exclusive partiality, neutral impartiality and committed impartiality. Synthesizing research and scholarship from a number of fields, the paper proposes committed impartiality as the most defensible teacher role.

From all points along the ideological spectrum teachers have been criticized for the values they are or are not transmitting. From the fundamentalist and political right, they are accused of inculcating a malignant secular humanism while commanded to inculcate the superiority of American capitalism and representative democracy. From the liberal and radical left, teachers are charged with perpetuating norms of racism, sexism, and cultural and economic domination. Seeing traditional values of equality and mobility as meritocratic myth, selective proponents of this critical perspective exhort teachers to become advocates for individual emancipation and social justice, while exposing the hegemonic distortions, inherent contradictions and structural inequalities which radicals claim characterize capitalist societies such as the United States.

Debates over the nature of social reality and the character of a decent society, and the school's role in uncovering and advancing these visions, strike at the heart of the social studies. Should social studies educators adopt a particular partisan position on such crucial issues as nuclear arms, religion in the schools, just principles of economic distribution, legitimate national security, or women's reproductive rights? If so, under what conditions and with what implications for the very language in which these issues should be framed (e.g. the child/fetus' right to live or the woman/mother's right of choice and control over her own body)? Hearing sirens from a different direction, should social studies educators heed those who advise steering a more neutral course between the Scylla of indoctrination and the

Charybdis of ethical relativism? What do such admonitions mean in practice? Or should social studies educators reassess the priorities of their discipline and support those who insist that sites other than schools are the proper places for the explicit transmission of cherished if controversial values?

How should the responsible social studies educator adjudicate these competing ideological and pedagogic demands? To assess these pressing questions, I will identify and critique four positions educators may assume in handling controversial issues. Borrowing the labels from Hill (1982) but substantially expanding upon his treatment, I will refer to these positions as exclusive neutrality, exclusive partiality, neutral impartiality and committed impartiality. While each has merit, my analysis will lead me to argue that the paradoxical position of committed impartiality is the most defensible one for teachers to uphold.

Exclusive Neutrality

Description

Advocates of this position contend that teachers should not introduce into the curriculum any topics which are controversial in the broader community. Schools have an implicit obligation to serve equally their varied publics. The mere inclusion of controversial topics, however, is likely to violate this tacit contractual agreement for several reasons: because it is difficult to give a fair or impartial hearing to all points of view; because the determination of particular personal, religious or political value positions is the sovereign task of other socializing institutions (Bereiter, 1973); and/or because classroom discussion of genuinely provocative values, in its inevitable volatility and unpredictability, undermines institutional norms of order on which teachers are evaluated (McNeil, 1983). Hence it is most appropriate that schools exclude these issues at the outset. Such exclusion allegedly preserves the nonpartisan or neutral status of school. Instead, schools should stick to the value-free teaching of that knowledge and set of skills which have been conclusively demonstrated to be true or important through rigorous scientific investigation or through broad consensus within the community.

Critique

The myth of a value-free education. While there is legitimate debate over what should be included in the school curriculum and what should be addressed by other institutions (Coleman, 1972; Goodlad, 1984; Newmann, 1977b), the position of exclusive neutrality seems both untenable and undesirable. The view that scientific discoveries and scientific methodologies inherently represent value-free truths and technologies—a tradition known as positivism—has been progressively and persuasively debunked by a host of scholars (Apple, 1979; Bernstein, 1978; Habermas, 1971; Kuhn, 1970; Lucas,

1983; Matthews, 1980; Polanyi, 1974). The similarity between teachers and scientists in this context is noteworthy. In the curricular or research goals they pursue and exclude, in the methods of instruction or investigation they use and reject, and in the presuppositions and consequences each embodies, teachers, like scientists, act in a value-infused context. The impossibility of neutrality as well as the likelihood of controversy at the level of concrete school experience is nicely captured in the following:

Values are taught by all teachers. Some teachers teach language arts, other social studies, others math or science, but all teachers teach values. Values are taught whenever an adult stands before children and acts, speaks, and reveals his convictions. Every teacher teaches something about values by the example he sets. When the teacher reveals the measure of his commitment to teaching by the care he takes in preparation, he teaches or misteaches his students about responsibility. By the amount of time he takes from class time to complain about other teachers or administrators in the school district, he teaches or misteaches the students about work. In the effort he shows toward student work, toward correcting students when they are wrong, he teaches them about what they can expect from the world beyond the schoolhouse door and he teaches or misteaches them about character, patience, and candor. By how he talks about the government and local elected officials and the laws, he teaches or misteaches students about citizenship. These "values" lessons occur in all classrooms and they are taught as well in places where athletics, debate, honor society, and newspaper work are done (Kohlberg, quoted in Sichel, 1982).

Research and other commentaries on the hidden curriculum (Anyon, 1980; Apple, 1979; Apple & Weis, 1983; Brophy, 1979; Rosenbaum, 1976; Rutter, Monghan, Mortimore, Outson, & Smith, 1979; Young, 1971) have repeatedly demonstrated that while teachers may choose to avoid explicit discussion of controversial issues, they cannot avoid practices which generate unintended and often provocative impacts on students. These outcomes involve such significant domains as self-esteem, future aspiration, the purpose of schooling, the nature of social and spiritual reality, and definitions of legitimate knowledge, work, authority and conflict.

Schools as tenable sites for a meaningful, civic education. Though it is impossible for teachers and curriculum to remain neutral in the sense of being value free and uncontroversial, it is not impossible for teachers and students to address value issues in a fair, impartial manner. Elaborated below in terms of a fair hearing and genuine dissent, the norm of impartiality is a tenable achievement in classrooms. While a lively discussion of a controversial issue may be more challenging to manage than a fact-oriented, tightly controlled teacher presentation, the former is by no means inherently chaotic or biased. Specific group discussion skills, instrumental to realizing

the norm of impartiality, are identifiable and teachable (Lockwood & Harris, 1985; Newmann & Oliver, 1970).

Not only is the fair handling of controversial issues in classrooms a feasible enterprise; it is also a compelling one. Two reasons are germane. The first, advocated by representatives of diverse political persuasions (Cagan, 1978; Giroux, 1983a; Goble & Brooks, 1983; Newmann, 1975; Will 1983) is that schools, particularly those publically financed and state supported in a democracy, have a moral responsibility to develop in their charges the understandings, competences and commitments to be effective citizens. Among other attributes, these broad goals include the ability to make informed judgments about public issues and to debate these judgments in a reasoned, if passionate, manner. In order to achieve these democratic goals, teachers need to include rather than exclude public controversy.

A second argument is more psychological in nature. Important controversial issues, by definition, reflect genuine concerns within a community which, either directly or indirectly, affect students. A school policy which preemptively excludes discussion of these real concerns communicates the irresponsible message that school work has no vital relationship to the real issues of life. Rather than serving as an opportunity to promote self-determination and a meaningful integration of school and community, a curriculum systematically devoid of genuine value controversies seems misguidingly conceived as a formula for intellectual sterility, leading to alienation from self, school and community (Newmann, 1981; Newmann & Kelly, 1983).

In sum, the search for a curriculum which maintains exclusive neutrality toward controversial issues is a futile and misguided enterprise. Values and value controversies inevitably permeate and rightfully belong in the school curriculum. What then should the teacher's role be?

Exclusive Partiality

Description

This position is characterized by a deliberate attempt to induce students into accepting as correct and preferable a particular position on a controversial issue through means which consciously or unconsciously preclude an adequate presentation of competing points of view. In the more authoritarian forms of exclusive partiality, teachers assert or assume the correctness of a particular point of view while competing perspectives are ignored, summarily dismissed, or punitively downgraded. Whether the advocacy and dismissal is done passionately or matter of factly, haphazardly or more systematically, the sum effect is a one-sided presentation where challenge to the preferred point of view is discouraged or precluded. In more subtle forms of partiality, the teacher may give the appearance of permitting genuine dialogue and dissent yet nevertheless attempt to stack the deck. These tactics might include selecting the most articulate or esteemed students to

represent the preferred position in a debate; deliberately inviting to class a representative of an alternative view whose personality or manner of presentation is likely to offend students or obscure the issues; using materials which are intentionally, though not flagrantly, weighted for the desired position; selectively praising responses supporting the preferred position without close attention to the merits of the students' actual contribution. Regardless of its form, the net effect of exclusive partiality is that advocacy tends to subsume and subvert autonomous rational critique rather than be its natural consequence.

Why do the strict partisans of exclusive partiality adhere to their position? Several explanations can be identified. For some, particular controversial issues are personally felt to be unproblematic. These individuals may not be adequately aware of contending positions or they may see them as evidence of prejudice or ignorance. In either case, they view truth on the matter to be self-evident, appropriately taught in a more or less straightforward, factual manner, something akin to Howard Cosell's "telling it like it is." Thus, without feeling or conveying the ambivalence or ambiguity others confront, these individuals can profess such unequivocal truths as abortion is murder; creationism is pure myth; Marxist revolutionary governments are merely Soviet puppets; capitalists are pigs; all Americans have equal rights; lying is always wrong.

Others believe that it is their paternalistic, moral or contractual duty to pass on the dominant religious, social or political values of the institutions in which they work. Whether true believers or loyal servants of that institution, they conceive their purpose as instilling in their charges the rightness of a certain set of beliefs and behaviors. Hence, public and private school teachers should promote students' commitment to principles of democracy and free enterprise just as parochial schools teachers should ensure students' commitment to the sacred doctrines of the particular religious denomination. While rational inquiry into values may not be considered irrelevant, a number of powerful forces are seen to make impartial examination of alternative positions exceedingly problematic if the ultimate goal of durable allegiance is to be achieved. Adherents of this position are alarmed by a culture where optimizing individual choice is king, where knowledge appears inconclusive in determining value choices, where statistics seem capable of supporting contradictory judgments, where ubiquitous and intrusive mass media can create and seduce impressionable minds. In short, in a culture where knowledge can confuse more than clarify, contaminate more than liberate, some believe that students need to be shielded from systematic exposure to potentially harmful alternative perspectives.

From a third perspective, exclusive partiality is seen as a necessary corrective to the pervasive distortions perpetuated by dominant social norms and practices. This perspective is held by various individuals advocating feminism, ethnic and black empowerment and neo-Marxist social recon-

struction. Often invoking the sanction of academic freedom and/or alerting students at the onset to their oppositional perspective (Berlak, 1985), these individuals essentially construe exclusive partiality as a warranted form of ideological affirmative action. Their views are reflected in the following perspective:

The people are indoctrinated by the conditions under which they live and which they do not transcend. To enable them to become autonomous, to find by themselves what is true and what is false for man in the existing society, they would have to be freed from the prevailing indoctrination . . . But this means that the trend would have to be reversed: they would have to get information slanted in the opposite direction. (Marcuse, quoted in Cohen, 1982, p. 89.)

To slant information in the opposite direction, teachers need systematically and exclusively to expose students to perspectives critical of mainstream values. Only a concentration of oppositional ideology will allow teachers to unearth the deep seeded roots of indoctrination in which common sense yet mystifying conceptions of social reality are grounded. In short, to advance the autonomous and rational pursuit of truth, prior indoctrination compels current counter indoctrination. Put in a different metaphor, we are unwitting captives of a pervasive, dimly understood programming. A radical and sustained process of deprogramming, very partisan in nature, becomes imperative if the urgent, emancipatory restructuring of consciousness and behavior is to be accomplished.

Critique

In certain respects, the different practitioners of exclusive partiality can be viewed with both sympathy and support. For example, amidst the incessant pelting of diverse values, it is understandable, if not condonable, that teachers feel compelled to provide students emotional shelter and the semblance of intellectual certainty. On the other hand, insisting that a commitment to truth and human emancipation obligates the penetration not the promotion of student insulation, radical left critics have developed insightful theories which dialectically interrelate dynamics of power, class, gender, race, culture and education (Anyon, 1980, 1984; Apple, 1982; Arnot, 1981; Cherryholmes, 1980; Freire, 1970; Genovese, 1972; Giroux, 1983a; Willis, 1981). These theories, while imperfect, represent a most significant challenge to mainstream conceptions of social reality and distributive justice. Disturbingly, theirs is a perspective which has too long been neglected or misrepresented in the traditional curriculum (Apple & Weis, 1983; Newmann, 1985; Taxel, 1984).

More generally, assuming a marketplace of diverse ideas, teacher preference (partiality, if you will) in the selection of curriculum meanings, materials and methods, short of outright deceit, is not necessarily incompatible with the cherished professional value of academic freedom (Macklin, 1980).

Despite a certain allure of this position, it seems misguided on a number of grounds. Below I address its intellectual, practical and moral shortcomings.

Intellectual grounds. A primary task of the intellect is to search for truth in all its complexity and subtlety. Rational inquiry by which this quest is pursued demands critical and continuing openness to the best evidence and argumentation available. When through intention or negligence teachers fail to inform their students about the complexity and diversity of perspectives on relevant issues, they severely undermine the vital mission of the intellect. Figuratively speaking, rather than fulfilling their role as educational Sherpa, helping students master and preserve a challenging topography, the practitioners of exclusive partiality effectively do neither. Whether by flattening the terrain of social reality or by obscuring the detours or alternative paths to the summit, these fraudulent guides attenuate rather than strengthen students' intellectual orienteering skills. As a result, students are equipped with the illusion rather than the reality of discovery and achievement.

Instrumental grounds. Not only is exclusive partiality a capital assault on the integrity and purpose of the intellect, it is also presumptuous and likely self-defeating in practical or instrumental terms. Instrumentally, the partisan's goal of durable allegiance to a particular doctrine is best realized when positions are informed, defensible and personally attractive (Straughan, 1982). By dint of their methods, however, these strict partisans preclude the opportunity for realizing an informed and defensible position. Hence, by failing to expose students to the best alternative arguments and how presumably these can be effectively rebutted, the strict partisan is essentially fabricating intellectual straw people vulnerable to the serious challenges from competent adversaries. In a tightly closed society where conflicting perspectives are suppressed or nonexistent, the strict partisan's approach to preserving allegiance is perhaps feasible. In a more pluralistic culture, however, a cocooned existence is constantly in jeopardy; as a result, the presumption that individuals can be protected from multiple influences is dubious, if not myopic. More likely, the fragile proteges of the strict partisan face any of several ominous futures: defensive dogmatism, disillusionment and paralysis, ethical relativism, or defection to the enemy camp. While the most appealing of these, dogmatism, may look and feel familiar to the strict partisan, it is an improbable strategy for durably seducing or converting an inquiring mind.

Revealed in Berlak's (1985) important research on liberatory pedagogy, certain members of the radical left appear subject to a similar criticism. While not clearly representative of a homogenous left perspective on pedagogy¹, these individuals tend to take as given the exploitive character of capitalist culture and argue that the pervasiveness of that culture liberates them from any necessary systematic coverage of its strengths. On the con-

trary, to provide students equal ideological time in the broader context of their lives, these teachers feel free or obliged to concentrate more exclusively on critiquing mainstream culture's weaknesses and exploring emancipatory alternatives (Berlak, 1985). This position, however, denies the incisiveness of their own critique: namely, that the permeation of the dominant ideology (i.e. hegemony) functions most effectively at the tacit or taken for granted level. The implication here is that it is presumptuous to assume that without explicit attention students can necessarily articulate the best case for capitalist values. As argued above, failure to encourage such a capacity is potentially to undermine the durability of any new adherent's allegiance to a preferred course, e.g. democratic socialism. In addition, for radicals supporting this form of partisan pedagogy, the authenticity of their professed commitment to critical literacy becomes problematic.

The meaning and implication of the term authenticity needs elaboration. Defined as the congruence between goals and practice (Newmann, 1977a), a quest for authenticity becomes crucial in light of the recurrent hidden curriculum findings that students are influenced by what teachers do in practice as much, if not more so, as by the vision teachers espouse, however compelling. Hence, despite the best of intentions, inauthentic pedagogy can subvert the perceived credibility of teachers and their goals. These are troubling implications for those teachers who practice exclusive partiality while professing a commitment to democracy. William Kilpatrick, a central figure in the great indoctrination debates of the 1920s and 1930s captures the several major points made in this sub-section.

To teach democracy . . . [so as] to foster uncritical acceptance would seem an odd way of fostering democracy. To indoctrinate a belief in democracy without including the reasons . . . and without building ability to think critically about it, is to make blindfolded adherents . . . Such people would not know the why of their practices of dogmas and consequently could not be trusted to apply the doctrines intelligently. When they grow up into active citizenship they might be easily induced, for example, to forbid the study of controversial issues in school. They might forbid the critical study of democratic doctrines and so prevent wise adaptation of these doctrines to new conditions . . . In one word, such indoctrination would make blind dogmatists. . . quite unfit to carry on the democratic process in a changing civilization. That way lies fanaticism (Raywid, 1980).

Moral grounds. Even if, on instrumental grounds, exclusive partiality were successful in inducing students to hold preferred beliefs, the practice would be nonetheless objectionable on moral grounds. Two arguments are central. First, the practice of exclusive partiality abridges students' human dignity by violating the Kantian imperative to treat people as ends in themselves and not merely as a means to someone else's ends. It does so by in-

fringing students' rights and their opportunity to make informed, independent judgments after due consideration of alternatives. In effect, strict partisans cast the student into the role of pawn in the perpetuation of their own ideological agenda.

Relatedly, exclusive partiality undermines students' identity as moral agents. Because ought implies can (Kohlberg, 1981), moral agency assumes that students have the necessary freedom and ability to control their own beliefs and behaviors. However, in seeking durable allegiance, the logic, if not the effect of the strict partisan's goals and methods, is to have students' beliefs and behaviors become resistant to ongoing scrutiny, regardless of the merits of competing arguments. In succeeding in this endeavor, the strict partisan's victory becomes the students' incapacitation. To the extent that students are unable progressively to reconstruct prior learnings in the light of new experience, they cannot act freely as the directive guardians of their own lives (Dewey, 1944; Raywid, 1980). However, as has been effectively argued (Dewey, 1944; Newmann, 1975), to empower students with this capacity, appropriately directed toward advancing the public interest, is precisely the primary goal of education in a democracy.

In summary, the posture of exclusive partiality is not without persuasive force. However, its intellectual, instrumental and moral shortcomings necessitate a greater commitment to an alternative ideal and practice, that of impartiality. Impartiality may take two quite dissimilar forms. These are discussed and critiqued in the following sections.

Neutral Impartiality

Description

A third position teachers might assume regarding controversial issues is that of neutral impartiality. The advocates of neutral impartiality differ from proponents of exclusive neutrality in their belief that students should be actively involved in discussions of controversial public issues as part of their education for citizenship. On the other hand, they differ from proponents of exclusive partiality in both the procedural ideals to which they subscribe and the role their own views should play in the discussion of values. The rationale for addressing controversy in citizenship education was noted in an earlier section. Explication and critique of neutral impartiality follow.

As an ideal in discussing controversial value issues, impartiality entails the related principles of a fair hearing and critical dialogue. In striving to insure that students have the opportunity to consider all relevant positions on an issue, teachers will attempt to honor the best case standard wherein the strongest arguments for competing points of view are presented and critiqued. Figuratively this standard would be achieved if the most sophisticated competing advocates on an issue could determine that their views received as fair and accessible a rendering as possible, given the developmental nature of the student population.

Diverse instructional experiences could be instrumental to achieving this standard of a fair hearing. These could include a combination of library research, guest speakers, films, role playing, community interviews, didactic presentations and group discussions. Overall, the teacher seeks to promote a classroom atmosphere where complexity of understanding, tolerance for ambiguity and responsiveness to constructive criticism are extended and where genuine dissent—the right to express an opposing view without ridicule, coercion or censure—flourishes. Challenging but achievable, this ideal of impartiality suggests a collaborative and passionate, if not conflict free, search for truth.

Neutrality in this context refers to the belief, and corresponding practice, that teachers should remain silent about their own views on controversial issues. On occasion, as one of a number of strategies suggested in the previous paragraph, teachers may subsume their own position under the role of a devil's advocate. This practice is not done to manipulate but, in the interest of impartiality, to insure that relevant views are duly considered. On other occasions, when directly queried by students, neutral teachers may reluctantly disclose their personal position. However, their disclosures are characteristically understated and usually qualified by repeated declarations that theirs is just one of several possible positions. In short, far from a positive ideal, the mere expression, much less advocacy, of their own point of view represents for the neutralist a practice to be optimally avoided.

This is a posture advocated by a number of influential educators (Elliott, 1973; Stenhouse, 1972). My professional experience also suggests that preservice teachers agree that the ideal stance of the teacher is to facilitate so that at the end of the discussion students do not know which side of the controversy the teacher stands.

Why would educators hold this point of view? Below I distinguish and then critique six different explanations, each of which may coexist in varying degrees within one individual. While none by itself supplies a full explanation, the set represents the major factors influencing the neutralist stance.

One explanation is the *public service orientation*. Tracing its roots to Plato and its intensified cultivation in the efficiency and positivist movements of the turn of the century, Bullough and his colleagues (Bullough, Gitlin, & Goldstein, 1984; Bullough, Goldstein, & Holt, 1984) explain how teachers are part of a long tradition of public service whose ideal virtues include being industrious, obedient, disinterested, unambitious and intelligently loyal. Capturing Plato's sentiments, they state, "as paradigmatic models of civic virtue, public servants must live in Spartan rigor, materially poor but spiritually rich in their selfless identification with the state's welfare" (Bullough, Gitlin, & Goldstein, 1984, p. 344).

The more contemporary manifestation of this tradition sees teachers as faithful subordinates in an institutional hierarchy. Their role is to execute

the curricular choices of others in an efficient, technically competent manner. Viewed more as anonymous team members than as autonomous individuals, as technicians rather than as philosophers, teachers must relegate their personal points of view on controversial issues to a minor, if not irrelevant, status.

While this brief explanation may partially illuminate a cultural tradition which nurtures a predisposition toward neutrality, it is possible to argue that a practice of exclusive and explicit partiality toward the interests of the state would also be a compatible teacher practice. However, the attraction of neutral impartiality is strengthened when we consider the dominant state's political philosophy of *liberal pluralism*. Advocates of liberal pluralism generally view human diversity as a social good—either as a source of social vitality and personal enrichment, or more protectively, as “a hedge against the totalitarian suppression of civil rights” (Greenbaum, 1974, quoted in Day, 1983). Social justice is measured by the efficacy of procedural mechanisms which permit these diverse and often competing interests fair opportunity for voice and fulfillment. Within this pluralist conception, the school is an important arena for the expression of diverse values and the teacher must assume the role of a nonpartisan referee, whose dominant interest is to ensure fair competition in the classroom marketplace of ideas. Any attempt by the teacher to influence the substantive outcome of the ideological market would constitute an exploitive appropriation of power and a gross breach of professional conduct. Thus, procedural fairness, the alleged basis for the legitimacy and stability of the social system, dictates teacher neutrality.

A third explanation, *political prudence*, looks at pluralism from a less sanguine perspective. In a competitive, pluralistic climate, where divisiveness threatens tolerance and where aggressive clients/consumers (i.e., parents and students) pose a perceived omnipresent threat of litigation, expressing one's views on a controversial issue can be seen, at minimum, as an avoidable invitation to unwanted conflict. At maximum, it can invoke job-threatening accusations of coercive indoctrination of blatant bigotry. To avoid these risks, silence may seem the most prudent posture.

A fourth position is that of the *ethical relativist*. In general, the ethical relativist holds the beliefs that (a) all values are of equal worth and (b) there are no morally definitive, objective standards to judge some value positions better than others (Brandt, 1959). For the relativist, ultimately what is right and good is highly personal and subjective. Thus, teacher relativists reject as authoritarian fallacy the role of “keepers of the nation's moral conscience” (May, 1983). Instead they view their central moral responsibility as helping students develop a personal code of ethics which the latter can find useful and satisfying. In sum, believing that their own personal values, no matter how fervently held, cannot claim the authoritarian status that their knowledge of more objective facts permit, these teachers can reason that it

is presumptuous and intrusive to assert their own view on a controversial issue.

A fifth explanation for assuming a neutral posture, *ambivalence*, also relates to the role of the teacher as perceived authority. More psychological in emphasis than previous explanations, this type is characterized by (a) an uncertainty about one's own position on particular controversial issues, and (b) a sense of guilt or embarrassment that they do not have a more definitive stance. These individuals assume that to expose their own ambivalence or ignorance is to act as a poor model and to subject themselves to student ridicule. To conceal their insecurity and protect their credibility as authority figures, they come to embrace the posture of silence and neutrality.

Others approach a neutral stance less unsure about their personal authority on issues, but quite concerned that certain colleagues do assume that role of keepers of the nation's moral conscience. Motivated by a certain conception of fairness characterized as the principle of equal restriction, holders of this perspective reason as follows: There are any number of teachers who are prejudiced and dogmatic on controversial value issues. Outspoken in class, these latter teachers continually violate principles of impartiality. Their behavior is a source of miseducation for students and embarrassment, if not disgrace, to the teaching profession. Norms curbing their behavior need to be established. In actuality, it may be the case that certain teachers have the principled self-discipline to express their personal views without undermining impartial inquiry. However, to be fair and consistent, it is necessary to restrict all teachers from expressing or advocating their personal points of view in controversial areas.

A final explanation, the *rationalist perspective*, focuses more exclusively on the alleged educational impact of teacher disclosure on students. Proponents of this position seek to empower students with the ability critically to analyze alternative value arguments and to develop defensible positions of their own. Because in their view, critical intelligence involves a recognition that some arguments, rooted in more universal principles, are superior to others, rationalists firmly reject the indiscriminate epistemological egalitarianism of ethical relativism. However, for both philosophical and pedagogical reasons, they also eschew a moral absolutism which sees correct moral choice as readily deducible from empirically validated ethical commandments. Such a positivistic position is untenable for at least three reasons. First, there is the inherent difficulty of predicting and controlling actual effects of particular policies on specific individuals. This problem is compounded when short- and long-range impact must be considered and where the number, diversity or anonymity of individuals affected is increased. Second, there is the difficulty of reaching consensus on the actual meaning of generally cherished values such as justice, freedom, equality, diversity, due process, the common good when these are applied in concrete cases. Much of the debate over Butts' (1980) proposal for reviving civic

education centered on this persistent issue (*Journal of Teacher Education*, 1983).

Finally, following Ross' infomative distinctions (1930), there are a number of prima facie duties that may conflict in concrete cases, making the determination of our actual duty in such instances quite problematic. For example, the prima facie duty to tell the truth or keep one's promises may compete in specific situations with other prima facie duties, as to avoid causing harm or suffering. What one's actual duty is in such cases can best, but perhaps not definitively, be judged only by a sensitive contextual consideration of the relevant moral principles and the presumed consequences of alternative actions.

It is both the existence of and conflict between certain prima facie duties which help define the pedagogic approach of the rationalists. For them, the fostering of autonomous critical intelligence is the best means of preparing students to recognize and meet the complexity, the ambiguity and the imperative in moral life without succumbing to the seduction of arbitrary individualism or the addiction of simplifying authoritarianism. Pedagogically, this goal implies nurturing the conditions of impartiality previously discussed. Of particular salience to the rationalists, it also implies that judgments are made on the merits of arguments and not, per se, on any spurious or nondiscursive (Habermas, 1971) features of the competing positions. General nondiscursive features could include favorable or unfavorable stereotyping associated with diverse advocates and positions; advocates' perceived power to affect students' welfare; or personal obligations implied by different arguments. The point here is not to argue that the status or ideology of different advocates or the behavioral impact of their ideas is unimportant in assessing the legitimacy and significance of their claims. Sensitivity to these considerations can be integral to informed critique. The problem occurs when these factors function as nonrational interference, short-circuiting rather than illuminating critical analysis.

Bringing this point home to the student-teacher relationship, rationalists can point to a number of potent forces within schooling which threaten the independence and soundness of students' judgments on controversial issues. These forces include a prevailing compulsory context, unequal power relations within classrooms, the alleged impressionability of the young, a generalized perspective of narrow instrumentalism and more specifically, incessant competition for good grades to gain competitive career advantage. Amidst these forces the teacher is thrust into a number of restricting roles: oppressive custodian to be resented and rejected; credential gatekeeper to be duped or appeased; infallible or intimidating authority to be emulated or feared. Each of these particular roles can intensify the likelihood that nonrational factors will be involved in students' judgments. To minimize these threats to critical intelligence, rationalists attempt to remove themselves from intrusive involvement as much as possible. In this

light, silence about their own position becomes a vital commitment. If students do not know where the teacher stands, they will be forced to rely more on their own critical intelligence. Indeed, for some, the goal of nurturing mature rational autonomy suggests a student-teacher relationship where the teacher become progressively dispensable (Rosen, 1980). Teacher neutrality in controversial matters, not self-disclosure or advocacy, seems best designed to achieve such an ideal.

Critique

A number of these explanations for neutral impartiality are interrelated. In the interests of economy and synthesis, I will critique them in a collective manner, where appropriate. Because the rationalist perspective subsumes some of the other arguments and is the most compelling in certain respects, I will address it first.

The rationalist position has several strengths: its primary focus on the best interests of students rather than teachers; its concern for advancing critical rationality as opposed to fixed doctrine as an ongoing resource for self-development; its explicit rejection of ethical relativism as a guiding philosophical position. However, despite its important strengths, this perspective suffers from problematic assumptions and a narrow rationalism, suggesting the need for a fourth major perspective on the teachers' role.

The presumption of nonrationality. As indicated above, the rationalist assumes that in the context of teachers' superior power and authority, students will be unlikely or less able to reflect rationally on the substantive issues of a controversy when presented with a teacher's explicit point of view. In important respects, however, this contention is problematic. It may well exaggerate the potency of teachers' influence on students' value formation. In a pluralistic society, youth are subjected to a number of diverse and conflicting influences. Not only are the messages transmitted within a given school and a classroom at times incongruous. Often schools and teachers must compete, often unsuccessfully, with values communicated in the home, peer group, church and mass media. Hence, while some students are assuredly impressionable, it may be a gross distortion to assume that most students will merely parrot what teachers say. Teachers know, with a mixture of relief and regret, that the image of children as either simple sponges or defiant resisters plainly is not the case. Adult authority figures should always be sensitive to the influence they may have on youth. Theirs is a dynamic relationship which deserves the utmost care and reflection, as well as substantially more research (Giroux, 1983b; Leming, 1981). However, while the empirical question remains open, there is reason to believe that enhanced student rationality is not dependent upon teachers concealing their own point of view.

The presumption of either impressionability or opposition becomes more dubious as the focus moves from young children to older youth. Generally,

as their capacity for autonomous reasoning matures, youth may view the neutral teacher with distrust and resentment for several reasons. One is the manipulative, gamey quality which can arise when the student is asked to become vulnerable and take risks by expressing personal views on controversial issues while the nominal leader who controls the action deliberately avoids doing so. As revealed in informal discussions with a number of preservice and seasoned teachers, this lack of reciprocity however well intended, can be viewed in a number of related and unfavorable ways: as a cowardly sign of teachers' evading legitimate challenge of their own views; as a frustrating denial of a potentially informative perspective; as evidence of fraudulent commitment to rational inquiry; as an indication that the subject really isn't that important; as an endorsement of straddling issues; or as an admission that certain views must be irrational or inferior for they cannot be reasonably articulated or defended.

Ironically, then, this silence of the teacher can be deafening rather than quieting. Instead of enlivening and legitimating discussions, it may deaden them as students feel manipulated, misled and denied the developmental opportunity to compare their own perspectives and refine their advocacy skills with an expressive, responsible adult. To the extent students, like teachers, fail to become fully engaged, their interest in and powers of critical intelligence can not be optimally realized.

Disputable means-end imperative. Even granting that children and youth are considerably impressionable, it is by no means self-evident that assuming a role of nondisclosing neutrality is either the most logical or instrumental response for safeguarding rational intelligence. To the contrary, it can be reasonably argued that to minimize manipulation or misinterpretation, the opposite teacher role of self-disclosure is needed. Thus, better to assess the merits and balance of particular ideas presented by the teacher, students could well profit from knowing teachers' general ideological persuasion and specific position on concrete issues. While knowledge of teachers' value orientation does not insure students' critical evaluation, it does offer important information with which to judge whether the espoused ideal of impartiality is being promoted in practice.

An additional counterargument can be made. As suggested earlier in the critique of exclusive neutrality, no matter how vigilant or circumspect a teacher may be, it is nearly impossible fully to conceal the nature of one's beliefs, particularly to those with whom one is in daily contact. Nonverbal cues and stray remarks will inevitably leak messages for student decoding. However, because this communication is unintended, indirect and hence likely fragmentary, the probability of misinterpretation is magnified. As Hill notes (1982, pp. 117-118), under these conditions students may emulate selected teacher behavior without adequately understanding the motivation and convictions which inspired the behavior. Relatedly, students may make erroneous interpretations about the teacher's beliefs resulting in confusion

about the relationship between one's beliefs and one's behavior. In either case, it seems reasonable to assume that, other things equal, distortion could be diminished and the interests of critical judgments advanced if teachers offered direct and honest explanations of their beliefs and behaviors.

Narrow rationalism. The previous points suggest that rationalists possess a limited understanding of the role their own behavior exerts in the process of values development. In a provocative article, Leming (1981) argues that as educators we need to be concerned not only with facilitating the soundness of students' reasoning on moral issues, but stimulating and reinforcing students' lived commitment to act in morally defensible ways. Integration of beliefs with behavior Leming saw as the measure of moral maturity, not just the development of rational decision-making skills.

How do educators catalyze this lived commitment? As Leming's review of prosocial behavior reveals, modeling is a powerful source of both short-and long-term behavior change. Of particular salience to the present discussion is the finding that "the power of the model to induce actual performance (as distinguished from acquisition) is strongly influenced by the observed consequences for the model of the exhibited behavior" (Leming, 1981, p. 10).

These findings support on several levels the expression rather than suppression of teachers' positions when discussing controversial issues. At the level of classroom verbal behavior, students need to observe and experience teachers who engage in critical discourse and respond with openness and conviction to dissent and needed refinement of their positions. Such exemplary teacher behavior can enlighten and inspire students. In essence it would authenticate teachers' alleged commitment to rational impartiality as a precondition to informed action. But constructive verbal classroom behavior does not compose the universe of prosocial action, nor on a host of controversial issues confronting youth as students and citizens (e.g. the potential of nuclear holocaust, military draft, school closings, busing, crime, drug abuse) is the classroom necessarily the optimal site to engage in the action consistent with the imperatives of their rational analysis.

While precluding classroom self-disclosure, does the rationalist position encourage teachers' modeling civic advocacy at broader school and community levels? Quite the contrary. Given both the permeable boundaries between classroom and community and their view that self-disclosure contaminates students' rational autonomy, the rationalist's surest safeguard against infecting students is to exorcise the disclosure at its roots, choking public expression on controversial issues together. In effect, then, the practice of classroom disguise, however imperfectly performed, threatens the civic identity of the neutralist. Put differently, the inner logic of classroom neutrality represents both an expansionist and an isolationist doctrine of teacher self-censorship. Designed to liberate students, its masterful execu-

tion may effectively miseducate them while it places the teacher in a civic straitjacket.

Fortunately, if my analysis is correct, there is no need to make the draconian choice implied in the previous comments. That is, if the civic interests of teachers were fundamentally incompatible with the intellectual and ethical interests of students, the results of a hardheaded utilitarian calculus might have us conclude that the self-alienating, sacrificial muteness of the neutralist, while regrettable, was nonetheless educationally sound. However, as I have attempted to demonstrate, in neither the logical, empirical or moral sense does the posture of neutrality necessarily lead to promoting students' autonomous critical intelligence. To the contrary, I have argued that the rationalist's advocacy of teacher neutrality is a self-deceptive and potentially self-defeating doctrine, for it can function to obscure and contradict rather than advance critical intelligence and enlightened action.

In the remaining part of this section, I will briefly critique the prudence, public service and ambivalence arguments for teacher neutrality. The principle of equal restriction is best discussed in the concluding section.

As both justification and explanation, prudence provides an important perspective on teacher neutrality. From a moral point of view, a premier decision-making principle is to give equal consideration to the rights and interests of all relevant parties to a conflict. This consideration naturally and reasonably can begin at home, with oneself and one's family. Hence, in general, prudence is rightfully considered a virtue. However, when conflict avoidance becomes the teacher's dominant priority around controversy, overwhelming considerations of student development, prudence seriously jeopardizes its virtuous status. Manifested in a neutral stance, this narrow form of political prudence assumes the character of an unprofessional egoism.

As public service professionals, teachers must fulfill their twin and paradoxical obligations of leading and being led. As leaders they need to recognize that they are charged with promoting students' reasoned commitment to democratic norms and practices; that they are perpetual role models; that youth often look to them for the guidance that experience and formal training presumably provide. Fulfilling this leadership role implies their rejecting a pattern of personal silence or unreflective compliance to institutional dictates under the banners of political prudence or noble loyalty.

Conversely, in the role of public servants teachers should be held accountable for the beliefs and behaviors they are intentionally or unintentionally fostering in their students. As consumers, guardians, taxpayers and supervisors, all students, parents and administrators have a right to know, discuss and to challenge teachers' school-related perspectives and influences. However, this professional obligation to express points of view on relevant issues, if rightly approached, can also be a positive opportunity for it contains the seeds for ongoing dialogue and personal development (Lickona, 1980).

The modeling/developmental arguments apply as well to those who would choose neutrality because of their ambivalence on specific issues. Endemic in a pluralistic society, ambivalence in the normative domain is an inevitable concomitant of intellectual development and need be no cause for apology or deception. Teachers' public discussion of their ambivalence can provide vital opportunity for students to receive validation and incentive in their own struggle to develop informed convictions. In addition, by obligating themselves to public disclosure, teachers present themselves with further incentive continually to refine their own positions. For these additional reasons, disclosure rather than silence is the more educative pedagogic position.

Committed Impartiality

Refuting the Critique: the Paradox without Contradiction

The arguments in the prior sections have advanced the view that when controversial issues arise, as they inevitably will and legitimately should, the role teachers assume with regard to the expression of their own value positions is of considerable educational significance. More specifically, I have argued that the positions of exclusive neutrality, exclusive partiality and neutral impartiality, while not without certain strengths, are seriously flawed conceptions of that proper teacher role. The critiques of these positions indicate the need for a more compelling fourth position. The character of this preferred role has been largely, though not completely, foreshadowed in the previous section. In this section I will highlight the meaning and significance of the more ideal role. Using Hill's (1982) language, this role is paradoxically termed committed impartiality.

Committed impartiality entails two beliefs. First, teachers should state rather than conceal their own views on controversial issues. Second, they should foster the pursuit of truth by insuring that competing perspectives receive a fair hearing through critical discourse. Discussed earlier, the ideal of impartiality needs no separate elaboration here. It is the first of these two terms, the notion of being committed, which needs explication.

It is important to define teacher disclosure carefully. What is recommended is that teacher expression of personal views represent a positive ideal. It is conduct which should be consciously included rather than avoided in the discussion of controversial issues. Teachers' views should be clearly owned, not consistently disguised under devil's advocacy or compromised with excessive humility or repeated qualification. This disclosure may be teacher initiated or a response to direct student inquiry; it may be conveyed in passionate or understated terms. Questions of initiation, timing and tone should be decided by the judicious teacher consistent with the imperatives of impartiality and personal witness. To recommend that teachers state their personal views on issues does not mean, however, that as a general rule, they repeatedly attempt to convince students of the superiority of their

own positions. To the extent that teacher self-disclosure becomes heavy-handed advocacy, it may reasonably be perceived by students as propaganda or psychological intimidation. In either case, the norm of impartiality would be undermined.

To acknowledge that there are certain clearcut cases of abuse of teacher self-disclosure is not to assert, at least successfully, that all teacher self-disclosure is clearly violative of impartiality. However, neutralist critics of committed impartiality essentially do make that claim. Their reasoning goes as follows. Conditions in the classrooms such as unequal power, compulsory attendance, and pressure for grades create an atmosphere where teacher self-disclosure, no matter how understated, is implicit advocacy. These same conditions, however, make any form of advocacy coercive. Coercion is undeniably incompatible with impartiality and should be deterred. Hence, teacher self-disclosure should be precluded. Committed impartiality is not just paradoxical; it is contradictory.

This neutralist conclusion needs further examination. The claim that committed impartiality is a contradiction could be based on one of two sets of assumptions. I will argue that both are unconvincing. One view focuses primarily on the teacher rather than the student. The claim that stating one's convictions precludes rational analysis seems based on the assumption that emotions and reason inhabit discrete and antagonistic spheres of existence. With regard to the nature of our beliefs, emotions are seen to function solely to distort clearheaded analysis. In this view, the population of the emotions is limited to villainous characters: prejudice, rationalization, resistance to reexamination and reasoned modification of position.

While it is true that, at times, nondiscursive or irrational factors interfere with reasoned judgments, it is certainly not always true that emotions generate such a toxic effect. This perspective does not account for the numerous cases where emotions and reason are compatible, even mutually reinforcing. Consider, for example a strong desire to impress a respected associate; or moral indignation in response to governmental abuse of power; or the intense antagonism and incentives embedded in the adversarial system of jurisprudence. In each instance, emotions could well trigger a more rational articulation of one's position. Rather than being inevitably subversive in nature, emotions have the potential to animate a search for truth, to compel action consistent with the provisional findings of that search, and/or, more intuitively, actually to inform that search.

Also, by implication, this critique of teacher self-disclosure elevates robotic rationality as the ideal stance. Given the view of emotions as wholly contaminating, it would seem to follow that the premier product of this dichotomous conception is a hollow embodiment of pure reason. It is an eviscerated, one-dimensional creature. It is hardly a compelling ideal, as real human beings need not apply for the throne.

A second basis for the claim that committed impartiality is a contradic-

tion focuses on students' alleged inability to remain rational in the face of the teacher's expressed views. This is the rationalist position critiqued in the preceding section. That critique need not be reiterated here. However, it should be noted that the rationalist's presumption of student impressionability is rooted in a belief that teachers can strongly influence student behavior. Granting this, it is crucial to emphasize that teachers also possess a potent weapon of a continuing affirmation of the value of impartiality that safeguards the rational process. That norm can be authenticated in practice as teachers praise reasoned oppositional viewpoints, push students to critique teachers' points of view, publically engage in self-critique, or critique students who merely parrot them. In short, teachers possess a set of strategic correctives which should be able to reduce threats to rational analysis potentially precipitated by teachers' disclosing their personal views.

Advancing the Positive: The Role of Personal Witness, Democratic Authority and Collegial Mentor

I have argued that there is no inherent contradiction between expressing one's commitments and maintaining the norm of impartiality. However, the absence of central contradictions does not translate into the presence of a compelling case. What positive reasons exist for recommending committed impartiality as the preferred teacher role? Much of the rationale can be inferred from preceding critiques. Here, I will summarize and, where necessary, elaborate the rationale under three interrelated ideas: personal witness, democratic authority and collegial mentor.

The idea of a personal witness is meant to convey the power of personal modeling and the imperative of personal integrity. Being a personal witness places emphasis on observation and example, two major modalities by which individuals learn. As G. S. Bilkin has noted, a teacher is not only one who imparts truths and skills by instruction by one who is a "truth for students" (Hill, 1982 p. 117). In the ideal, teachers as personal witnesses are those who possess and live reasoned convictions and believe youth should do so too. In both a personal and civic sense, they recognize authenticity and integrity as the best nutrients for sustaining intellectual, psychological and ethical health. Consequently, they reject as fraudulent and/or miseducative for self, subject and student roles of neutrality and partiality previously critiqued. Conscious of teaching by example, they attempt to live exemplarily. Inevitably failing to do so, they forthrightly address their imperfections, and ironically, in so doing, they exemplify a distinctively human achievement (Perkinson, 1984).

The notion of a democratic authority emphasizes learning through direct experience and is meant to convey several related ideas. Philosophically, as the governors in a democracy, we are all authorities. Hence, practically we need to perceive and practice ourselves in that role. In particular, as has been argued persuasively (Newmann, 1975; Newmann, Bertocci, & Landsness, 1977; Wood, 1985), schools need to be sites and sponsors for youthful

citizens-in-training. In order to develop fully the requisite competence and identity, youth in school will need to experience opportunities where they can confront authority in a genuine yet supportive manner.

While not the only source, teachers practicing committed impartiality excellently offer such an experience. As personal witnesses giving voice to themselves and permitting fair hearing to youth, these teachers, in theory, both embody and help empower democratic authorities. When students have access to the authority's ideas, are repeatedly pressed to challenge the validity of those ideas, and are coached in the process free from spurious sanctions, they are assuredly engaged in an educationally enriched environment where civic commitments, competences and courage are born and renewed (Mosher, 1982).

A third set of reasons why the practice of committed impartiality is to be preferred is rooted in developmental research. The most comprehensive research on experimental education programs with adolescents (Conrad & Hedin, 1981) reveals that the type of adult-youth relationship most associated with youthful social development is captured by the phrase collegial mentor. Like committed impartiality the term is paradoxical. Additionally, although not identically, both phrases unconventionally assert a harmony of interests between dynamics of equality and expertise. For our present purposes, it is the concept of collegiality, not mentorship, which is most relevant.

Collegiality contains a number of developmentally rich dynamics common to the role of committed impartiality. These can be grouped under two related headings of mutuality and multidimensionality. Mutuality involves teachers' belief that students can make useful contributions to the learning process. Teachers show genuine respect for students' knowledge and interests, manifested in a nonimpositional, nonpatronizing style of interaction. This style and set of beliefs unite the collegial mentors and practitioners of committed impartiality in opposition to those who practice exclusive partiality.

The quality of multidimensionality on the other hand unites the collegial mentor and practitioner of committed impartiality in opposition to the impartial neutralist. Harmonious with a personal witness perspective, multidimensionality involves relating to others in an authentic, nonposturing way. Interaction is characterized by an engagement at diverse levels of experience between people seen as individuals, not merely role incumbents. Like the practitioners of committed impartiality, teachers who convey a collegial attitude toward youth see it as natural and appropriate to share honestly their personal feelings and beliefs on relevant matters. In this process, youth tend to feel entrusted and enhanced. Treated as a colleague, youth begin to see themselves as more adult. So it is that the collegial dynamics of mutuality and multidimensionality, conveyed too by the practitioner of committed impartiality, intertwine in a developmentally reinforcing pattern of affirmative transaction.

Conclusion

In the preceding section I drew on the perspectives of personal witness, democratic authority and collegial mentor to argue that the paradoxical role of committed impartiality is most proper for teachers to assume in discussing controversial issues. It is most proper because it presents a model of a fully functioning human being, one who expresses and acts upon reasoned convictions. Sensitively encouraging the same in students through the dynamics of modeling and the norms of impartiality, this teacher creates an educative culture in which relevant controversial issues of important curricula are legitimately confronted without undermining the integrity of either subject content or self. Given this affirmative rationale, the neutralist principle of equal restriction emerges as a significantly misplaced self-censorship. To suppress one's own self-disclosure as an alleged requirement for seeking the fair censure of a strict partisan is analogous to throwing out the healthy baby of committed impartiality with the fouled bath water of exclusive or prejudicial partiality. As this paper has systematically attempted to demonstrate, the compelling tasks for social educators are neither to choke teacher self-disclosure nor to concede to irresponsible partiality. Rather they are to reject as myth and misguidance a value-free and nondisclosing neutrality and to work continually to infuse classroom discourse with that balance of personal commitment and impartiality which promises to catalyze the critical intelligence and civic courage of both our youthful citizens and ourselves. These are persisting challenges we cannot, and should not, avoid.²

Endnotes

1. I want to emphasize the qualification made here. I am not arguing that all members of the radical left are strict partisans as I am using that concept. Berlak explicitly rejects claims to representativeness of her sample. She sees her data, drawn from a sample of thirty higher education teachers, as more impressionistic and suggestive of further study than systematic and conclusive in themselves. In addition, her data indicate that a number of these teachers viewed their own oppositional perspectives as problematic, subject to continuing interrogation. The reader interested in more detail than provided here is encouraged to examine her work directly. For a recent example of a systematic and balanced rendering of both mainstream and oppositional perspectives by a left scholar, see Beyer (1984).

2. A number of colleagues have been helpful in the development of this paper. They include John Boatright, Jim Dague, Jeff Edelson, Jean Erdman, Andrew Gitlin, Jesse Goodman, Jerry Jorgenson, Pat Kearney, Alan Lockwood, Fred Newmann, Marcie Stein, Bow Sweeney, Gary Wehlage, Sally Wertheim, Erik Olin Wright and Michael Wright. In addition, many students, too numerous to list, made their valuable contributions. Finally, I would like to thank Rita McCauley for her excellent manuscript preparation.

References

- Anyon, J. (1980). Social class and the hidden curriculum of work. *Journal of Education*, 162, 67-92.
- Anyon, J. (1984). Intersections of gender and class: Accommodation and resistance

- by working class and affluent females to contradictory sex role ideologies. *Journal of Education*, 166, 25-48.
- Apple, M. W. (1979). *Ideology and curriculum*. London: Routledge & Kegan Paul.
- Apple, M. W. (1982). *Education and power*. Boston: Routledge & Kegan Paul.
- Apple, M. W., & Weis, L. (Eds.) (1983). *Ideology and practice in schooling*. Philadelphia: Temple University Press.
- Arnot, M. (1981). *Class, gender and education*. Milton Keynes: Open University Press.
- Bereiter, C. (1973). *Must we educate?* Englewood Cliffs, NJ: Prentice-Hall.
- Bernstein, R. J. (1978). *The restructuring of social and political theory*. Philadelphia: University of Pennsylvania Press.
- Berlak, A. (1985). Back to the basics: Liberatory pedagogy and the liberal arts. *Thought and Action*, Spring.
- Beyer, L. (1984). Education reform: The political roots of national risk. *Curriculum Inquiry*, 15, 37-56.
- Brandt, R. B. (1959). *Ethical theory: The problems of normative and critical ethics*. Englewood Cliffs, NJ: Prentice-Hall.
- Brophy, J. (1979). Teacher behavior and its effects. *Journal of Educational Psychology*, 71, 733-750.
- Bullough, R. V., Gitlin, A. D., & Goldstein, S. L. (1984). Ideology, teacher role, and resistance. *Teacher's College Record*, 86, 339-358.
- Bullough, R. V., Goldstein, L., & Holt, L. (1984). *Human interests in the curriculum*. New York: Teachers College Press.
- Butts, R. F. (1980). *The revival of civic learning: A rationale for citizenship in American schools*. Bloomington, IN: Phi Delta Kappa.
- Cagan, E. (1978). Individualism, collectivism and radical educational reform. *Harvard Educational Review*, 48, 227-266.
- Cherryholmes, C. H. (1980). Social knowledge and citizenship education: Two views of truth and criticism. *Curriculum Inquiry*, 10, 115-141.
- Cohen, B. (1982). *Means and ends in education*. London: George Allen & Unwin.
- Coleman, J. S. (1972). *Youth: Transition to adulthood*. Report of the Panel on Youth of the President's Science Advisory Committee. Chicago: University of Chicago Press.
- Conrad, D., & Hedin, D. (1981). *Executive summary of experimental education evaluation project*. Minnesota: Center for Youth Development and Research.
- Day, D. (1983). Agreeing to differ: The logic of pluralism. In F. Coffield & R. Goodings (Eds.), *Sacred cows in education*. University of Durham: Edinburgh University Press

- Dewey, J. E. (1944). *Democracy and education*. Toronto: Macmillan.
- Elliott, J. (1973). Neutrality, rationality and the role of the teacher. In R. S. Peters, P. H. Hurst, & H. T. Sockett (Eds.), *Proceedings of the philosophy of education society of Great Britain*. Oxford: Basil Blackwell.
- Freire, P. (1970). *Pedagogy of the oppressed*. Trans. Myra Bergman Ramos. London: Herder & Herder.
- Genovese, E. (1972). *Roll, Jordan, roll: The world the slaves made*. New York: Vintage.
- Giroux, H. A. (1983a). *Theory and resistance in education: A pedagogy for the opposition*. South Hadley, MA: Bergin & Garvey.
- Giroux, H. A. (1983b). Theories of reproduction and resistance in the new sociology of education: A critical analysis. *Harvard Education Review*, 53, 257-293.
- Goble, F. G., & Brooks, B. D. (1983). *The case for character education*. Ottawa, IL: Green Hill Publishers.
- Goodlad, J. I. (1984). *A place called school: Prospects for the future*. New York: McGraw-Hill.
- Greenbaum, W. (1974). America in search of a new ideal: An essay on the rise of pluralism. *Harvard Educational Review*, 44, 411-440.
- Habermas, J. (1971). *Knowledge and human interests*. Trans. Jeremy J. Shapiro. New York: Beacon Press.
- Hill, B. V. (1982). *Faith at the blackboard: Issues confronting the Christian teacher*. Grand Rapids, MI: William B. Eerdmans Publishing Company.
- Journal of Teacher Education*. (1983). 34(6).
- Kohlberg, L. (1981). *The philosophy of education*. San Francisco: Harper & Row.
- Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Leming, J. S. (1981). On the limits of rational moral education. *Theory and Research in Social Education*, 9(1), 7-34.
- Lickona, T. (1980). What does moral psychology have to say to the teacher of ethics? In d. Callahan & S. Bok (Eds.), *Ethics teaching in higher education*. New York: Plenum Press.
- Lockwood, A., & Harris, D. (1985). *Reasoning with democratic values: Ethical problems in United States history. Vol. I, II and Instructors Manual*. New York: Teachers College Press.
- Lucas, C. (1983). Positivist socio-pathology and the contemporary plight of the humanities. *The humanities in education: Rebirth or burial in the 1980's*. Ann Arbor, MI: Prakken, 20-30.
- Macklin, R. (1980). Problems in the teaching of ethics: Pluralism and indoctrination. In D. Callahan & S. Bok (Eds.), *Ethics teaching in higher education*. New York: Plenum Press.

- May, P. R. (1983). The teaching of moral values. In F. Coffield & R. Goodings (Eds.), *Sacred cows in education*. University of Durham: Edinburgh University Press.
- McNeil, L. M. (1983). Defensive teaching and classroom control. In M. W. Apple & L. Weis (Eds.), *Ideology and practice in schooling*. Philadelphia: Temple University Press.
- Mosher, R. (1982). Curricula and programs in secondary school which positively influence adolescents' development: Where there's the will, there's a way. In F. M. Newmann & C. E. Sleeter (Eds.), *Adolescent development and secondary schooling*. Madison, WI: University of Wisconsin R & D Center.
- Newmann, F. M. (1975). *Education for citizen action: Challenges for secondary curriculum*. Berkeley: McCutchan.
- Newmann, F. M. (1977a). Alternative approaches to citizenship education: A search for authenticity. In B. F. Brown (Ed.), *Education for responsible citizenship*. New York: McGraw-Hill.
- Newmann, F. M. (1977b). Building a rationale for civic education. In J. P. Shaver (Ed.), *Building rationales for citizenship education*. Washington, DC: National Council for the Social Studies, Bulletin #52.
- Newmann, F. M. (1981). Reducing alienation in high schools: Implications of theory. *Harvard Educational Review*, 54, 546-564.
- Newmann, F. M. (1985). The radical perspective on social studies: A synthesis and critique. *Theory and Research in Social Education*, 31(1), 1-18.
- Newmann, F. M., Bertocci, T. A., & Landsness, R. M. (1977). *Skills in citizen action: An English-social studies program for secondary schools*. Skokie, IL: National Textbook Co.
- Newmann, F. M., & Kelly, T. E. (1983). *Human dignity and excellence in education: Guidelines for curriculum policy*. Washington, DC: Report to the National Institute of Education.
- Newmann, F. M., & Oliver, D. W. (1970). *Clarifying public controversy: An approach to teaching social studies*. Boston: Little, Brown.
- Perkinson, H. (1984). *Learning from our mistakes: A reinterpretation of twentieth century educational theory*. Westport, CT: Greenwood Press.
- Polanyi, M. (1974). *Personal knowledge*. Chicago: University of Chicago Press.
- Raywid, M. A. (1980). The discovery and rejection of indoctrination. *Educational Theory*, 30(1), 1-10.
- Rosen, B. (1980). The teaching of ethics in undergraduate nonethics courses. In D. Callahan & S. Bok (Eds.), *Ethics teaching in higher education*. New York: Plenum Press.
- Rosenbaum, J. E. (1976). *Making inequality*. New York: Wiley.
- Ross, W. D. (1930). *The right and the good*. Oxford: Clarendon Press.

- Rutter, M., Maughan, B., Mortimore, P. & Outson, J., with Smith, A. (1982). *Fifteen thousand hours*. Cambridge, MA: Harvard University Press.
- Sichel, B. A. (1982). *Value education for an age of crisis*. Washington, DC: University Press of America.
- Stenhouse, L. (1972). The idea of neutrality. *Times Educational Supplement*, 4(2).
- Straughn, R. (1982). *I ought to, but . . .* Windsor, Berks: NFER-Nelson.
- Taxel, J. (1984). The American revolution in children's fiction: An analysis of historical meaning and narrative structure. *Curriculum Inquiry*, 14, 7-55.
- Will, G. F. (1983). *Statecraft as soulcraft: What government does*. New York: Simon & Schuster.
- Willis, P. (1981). *Learning to labor: How working class kids get working class jobs*. New York: Columbia University Press.
- Wood, G. H. (1985). Education for democratic participation: Democratic values and the nuclear freeze campaign. *Theory and Research in Social Education*, 12(4), 39-56.
- Young, M. F. D. (Ed.) (1971). *Knowledge and control*. London: Collier-Macmillan.

Rethinking Social Studies Research and the Goals of Social Education

James S. Leming
Southern Illinois University

Abstract

This paper examines the research on social studies curriculum's influence on the social, moral, and political attitudes of youth. It is argued that it is difficult to make a case for the social or educational significance of these findings given their small magnitude. Four alternative interpretations of this genre of research are presented. It is concluded that the social studies profession should focus primarily on the achievement of cognitive goals and that further research into curricular effectiveness without longitudinal data is of limited value. As an alternative to curricular effectiveness research, it is proposed that descriptions of the workings of exemplary programs become a major research priority.

The concept of citizenship has long been at the heart of social studies education. Central to the achievement of this goal has been the belief that social education should foster the development of positive democratic orientations in youth. It has likewise been accepted by the research community that social studies research should assess the extent to which social studies curriculum has, or has not, contributed to the achievement of this goal. The purpose of this paper is to review briefly recent relevant research and to explore alternative views on the goals of social studies and social studies research that are suggested by this review.

The brief review of research presented below is based on a more extended review recently published (Leming, 1985). The purpose of this more extended review was to examine the literature between 1976 and 1984 on the influence of social studies curriculum on the social orientations of students: students' reasoning, attitudes, dispositions, and behaviors regarding moral, social, and political life. This review was limited to areas where more than one or two isolated studies existed. As a result, the four areas of research selected were those where a research base existed from which one could, with some degree of confidence, draw generalizations. This paper did not report the results of areas where the research is notoriously poor (e.g., values clarification) or areas where only isolated studies exist (e.g., values anal-

ysis). For each of the four areas identified I will briefly describe the nature of the instructional practice under study, the outcomes analyzed in the studies, and the general pattern and magnitude of findings. The report of the findings will necessarily be brief.

Research Areas in the Socio-Moral-Political Domain

Moral Dilemma Discussion Approach

Most professional educators are well familiar with the Kohlbergian approach to moral development and its implications for social studies education. The approach claims that there exist five naturally occurring stages of moral reasoning that comprise an invariant developmental sequence. The essential conditions for moral development consist of cognitive disequilibrium (through deliberation about moral dilemmas) and exposure to reasoning at the next highest stage. Kohlberg (1981) claims that the higher stages are morally superior stages and therefore, stimulating cognitive development is a desirable educational goal.

The research reviewed involved the incorporation of the discussion of moral dilemmas on a weekly basis within the regular school curriculum. Stage of moral reasoning (the dependent variable) was typically ascertained by the scoring of a recorded transcript of student response to interviewer probes of responses to a moral dilemma (the Moral Judgment Interview) or by an objective paper and pencil assessment of moral judgment (the Defining Issues Test). Although student responses in the Moral Judgment Interview are occasionally reported as a global (modal) score, most commonly responses are reported as a moral maturity score (MMS). Reporting the data in this form yields a score between 100 (pure stage one) and 500 (pure stage five). The MMS is a mean statistic and, as a result, is an average of a number of responses. The same score in two different subjects may reflect either a homogeneous or heterogenous pattern of subject reasoning.

The research suggests that the stimulation of development in moral reasoning is an achievable goal. In approximately 50% of the studies identified, the regular discussion of moral dilemmas in the classroom resulted in statistically significant upward change in MMS. If one eliminates studies of duration of less than one semester, successful treatment rises to around 70%. If one further eliminates from the longer studies those that incorporate moral dilemmas dealing with religious issues and questions of sexual morality, 90% of the studies find upward stage movement. The magnitude of change ranges between 1/7 and 1/2 stage or between 15 and 50 points on a 400 point scale (4%–12% of the scale's range).

Schaepli, Rest, and Thoma (1985) have recently completed a review of literature on the effect of the dilemma discussion approach in studies that utilized the Defining Issues Test. This test yields a P score, or percentage of moral reasoning at the principled level. Thoma assessed the power of the treatment by using the standard effect size statistic. This statistic represents the magnitude of treatment effects in a standard score form (a percentage of

standard deviation from pre- to post-test). In the analysis of 40 studies, it was found that overall the magnitude of the effect was small ($d = .28$). This effect size increases, however, as those studies of weak design are eliminated (i.e., studies of short duration or those not pointedly focused on moral dilemmas). This study also found that the effect size was greatest for adult and college age youth while studies using high school and junior high youth yielded effect sizes of .23 and .22, respectively.

Classroom Climate and Controversial Issues in the Social Studies

Social studies classrooms may have a desirable influence on student attitudes, provided there exists regular discussion of controversial issues, a classroom climate where students feel free to express their opinions, and a teacher who actively solicits and positively values student contributions. This body of research consists of far fewer studies with a wider variety of dependent variables when compared with the research on the moral discussion approach. Generally, the incorporation of controversial issues as a part of the social studies classroom fosters greater respect for civil liberties, and has a positive impact on students' political attitudes. The attitudes most commonly measured were: political trust—belief that human behavior is constant and governed by positive emotions; social/political integration—belief that one is connected, not alienated from one's environment; political interest—one is predisposed to respond positively to political matters; and political efficacy—belief that one's actions can have an effect on the political environment.

The attitudes in this body of literature are typically measured by a five-point Likert scale with students responding to items such as "People like me can influence political decisions." The findings from this body of literature are reported in a variety of ways such as Z-scores, group means, or percent of subjects at one end of a scale. Also, the number and makeup of items vary widely among scales. Nevertheless, the pattern of findings is consistent. Usually, on all political attitude variables with the exception of political efficacy (confidence), small positive increases in attitudes were detected. After courses that involved discussion of controversial issues, students became slightly less confident about their ability to influence political events. One author, Ehman (1980), attributes this declining efficacy to a sense of political realism.

A related body of research focuses on the influence of democratic classrooms and democratic alternative schools on the political attitudes of students. These studies generally found that sense of political efficacy and interest were slightly enhanced; however, a small but significant increase in cynicism also frequently was detected.

Community Service/Social Action

The programs reviewed in this section place students outside of the school and in the community where they assume participatory roles with real social consequences. These roles, depending upon the type of program, range

from performing public service activities to participation intended to influence social/political policy or action. The rationale for such programs typically anticipates one or more of the following outcomes: the development of positive attitudes related to civic and political participation, the acquisition of relevant interpersonal, intellectual, and political action skills, and increased desire for participation in civic life. The research on this approach to social education comes from three primary sources: Jones (1974), Conrad and Hedin (1981), and Newmann and Rutter (1983). Jones found only a slight positive influence of such programs on efficacy on measures of political awareness, perception of public officials, and desire to be a public official. Conrad and Hedin found slight increases favoring experiential programs on moral reasoning, sense of duty, concern for others, sense of efficacy, and responsible action. However, the Newmann and Rutter review is critical of the Conrad and Hedin study, questioning the statistical treatment of the data and pointing out that the comparison group scores declined from pretest to posttest on every variable, thereby increasing the probability of achieving positive results for program students.

Newmann and Rutter (1983) examined the impact of community service programs in eight schools. The variables measured were taken from a researcher constructed test of sense of community and school responsibility, sense of social competence, political efficacy, anticipated future community involvement, and anticipated future political participation. The study concluded that community service programs increase students' sense of community responsibility and sense of personal competence in a very modest way (mean movement of about 1.5% of the range of a five-point scale), but have no impact on the other variables studied.

Cooperative Learning

Cooperative learning is a recent approach that has accumulated an impressive body of research that suggests promise for social education. The approach attempts to redress what is seen as an overemphasis on competitive and individualistic structures in schooling by placing students in cooperative learning situations (Slavin, 1983). Typically, groups of four to five students are comprised in such a way as to include a wide range of academic, social and racial backgrounds and students are rewarded for the success of the group as a whole. The extensive field-based research can be summarized as follows: Cooperative learning methods in general, when compared with individualistic and competitive methods, produce greater academic learning, better intergroup relations among black, white, and Hispanic students, enhanced self-esteem, improved relationships between mainstreamed academically handicapped students and normal progress students, general mutual concern and interpersonal trust among students, and increased propensity for prosocial behavior. While the magnitude of the effect size is comparable with the other research reported, the consistency of the findings is impressive.

Interpreting the Research

Any research findings are socially meaningless until they are placed within a framework consisting of agreed upon purpose and value regarding the nature of education. For example, the finding that experience in social studies classroom X results in a mean change in variable Y among students becomes meaningful only within a framework from which importance is attached to finding Y based on some view of Y's relationship to a desired personal or social end-state (value). Once this framework from which the interpretation takes place is established, a further question concerning the significance of the magnitude of the change in the variables must be addressed. That is, once it is agreed that variables of the sort Y are valid outcomes (goals) for social education, then one must decide upon what constitutes significant change as a result of a curricular intervention.

It is my position, in this paper, that the nature of what constitutes an educationally and socially significant finding needs clarification. In addition, the relationship between the variables utilized in the research and the goals of social education need more careful examination. In this section I will address two points but limit discussion to that research which focuses on the development of reasoning about political questions and on democratic political orientations: political efficacy, interest, trust, and participation. I adopt this perspective both to give the remaining sections focus and because of the alleged relationship of these variables to the traditional social studies goal of citizenship education.

I turn first to the question of the significance of the findings. It appears to me that there are four potential interpretations regarding the question of the educational and social significance of the changes found.

The trivial findings interpretation. The research findings suggest that the influence of social studies programs in the socio-moral domain are trivial and completely without educational or social significance. This perspective holds that the magnitude of the changes are so small that no reasonable case can be made for continuing the investment of time and energy expended on any given approach. Not only has it been shown that the changes are miniscule, but the research has also failed to demonstrate that these changes persist after the treatment ends. Also, no link between the paper and pencil data collected and real-world social or political behavior has been established. Given these findings, the profession would best use its time and energy concentrating on what it does best, increasing student knowledge.

The resistant variables interpretation. It is granted that the changes detected in the research are small; however, given the nature of the variables, this is not unexpected. Behavioral science research has long found the areas of attitudes and values to be among the most resistant to change. With the well established potent influence of the home and the somewhat less potent media and culture, it would appear unrealistic to expect schooling experi-

ences to dramatically impact on political attitudes. The fact that any positive influence is detected is encouraging and sufficient warrant for holding the intervention as effective.

The small but socially significant interpretation. This interpretation is built around the observation that sometimes the small changes found as a result of an intervention may have great social significance. The best example of this line of argument comes from the field of medical research. Gage's (1984) recent discussion of small effects in teaching effectiveness research makes this medical analogy. One study cited by Gage is the research reporting that through restriction in cholesterol and treatment with drugs the incidence of heart attack can be reduced. Although the above treatment produced only a 1.7% reduction in heart attacks and accounted for only .1% of the variance in heart attacks, the study was widely cited as of profound significance for the practice of medicine. If we can argue that the changes found in social studies research are of equal social significance, then the research in the field takes on a much more salutary appearance.

The developmental interpretation. This interpretation assumes that political reasoning and attitudes are developmental in nature, that is, they proceed through an invariant progression of stages from an immature egocentric perspective to the highest stages where attitudes and reasoning of individuals are consistent with and supportive of the highest principals of our constitutional democratic system of government. From this perspective one would positively interpret the small gains induced as a result of different social studies interventions as making an essential contribution to the democratic development of youth. The small increments observed are seen as essential steps toward more mature and democratic form of belief and reasoning. Since development is invariant and progressive, these small gains will not be lost, but rather are a small socially significant step toward a desirable goal of a developmentally mature populace that possesses a sophisticated understanding of our government and a deep affective allegiance to its principles. The Kohlbergian research on reasoning easily fits the interpretation spelled out above. To attempt to fit political attitudes to this template is, as will be seen below, a much more difficult matter.

Which of the above interpretations is the most plausible? The trivial findings interpretation is probably the most distasteful to social studies researchers (for it suggests that real estate sales may well be a more socially productive career), but it is a difficult notion to dispel. The critical information needed to reject this proposal is longitudinal; that is, if it can be shown that these incremental changes persist over time, or perhaps lie dormant for a period of time, but eventually have impact, or in some way accrete to that constellation of attributes we typically refer to as good citizenship, then there exists nontrivial social value in the enterprise of social education. While there is some evidence that this may be the case, there is

other evidence that suggests a sanguine mood is not in order. One piece of encouraging research is drawn from Almond and Verba's (1963) finding that adults in their sample who scored highest in sense of political competence (believe that they are able to participate in politics) are those who with the greatest frequency remember that they could participate in class discussion. There also exists a body of research that finds that active adult citizens report political and civic participation as youth. These findings, however, are only suggestive; they are not longitudinal and the exact nature of the youthful experiences are shrouded in the mists of the remembered past. True longitudinal studies on this issue are, to my knowledge, non-existent or unreliable. I could identify no study that precisely documents the nature of the student's social studies experiences and then follows up over time the development of student political attitudes. Ehman (1980) followed students over a three-year period (while they were in school) and found the open classroom climate factor accentuating development of selected political attitudes over time. Whether or not these attitude changes persisted after exposure to an open climate ended was not studied. Miller (1985) recently reported an analysis of the National Longitudinal Study of the Class of 1972 (sponsored by the U.S. Office of Education and the National Center for Educational Statistics) where he used the data base to explore the influence of high school social studies courses on adult political participation. In this sample of 16,000 subjects, data were collected seven years after high school graduation on reported political participation. He concluded that the number of high school social studies classes taken has no significant influence on subsequent political participation. While the study addresses a most important question, the failure to describe the nature of the different experiences within the social studies classes and lumping the sample into only two groups (two years of social studies and more than two years) suggest that a cautious approach to the findings is warranted. To sum up, the evidence does not warrant dismissing the first interpretation. There is some data, retrospective in nature, that suggests some lingering influence of school experiences on citizenship attitudes, but careful longitudinal data are needed.

The resistant variables interpretation seems warranted. Attitudes are difficult to change and when one compares the school and its potential influence with that of the family, peers, and the cultural milieu, one must be conservative about what reasonably can be expected from social studies instruction in this regard. This interpretation, however, while making us less idealistic about what can be accomplished, does not establish any grounds for saying that the small influence of social education on student attitudes has social value, for the same possibility of ephemeral changes as discussed under the trivial interpretation persists.

The small but socially significant (medical analogy) interpretation is engaging, but unfortunately such a critical and significant outcome as preserv-

ing a human life simply does not exist in social studies research. While few will argue that development of commitment to democratic principles and citizenship skills are not highly significant social outcomes, difficulty exists in showing that the social studies contributes in any meaningful and lasting way to that development. It is much like a physician arguing for the effectiveness of a given treatment by only reporting that immediately after leaving the office, the patient reported feeling better.

Finally, the developmental interpretation is at least partially encouraging. By now, from my perspective, one can say with confidence that in school settings it is possible to facilitate the development of moral reasoning. A variety of research studies have also shown that reasoning about political issues follows hand in hand with development in moral reasoning. Kohlberg (1981) has also shown that his highest stage of moral reasoning (stage five) embodies the morality of the U.S. Constitution, and to fully understand our constitution requires the highest attainable stage. Therefore, since the small changes noted in the dilemma discussion classrooms appear irreversible and provide the foundation for subsequent growth, and since they lead to a fuller understanding and appreciation of our democratic constitutional system of government, one can establish some social value for the achievement of these outcomes. However, the relationship between reasoning, attitudes, and social/political behavior has not been demonstrated. It seems obvious that one cannot talk about political attitudes in any strict developmental sense. It cannot be said that there exists an invariant, progressive sequence of political attitudes toward some developmental end point, for it makes perfect sense to talk about losing interest, losing one's sense of efficacy, losing the desire to participate, and developing mistrust. In addition, Jennings and Niemi (1981) demonstrated that historical factors such as the Vietnam war and Watergate can have a negative influence on factors such as political trust. In their study, in response to the statement "Government is run for the benefit of all the people," a parent and youth panel fell from 87% and 75% agreement respectively in 1967, to 39% and 45% agreement in 1973. In addition, a comparison of the responses of the 1965 and 1973 senior classes found 78% agreement among the 1965 class, and only 36% agreement in 1973. Clearly, historical events can have negative impact on political attitudes. Thus, I conclude that while the developmental interpretation may offer some solace regarding cognitive goals, it cannot account for outcomes in the area of political attitudes. It should be pointed out, however, that an exploratory body of literature exists that suggests there may be developmental perspectives appropriate to social education. Work by Durio (1976), Hogan and Mills (1976), Rawls (1971), Durkheim (1973), Tapp and Kohlberg (1971), and Dynneson and Gross (1985) offer differing perspectives of a developmental nature on educating for citizenship. While it is beyond the scope of this paper to attempt a developmental perspective on the affective and behavioral dimensions of citizenship, the area needs analytic and empirical exploration.

Rethinking the Goals of Social Studies Education and Social Studies Research

In this final section of the paper I will attempt to place social studies education and social studies research in a perspective that, in my judgement, is more realistic regarding the proper goals of social studies education. In addition, I will suggest an alternative perspective for the practice of social studies research that could result in more useful and data for the practice of social studies education.

Setting standards for the practice of social studies education toward which we strive as a profession and by which we are held accountable by society is essentially a question of values and not of research. That is, the educational process is designed by society and serves societal needs. The standards set for schools reflect the values of a given society. Thus, in the United States our democratic system of government requires citizens with knowledge, attitudes and behavior suitable for life in a democracy. As a result, the traditional and persisting emphasis in social studies education is on citizenship.

The research summarized in this paper cannot, nor should it, attempt to determine the debate over appropriate standards for social studies. The research can, however, provide a strong dose of realism regarding the feasibility of achieving the standards chosen. In addition, the research can assist the profession in evaluating competing views on the purposes of social education. Finally, if the profession is to be held accountable by society for its performance vis a vis these standards, research will shape this evaluation. Thus, standards represent a two-edged sword for the profession. On the one hand they reflect society's values, give direction and purpose to teachers and schools, and enlist public support for the endeavor. On the other hand, we are also judged publically by the extent to which we achieve those standards.

Excellence is the state of possessing good qualities to an eminent degree. Excellence in social studies education as traditionally conceived consists of practices that produce citizens possessing to an eminent degree citizenship knowledge, attitudes, and behaviors. I have reluctantly come to the conclusion that excellence may not be an achievable goal in social studies education if we continue to hold to a rational-activist/New England town meeting model of what it means to be a good citizen. The social studies profession may well be involved in a Catch-22 concerning this issue. That is, the higher we set our sights, the greater the inevitable gap between standards and outcomes. This paradox of citizenship education will not be easily resolved by newer, more effective strategies, for the etiology of political attitudes indicates that factors with greatest potential for forming those attitudes lies outside the influence of the profession. If the profession is hard pressed to demonstrate that it has substantial influence on other than cognitive outcomes, should it continue to hold itself accountable for these outcomes?

One question I wish to raise in the remainder of the paper is whether the social studies should set more modest standards for itself, namely focusing on cognitive goals. An affirmative answer to this question requires the demonstration that a more limited view on goals of citizenship education will not place democracy at risk.

Below I summarize in four points the arguments of a variety of authors whose observations suggest the more modest cognitive view of the goals of citizenship education is consistent with our current system of government:

1. Increased citizen competence may result in greater social and economic inequality. If all interests in society effectively use their resources, what hope is there for the downtrodden (Weissberg, 1981)? Class differences will likely be further exacerbated. Higher economic status children are more predisposed to participation the first place (Miller, 1985; Verba & Nie, 1972), and will be better able to master citizenship competencies. The inevitable result will be that lower SES interests will be blunted.

2. Active citizenship conceptions of democracy rest on unrealistic grounds in that they ignore the general desire to absolve oneself of decision-making responsibility in favor of protective decisions of a leader. The most important element of a democratic system is not citizen participation in all aspects of social and political life, but rather is in the election process through which nonelites choose governing elites. The public's role is to verify whether their political elites are practicing self or group interests (Dahl, 1956). Democracy best functions as competition among decision-makers for public support (Schumpeter, 1962). Our system doesn't depend on or require increased citizen participation. In fact, it runs better without it. Witness the recent litigation explosion. Has increased citizenship activism resulted in a more just society?

3. The political apathy of the public is a plus in that it prevents those with limited interest and expertise from creating undue stress on the system. Those with the least democratic attitudes participate the least. Nonparticipation is a positive goal in that it prevents those with limited interest and expertise from creating undue stress on the system (Berelson, 1952).

4. Participation and citizenship skills are normally associated with liberal or enlightened policies; however, there is no guarantee that increased participation might not result in more authoritarian and repressive policies as witnessed in the suggested social and political policies of the troika of Helms, Falwell, and Schafley (Weissberg, 1981).

The arguments summarized above suggest that should citizenship education be effective in its goal of producing an entire population of rational/activist citizens unintended consequences may result that conflict with basic democratic values. It is somewhat puzzling that this more limited view of citizenship, which has such dramatic implications for the conceptualization of social studies education, is seldom discussed by a profession committed to critical inquiry.

The more modest view of the goals of social studies education I wish to suggest places the greatest emphasis for practice and research on student knowledge, understanding, and reasoning regarding our history and system of democratic governance. If we are held to be publically accountable for fostering political attitudes and citizen skills in youth, we will have a difficult time making our case based on available practice and research. Clearly, increasing student knowledge and facilitating cognitive development are goals that the profession can achieve. For example, the profession can point with some satisfaction to the recent National Assessment of Educational Progress (NAEP, 1983) finding that on a wide range of questions involving knowledge and understanding of the workings of government student scores increased from age 13 to 17. In addition, all students showed improved performance between 1976 and 1982. Attempting to influence future citizens' attitudes and behavior is a much more complex and difficult task and one which the personal and political costs and benefits of the energies expended must receive careful attention.

Upon what then should social studies research focus? I would suggest that one area where research efforts should not concentrate is on the effects of social studies curriculum, *per se*. Such research will not be helpful, other than on a local basis, because we can safely say it is unlikely to yield educationally or socially significant results. The evaluation of social studies education, like the evaluation of education in general, requires longitudinal investigation. Clearly, this type of research lies beyond the abilities, energy, financial commitment and interest of school districts, graduate students, and faculty members. The necessary funding required alone makes it a highly improbable endeavor, and besides, we already know what the results would be—education does make a difference (Hyman & Wright, 1979). Therefore, the path to increasing knowledge in social studies education may be first to assume that it is desirable in itself and that its positive effects can largely be taken for granted. That the American experiment of democratic government has been successful is an understatement. Our political system has resulted in a highly stable and effective form of government; one that has insured a greater voice for the people, opportunity for self-improvement, and respect for human rights than any government previously known in human history. It is obvious that the educational process (formal and informal) has been effective in its role of citizenship education in that the system remains strong, flexible and vibrant. It can also be argued that the democratic orientations that develop in youth are, *ipso facto*, desirable, for the result of the educational and socialization process is a level of support and commitment that assures continued effective functioning of the system. An educational system that yielded greater or lesser political commitment, trust, or participation would put stress on the existing system with uncertain results. The political system works and the educational system contributes positively to it. Only utopian views can argue for change in either, and clearly at

this point in history such utopian visions pose greater risks than opportunities for our freedom and way of life.

What social studies research then is of most worth? In my judgement, a primary focus of social studies research should be on the study of exemplary programs as judged by the enthusiasm of students, teachers and community. Once such programs are identified, the focus of research should be to describe how the program works, why it is perceived as successful, and to the extent possible the effects of the curriculum on students (short and long term). Such research will blend qualitative and quantitative methods of inquiry. It is my hunch that this focus of research will result in an increased appreciation of the importance of the charisma, intellect, energy and character of the individual teacher. In my judgment this dimension of the practice of social education has been overlooked. Generally, the public and the profession stand to gain little from further curricular effectiveness research. In all likelihood, the results of such studies are in the aggregate highly predictable, especially with regard to political attitudes. Research does not need to buttress public support for the practice of social education. Public support exists for citizenship education and existing research, if honestly presented, could raise more questions and potentially do damage to that support. A focus on exemplary programs and rich descriptions of the infrastructure of those programs offer greater potential for understanding social education and communicating that understanding to practitioners, public, and policy makers.

In conclusion, I do not wish to suggest that there is no place in social education for well designed experimental research. As a profession, we have a responsibility to guard against the dangers of unbridled ego and uncritical demagogery that frequently accompany questions of curriculum theory, development, and evaluation. In addition, classical experimental procedures still offer the most reliable and valid method of adjudicating differences of opinion over many important questions in the field. However, until longitudinal research becomes common in the field of social education, we will gain only minor benefits from curriculum effectiveness research. In lieu of this kind of research, studies on exemplary programs and individuals in the field promise more productive means of expanding our view of effective social education.

Endnote

1. An earlier version of this paper was presented at the annual meeting of the Social Science Education Consortium, Wingspread Conference Center, Racine, WI, June 7, 1985. I wish to express my appreciation to the participants of that conference for their helpful comments.

References

Berelson, B. R. (1952). Democratic theory and public opinion. *Public Opinion Quarterly*, 16, 313-330.

- Conrad, D., & Hedin, D. (1981). *Executive summary of the final report of the experiential education project*. St. Paul, MN: Center for Youth Development and Research, University of Minnesota.
- Dahl, R. A. (1956). *Preface to democratic theory*. Chicago: University of Chicago Press.
- Durio, H. F. (1976). A taxonomy of democratic development. *Human Development, 19*, 197-219.
- Durkheim, E. (1925/1973). *Moral education*. Glencoe, IL: The Free Press.
- Dynneson, T. L., & Gross, R. E. (1985). An eclectic approach to citizenship: Developmental stages. *The Social Studies, 76*(1) 23-27.
- Ehman, L. (1980). Change in high school students' political attitudes as a function of social studies classroom climate. *American Educational Research Journal, 17*, 253-265.
- Gage, N. L. (1984). What do we know about teaching effectiveness? *Phi Delta Kappan, 66*, 87-93.
- Hyman, H. H., & Wright, C. R. (1974). *Education's lasting influence on values*. Chicago: University of Chicago Press.
- Hogan, R., & Mills, C. (1976). Legal socialization. *Human Development, 19*, 261-276.
- Jennings, K., & Niemi, R. (1981). *Generalizations and politics*. Princeton, NJ: Princeton University Press.
- Jones, R. S. (1974). Student political involvement and attitude change. *Social Science Quarterly, 55*, 439-450.
- Kohlberg, L. (1981). *The philosophy of moral development*. San Francisco: Harper & Row.
- Leming, J. S. (1985). Research on social studies curriculum and instruction: Interventions and outcomes in the socio-moral domain. In W. Stanley (Ed.), *Review of Research in social studies education: 1975-1983*. Washington, DC and Boulder, CO: National Council for the Social Studies and Social Studies Education Consortium.
- National Assessment of Educational Progress. (1983). *Citizenship and social studies achievement of young Americans: 1981-1982. Performance and changes between 1976 and 1982*. Denver, CO: Educational Commission of the States.
- Newmann, F., & Rutter, R. A. (1983). *Effects of high school community service programs on students' social development*. (Final report to the National Institute of Education, Grant No. NIE-G-81-0009). Madison, WI: Wisconsin Center for Educational Research.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Schlaefli, A., Rest, J. R., & Thoma, S. J. (1985). Does moral education improve

moral judgment? A meta-analysis of intervention studies using the Defining Issues Test. *Review of Educational Research*, 55, 319–352.

Schumpeter, J. A. (1942/1962). *Capitalism, socialism and democracy* (3rd ed.). New York: Harper & Row.

Tapp, J. L., & Kohlberg, L. (1971). Developing senses of law and legal justice. *Journal of Social Issues*, 27, 65–91.

Verba, S., & Nie, N. (1972). *Participation in America*. Ann Arbor, MI: University of Michigan Press.

Weissberg, R. (1981). The politics of political competence education. In D. Heater & J. Gillespie (Eds.), *Political education in flux* (pp. 127–148). Beverly Hills, CA: Sage Publications.

Effective Questions and Questioning: A Research Review

William W. Wilen
Ambrose A. Clegg, Jr.
Kent State University

Abstract

Questioning is considered an influential teaching act because it is the most basic way teachers use to stimulate participation, thinking and learning in the classroom. This research literature review focuses on the effective teaching studies conducted particularly during the past 15 years to determine the impact of teachers' questions and questioning techniques on students' learning outcomes. Eleven questioning practices that correlate positively with achievement are identified from five major reviews and are supported additionally with a sampling of individual research studies. Implications for future research in the social studies are discussed.

The purpose of this paper is to identify effective questioning practices synthesized from research findings related to the types of questions teachers ask and the questioning techniques they employ. Questions and questioning were selected because of their pervasive impact on stimulating students' participation in the classroom. Questions are the stock in trade of every social studies teacher. In fact it is hard to imagine a social studies teacher not asking questions. Research has demonstrated that teachers ask a high frequency of questions, about 300-400 during a typical day (Levin & Long, 1981). One social studies educator referred to the questions a teacher verbalizes in the classroom as "the most influential single teaching act" because of the power of the question to impact student thinking and learning (Taba, 1966).

The focus of this review is primarily on studies to determine the influence of teachers' questions and questioning techniques on student achievement. Since the turn of the century questioning has been a major concern of researchers. During the first 50 years research related to questioning focused on describing and evaluating teachers' use of questions in the classroom. For the next 20 years sophisticated methods of systematic observation and analysis were developed and used by researchers in classrooms to identify objectively teacher behaviors and to train preservice and inservice teachers in questioning skills. Starting about 1970, researchers turned their attention

toward determining the impact teacher behaviors have on student achievement. Generally referred to as process-product research, the findings from these correlational and experimental studies suggest specific procedures teachers can apply to increase student achievement. Although studies of teacher effectiveness have focused almost exclusively on the learner outcome of academic achievement, the impact of these findings on teacher education and inservice programs is growing (Wilten, 1982).

Each of the studies reported here relate to determining which types of teacher questions and techniques influence student achievement. The practices were synthesized from five major current reviews of research on effective teaching approaches that had questions and questioning as part of their focus. The eleven practices are then supported by specific research studies to provide information on the context of the findings. The studies reported were intended to be representative of the research conducted, not inclusive.

Presented below are eleven questioning practices that correlate positively with student achievement as synthesized from research conducted over the past 15 years. Each practice is supported, both as a finding of at least one major current review of research literature and by individual studies.

Effective Questioning Practice

1. Effective teachers phrase questions clearly (Berliner, 1984; Brophy & Good, 1985; Gall, 1984).

Clearly expressed and transmitted questions reduce the possibility of student confusion and frustration. If the question does not specify the conditions to which students are to respond, time is wasted attempting to determine what the teacher is expecting. A major source of ambiguity is the use of run-on questions. In this case two or more questions are asked in an uninterrupted series and the students do not know which question the teacher wants answered.

Studies indicate that there is only about a 50% relationship between the cognitive level of teachers' questions and students' responses (Dillon, 1982; Mills, Rice, Berliner, & Rousseau, 1980). Winne and Marx (1979) suggest that students' answers to teachers' questions may be incongruous because students either have not learned the content or, more likely, they do not understand what kind of mental activity is required to answer the question. To facilitate communication and make intentions clear to the students, teachers may use verbal cues. These prompts help students identify the thought level requested by a question. Another approach is to train students in a question classification system to understand more clearly the relationship between questions and thinking levels (Mills et al., 1980). Question clarity increases the frequency of precise and accurate responses from students.

2. Effective teachers ask questions which are primarily academic (Levin & Long, 1981; Weil & Murphy, 1982).

Nonacademic questions which are procedural, affective, or personal do not enhance student achievement. A review of several major studies found a positive relationship between student gain in achievement and focus on academic questions (Rosenshine, 1976, 1979). In their Project Follow Through evaluation study, Stallings and Kaskowitz (1974) found that low SES first and third grade students increased their performance over other students in the basic skill areas of reading and arithmetic when academically focused questions and responses were emphasized. Soar and Soar (1979) also obtained similar findings in urban classrooms of heterogeneous students in northern Florida and South Carolina for grades three through six.

3. Effective teachers ask high frequencies of low cognitive level questions with low SES students in elementary settings (Berliner, 1984; Brophy & Good, 1985; Levin & Long, 1981; Weil & Murphy, 1982).

Low cognitive level questions are those that require students to engage in low convergent thinking where the emphasis is on recalling knowledge or information. High frequencies of low level questions are characteristic of recitation, review, and drill.

Several studies indicate that low cognitive level questions promote greater achievement gains than do higher level questions, particularly with primary children in the basic skill areas of reading and arithmetic (Medley, 1977; Rosenshine, 1979). One major study (Soar & Soar, 1979) found that the frequency of broad, higher cognitive level questions correlated negatively with achievement and that lower level, convergent questions correlate positively with achievement. It was concluded that these questions were best for the wide range of SES urban elementary students involved in the basic skill areas. Gall et al. (1978) obtained a similar finding in an experimental study involving sixth grade students.

The major study supporting the finding related to the frequency of low level questions was Brophy and Evertson's (1976) two-year Texas Teacher Effectiveness Project. Involved were experienced second and third grade teachers in Austin, Texas where the student population in 70% Anglo, 17% Mexican-American and 13% Black. It was found that teachers who asked high numbers of low level questions had students who achieved more than the students of other teachers not asking high frequencies of questions. This finding was statistically significant with low SES students.

4. Effective teachers ask high cognitive level questions (Berliner, 1984; Brophy & Good, 1985).

High cognitive level questions are those that require students to engage in high convergent and divergent thinking. At the high convergent level students organize material mentally and apply learned information; comprehension and application levels of thinking. Low and high divergent questions require students to think critically about information and perform original and evaluative thinking, analysis, synthesis and evaluation levels of thought. Teachers generally use higher cognitive level questions in reflective discussions (Wilens, 1985).

Of all the questioning practices presented related to effective teaching research, this is the most tentative because of the conflicting findings. Two major meta-analyses were conducted on studies which examined the relationship between level of teacher questions and students achievement at a wide range of ability and grade levels. Winne (1979) reviewed 18 experimental and quasi-experimental studies and concluded that teachers' predominant use of low or high level questions had no influence on student achievement. Two years later Redfield and Rousseau (1981) reanalyzed Winne's 18 studies and added two others conducted since 1979 and drew the opposite conclusion. They found that gains in achievement were noticed in classrooms where higher cognitive level questions assumed a predominant role. Although the research findings are mixed, the latest review provides evidence that asking higher level questions is related to effective teaching.

5. Effective teachers allow 3-5 seconds of wait time after asking a question before requesting a response, particularly when higher cognitive level questions are asked (Berliner, 1984; Brophy & Good, 1985; Weil & Murphy, 1982).

The use of wait-time to stimulate thinking is particularly beneficial during discussions. Wait-time is generally not necessary during recitations, where the emphasis is on recall of information. When wait-time is increased to 3-5 seconds, students respond with more confidence, more appropriate and longer responses, higher cognitive level responses, and they tend to ask more questions (Rowe, 1974). Also, waiting 3 seconds after a student has responded increases the frequency of extended responses.

In two studies (Tobin, 1980; Tobin & Capie, 1982) involving 10-13 year old Australian and U.S. (Georgia) students in science classes, both forms of wait-time (wait after teacher question, and wait after student response) were studied. Wait-time correlated significantly with achievement.

6. Effective teachers encourage students to respond in some way to each question asked (Brophy & Good, 1985; Weil & Murphy, 1982).

This establishes the expectation that the teacher wants students to reflect and respond to questions. Probing can be used to get a minimal response and unanswered questions can be redirected to other students.

Several studies have found a negative relationship between students' failure to respond and achievement. Wright and Nuthall (1970) found that the higher the percentage of questions answered by students, the higher the achievement. They also found that redirection of questions to other students was significantly related to achievement. Dunkin and Biddle (1974), in their review, also found a positive relationship between redirection and achievement in the basic skill areas. Evertson, Anderson, Anderson, and Brophy (1980) involved urban seventh and eighth grade mathematics and English students in a study of teacher behaviors and student outcomes. They found that student failure to respond to questions correlated negatively with achievement. In contrast, Gall et al. (1978) found no relationship between redirection and achievement.

7. Effective teachers balance responses from volunteering and nonvolunteering students (Brophy & Good, 1985).

Nonvolunteering students' contributions should be encouraged when there is a high probability that they know the answer. Too often in recitations and discussions, only a few students actively participate. Calling on nonvolunteers forces all students to keep mentally attentive and involved.

Although some studies have found that teachers who depend heavily on volunteering students tend to produce higher achievement (Evertson et al., 1980), achievement gains have also been evident in classes in which an emphasis has been placed on nonvolunteering students. Anderson, Evertson, and Brophy (1979) found that ordered turns in small-group reading instruction resulted in higher achievement for experimental classes of first graders. One of the techniques of ordering turns was directing questions to nonvolunteering students.

8. Effective teachers permit student call-outs in low SES classes while suppressing call-outs in high SES classes (Brophy & Good, 1985).

Call-outs can be very useful in stimulating student participation and interest, especially in reticent classes. Call-outs need to be controlled more in eager classes. Brophy and Evertson (1976) found that relevant called out answers correlated positively with achievement in low SES classes. Calling out was discouraged in the high SES classes because of the role of the more assertive and confident student.

9. Effective teachers encourage a high percentage of correct responses from students and assist with incorrect responses (Brophy & Good, 1985; Levin & Long, 1981; Weil & Murphy, 1982).

A high percentage of correct responses can be attained in recitations when the emphasis is on student recall of information. It is assumed, of course, that the students have read or studied sufficiently to answer the questions. The teacher can assist students who are incorrect by acknowledging the correct portion of the answer and following up by rephrasing the question or giving clues.

In a suburban environment, Good and Grouws (1977) compared the teachers of high and low achieving fourth grade mathematics students. They found that teachers of high achievers asked fewer questions that received incorrect answers. When incorrect answers were received, teachers helped students obtain correct answers. Brophy and Evertson (1976) found that high SES students in urban Austin, Texas performed best when they answered correctly about 70% of the questions; low SES students did best when they answered 80% correctly.

10. Effective teachers probe students' responses for clarification, support for a point of view, or to stimulate thinking (Brophy & Good, 1985; Weil & Murphy, 1982).

This finding applied primarily to teachers conducting reflective discussions in which students' higher cognitive level thinking was encouraged. If students are not accustomed to responding at higher cognitive levels, their

initial responses might be incomplete, superficial or ambiguous. Probing helped students to clarify, support or expand initial responses.

Most studies support the use of probing questions. Brophy and Evertson (1976) found that probing improved low SES, but not high SES, second and third grade students' responses, and was significantly related to achievement. Probing was also an important correlate of achievement in the Clark et al. (1979) study involving sixth grade students. Feedback was used to make responses more complete or correct. Contrary to the other findings, Gall et al. (1978) found that the use of probing in recitation did not help students process knowledge or respond at higher cognitive levels.

11. Effective teachers acknowledge correct responses from students but are specific and discriminating in their use of praise (Brophy & Good, 1985; Weil & Murphy, 1982).

Teachers tend to use praise when a simple verbal or nonverbal acknowledgement to a student's correct or appropriate response is all that is necessary. Praise should be used genuinely, sparingly, and it should be specific. Good, Ebmeier, and Beckman (1978) compared two major sets of data based on studies conducted in urban settings. They found that praise of student responses related more to achievement in low SES situations and less in high SES classrooms. In the Teacher Effectiveness Project, Brophy and Evertson (1976) found that praise correlated positively with achievement in low SES classes but correlated negatively in high SES classes.

Thus far we have summarized the major research on questioning during the past 15 years. Where do we go from here?

Most of the studies cited above have dealt with a limited array of questioning techniques; effectiveness has often been tied to the criterion of measurable results on standardized tests. But there is still a larger domain of questioning that has not been adequately studied. This includes the more complex cognitive thinking tasks of conceptualizing, generalizing and hypothesizing, as well as those open-ended questions that seek creative solutions or multiple responses to problem solving. Another underexamined area is analysis, in which probing questions may help students identify competing value elements in important political, economic, or social issues.

It is a curious anomaly that effective strategies for teaching, such as higher level cognitive and affective strategies, have been widely available for many years. Taba's (1966) seminal studies provided detailed questioning strategies that effectively lead students through higher order thinking tasks. Yet, despite widescale dissemination efforts, as in the Addison-Wesley social studies series (Taba, Durkin, Fraenkel, & McNaughton, 1971), and advocacy in many texts on teaching methods (Armstrong, 1980; Banks & Clegg, 1985), such strategies were not tested in any of the major standardized tests used as evaluation measures in the studies reported above. Similarly, Hunt and Metcalf (1955) developed effective strategies for the study of those problems that society labeled as closed areas. The task ahead

for researchers is to examine the impact of questions within these strategies, many of which have been demonstrated to be effective in earlier, small scale studies. It is incumbent upon the research community to encourage the measurement of analytic skills in standard test batteries.

Little research has been devoted to the role of student questioning. Most of the studies cited above focused on a typical format of teacher question and student response, a teacher dominated strategy. What happens when students take charge of the strategy in group-oriented discussions and control both the type of questions asked and the data bank of answers? Many content areas such as social studies, language arts, and science claim to place a high premium upon group discussion, but there is little evidence that such discussions lead to high levels of independent thinking, critical analyses of previous students' statements, creative approaches to new issues, or divergent solutions to problems. Such classroom discussions, valuable as they may be, are seldom evaluated objectively, nor are they incorporated into the current models of standardized tests. Student-directed questioning strategies in classroom discussions, such as those suggested by Hunkins (1976), are a potentially fruitful and needed area of research in the years ahead.

In short, the previous research on classroom questioning has provided data on those approaches that appear most effective in rather traditional classroom settings. The challenge now is to study more intensely those questioning strategies that lead to higher order thinking, to value analysis, to creative responses to new situations, and to independent thinking in student directed discussions.

References

- Anderson, L., Evertson, C., & Brophy, J. (1979). An experimental study of effective teaching in first-grade reading groups. *Elementary School Journal*, 79, 193-223.
- Armstrong, D. G. (1980). *Social studies in secondary education*. New York: Macmillan.
- Banks, J. A., & Clegg, A. A. (1985). *Teaching strategies for the social studies* (3rd ed.). New York: Longman.
- Berliner, D. (1984). The half-full glass: A review of research on teaching. In Philip L. Hosford (Ed.), *Using what we know about teaching*. Washington, DC: Association for Supervision and Curriculum Development.
- Brophy, J., & Evertson, C. (1976). *Learning from teaching*. Boston: Allyn & Bacon.
- Brophy, J., & Good, T. (1985). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Third handbook of research on teaching*. New York: Macmillan.
- Clark, C., Gage, N., Marx, R., Peterson, P., Stayrook, N., & Winne, P. (1979). A

- factorial experiment on teacher structuring, soliciting and reacting. *Journal of Education Psychology*, 71, 534–552.
- Dillon, J. (1982). Cognitive correspondence between question statement and response. *American Educational Research Journal*, 19, 540–551.
- Dunkin, M., & Biddle, B. (1974). *The study of teaching*. New York: Holt, Rinehart & Winston.
- Evertson, C., Anderson, C., Anderson, L., & Brophy, J. (1980). Relationships between classroom behaviors and student outcomes in junior high mathematics and English classes. *American Educational Research Journal*, 17, 43–60.
- Gall, M. (1984). Synthesis of research on teachers' questioning. *Educational Leadership*, 42, 40–47.
- Gall, M., Ward, B., Berliner, D., Cahen, L., Winne, P., Elashoff, J., & Stanton, G. (1978). Effects of questioning techniques and recitation on student learning. *American Educational Research Journal*, 15, 175–199.
- Good, T., & Grouws, D. (1977). Teaching effects: A process-product study in fourth-grade mathematics classrooms. *Journal of Teacher Education*, 28(3), 49–54.
- Good, T., Ebmeier, H., & Beckman, T. (1978). Teaching mathematics in high and low SES classrooms: An empirical comparison. *Journal of Teacher Education*, 29(5), 85–90.
- Hunkins, F. P. (1976). *Involving students in questioning*. Boston: Allyn & Bacon.
- Hunt, M., & Metcalf, L. (1955). *Teaching high school social studies*. New York: Harper.
- Levin, T., with Long, R. (1981). *Effectiveness instruction*. Washington, DC: Association for Supervision and Curriculum Development.
- Medley, D. (1977). *Teacher competencies and teacher effectiveness*. Washington, DC: American Association of Colleges of Teacher Education.
- Mills, S., Rice, C., Berliner, D., & Rousseau, E. (1980). Correspondence between teacher questions and student answers in classroom discourse. *Journal of Experimental Education*, 48, 194–204.
- Redfield, D., & Rousseau, E. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research*, 51, 237–245.
- Rosenshine, B. (1976). Classroom instruction. In N. Gage (Ed.), *The psychology of teaching methods*. Chicago: University of Chicago Press.
- Rosenshine, B. (1979). Content, time, and direct instruction. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings and implications*. Berkeley, CA: McCutchan.
- Rowe, M. (1974). Wait-time and reward as instructional variables, their influence on language, logic and fate control: Part one—wait-time. *Journal of Research on Science Teaching*, 11, 81–94.

- Soar, R., & Soar, R. (1979). Emotional climate and management. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings and implications*. Berkeley, CA: McCutchan.
- Stallings, J., & Kaskowitz, D. (1974). *Follow-through classroom observation evaluation*. Menlo Park, CA: Stanford Research Institute.
- Taba, H. (1966). *Teaching strategies and cognitive functioning in elementary school children* (Cooperative Research Project No. 2404). San Francisco: San Francisco State College.
- Taba, H., Durkin, M. C., Fraenkel, J. R. & McNaughton, A. H. (1971). *A teacher's handbook to elementary social studies: An inductive approach*. Reading, MA: Addison-Wesley.
- Tobin, K. (1980). The effect of extended teacher wait-time on science achievement. *Journal of Research in Science Teaching*, 17, 469-475.
- Tobin, K., & Capie, W. (1982). Relationships between classroom process variables and middle-school science achievement. *Journal of Educational Psychology*, 124, 441-454.
- Weil, M., & Murphy, J. (1982). Instructional processes. *Encyclopedia of Educational Research* (5th ed.). New York: Macmillan.
- Wilén, W. (1982). *Questioning skills for teachers*. Washington, DC: National Education Association.
- Wilén, W. (1985). Questioning, thinking and effective citizenship. *Social Science Record*, 22, 4-6.
- Winne, P. (1979). Experiments relating teachers' use of higher cognitive questions to student achievement. *Review of Educational Research*, 49, 13-50.
- Winne, P., & Marx, R. (1979). Perceptual problem solving. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings and implications*. Berkeley, CA: McCutchan.
- Wright, C., & Nuthall, G. (1970). Relationship between teacher behaviors and pupil achievement in three elementary science lessons. *American Educational Research Journal*, 7, 477-493.

Computers in a School Environment: Preliminary Report of the Social Consequences

Richard A. Diem
The University of Texas at San Antonio

Abstract

Upper elementary school classrooms in four school field sites were visited weekly during a nine-month academic year to study student social behavior during computer instruction. Conclusions were drawn concerning roles played within student-initiated problem solving groups, role differentiation by sex, and apparent effects of computers on students' value structures.

In the past few years studies involving technology and its effects on school settings have focused mainly on performance gains criteria. The literature is now replete with papers showing effects on standard literacy measures of computer assisted instruction, computer managed instruction, and of late, the uses of microcomputers during instructional interactions (Alderman, 1978; Burns & Bozeman, 1981; Edwards, Norton, Taylor, Van Dusseldorp, & Weiss, 1974; Poulsen & Macken, 1978; Smith, 1973; Thomas, 1979). There are also a number of other articles and reviews relating the results of computer software applications in classroom situations (Edwards, 1982; Hunter, 1983). Recently, a series of books and articles discussing the nature of technology and its cultural influences on homes, businesses, and schools have also emerged (Bolter, 1984; Turkle, 1984). As these have appeared a new look at the social consequences of technological interventions in educational settings has begun to take place.

These initial inquiries have concentrated on the cultural impact that technology has had on educational environments in a global manner. That is, they have looked at technology and its change agent effects on a broad scale within school settings. These include the ways in which children interact with technology in a variety of school/social settings; language patterns and speech structure as related to technology; and behavior pattern development as a result of the imposition of technology (Diem, 1985; Diem & Searles, 1985; Hess & Tenezakis, 1970; Holmes, 1984; Ragosta, Holland, &

Jamison, 1982; Saracho, 1982). Most of these studies have used Pappert's microworld conceptual framework as the theoretical basis of their work (Pappert, 1980). Drawing on Piagetian ideas, these concepts relate how children are building their views of the technological world based on the information systems interactions they encounter.

Research Design

The study reported in this paper is drawn on questions developed out of these previous discussions. Using an ethnographic research design, an examination was conducted of the manner in which a technology, specifically microcomputer technology, affected social behavior patterns in upper elementary school classrooms. These behavior patterns encompassed social interactions between and among students, teachers, and administrators as well as value judgements about the technology that these groups elicited.

Beginning with the opening of a school year, the researcher visited four school field sites on a weekly basis. During these visits an observation schedule was developed and maintained. All events that were observed and recorded took place in specifically chosen classrooms. These included: (1) preclass computer laboratory instruction; (2) hourly computer laboratory and classroom instruction; and (3) postcomputer laboratory debriefing sessions. Following the school calendar, the observations continued throughout a nine-month period.

As an adjunct to the observational studies, informational data about the tracked students and school personnel engaged in working with these students were also gathered. This information included identifying data (age, sex, etc.), technological information (access to a home computer) as well as on-site interviews with various students, teachers and administrators. These interviews were conducted throughout the study and centered on various aspects of the observed socio-technical interactions.

Study Site/Participant Demographics

The four schools in this study are located in a culturally and economically diverse suburban community in a large Southwestern city. This school district was chosen because it had used computers in its elementary schools prior to this study, required its elementary school teachers and staff to be trained in computer literacy and classroom applications of technology, and had mandated that all of its sixth grade students be required to spend a minimum of two hours per week in computer literacy instruction. This instructional requirement was exclusive of state guidelines for mandated computer literacy training.

These variables were imposed as an external measure to eliminate school districts who were either just implementing computer technology or had not yet developed a computer literacy curriculum. This was invoked because it was felt that the types of social interactions that took place in an educational environment that had established techno-cultural interaction patterns

would differ significantly from one in which a cultural change agent, technology, was first being introduced.

Subjects were 138 students who participated in all, or parts, of this study. Due to shifts in class rosters and moves out of the school field site settings, only 121 remained constant from the study's inception until its conclusion. Descriptions of the demographic makeup of this study are based on this constant group, of whom 52% were female and 48% male. The median age was 11 years, 3 months with the range being 10 years to 13 years, 5 months.

At home, 73 (61%) of the students had access to personal computers. At school, 119 (99%) had previously used a computer. The school experiences had ranged from drill and practice lessons in mathematics and English, mainstreamed topics in special education to lessons in BASIC Computer Language program development in gifted and talented classes.

The teachers observed throughout this study were all female. Their ages ranged from 24 to 41 with the median age being 31. All had purchased their own computers and had previously taught computer literacy at the elementary school level for at least two years. Prior to teaching computer literacy three of the teachers had taught mathematics at the elementary level and one had taught science and reading.

Findings

The observational data that were gathered indicated that:

- (1) Students tended to cluster in groups as they solved technologically related problems. These clusters included an expert—a peer students used in an effort to overcome a particularly vexing problem, an experimenter—a peer who was willing to try almost anything in an effort to solve a problem in a quick manner, and several observers who tended to watch, similar to sidewalk superintendents, and be passive during technological interactions.
- (2) Students used technology as part of their social and value structure when dealing with problem solving exercises.
- (3) Students used technology during their social interactions.
- (4) Teachers used technology as part of their social reward structure.
- (5) Students understood both the social and economic value of the information they used and produced.

The above are listed in order of occurrence/frequency throughout the course of the study.

Discussion

Group problem solving efforts. At each of the field sites a school-wide computer laboratory had been operationlized. Using 30 stand-alone microcomputers and 5 printers, these laboratories had been available for instructional purposes for a year. Several teachers informed me that during most of the first year in the laboratory there had been a wide range of problems relating to classroom and instructional procedures and technological break-

downs. They all expressed confidence that the problems had been worked out and that they could now concentrate their efforts on instructional activities rather than on administrative and procedural ones.

Each of the computer literacy classes had a maximum enrollment of 30 so that each student could have exclusive use of a microcomputer. Machines were numbered and assigned to specific students. The students were told that they were to complete their assignments on their own machines. Any group work was to be preapproved, as were any extra efforts such as graphic, word processing, or programming activities. The teachers indicated that these procedures had evolved out of their previous experiences. They all felt that a structured instructional environment was needed to accommodate the wide range of academic and computer skills of the students, as well as for classroom control.

Despite admonitions to the contrary, group problem-solving schemes were initiated almost from the students' first encounters in the laboratory. With a single instructor monitoring 30 students it was virtually impossible to answer and assist individual questions at a rapid pace. Students who either understood the assignments or who had previous knowledge in computers aided their peers with little or no prompting from their teachers. They did so without rancor or disrespect to another's lack of technological expertise.

At the same time that instruction on machine usage and implementation was occurring, classroom and laboratory assignments also were being made. These were supposed to be completed on an individual basis but were rarely done so in that manner. Despite the threat, or encouragement, of the grading system, students actively assisted one another in all phases of problem solving activities.

This assistance came in the form of study clusters. These clusters usually consisted of three types of students: the expert, the experimenter and the observer. As previously described, the roles of these students were delineated early in the school year. However, throughout the study most of the observed students took on all of the role attributes within shifting clusters at one time or another, depending on the problem being studied or the social situation that evolved. The more demanding the problem, the more that students relied on one another for both intellectual and moral support.

Within the role structure males dominated as experts and experimenters. Females were not overtly dissuaded from assuming these positions. However, when students formed groups to solve problems it was very rare for a female student to assume a leadership position. This did not influence test or grade averages in these classes as these were comparable across sexes. The males who led these groups were looked upon as school-boy types. While they were not social outcasts, they did not impress their peers as being potential social leaders.

Social/value structure. Throughout the variety of instructional sequences

that were observed in the course of this study the inculcation of technology within the social and value structure of both students and teachers manifested itself in a number of ways. The most obvious was the ingestion of technological and computer jargon into daily speech patterns. Terms such as K and ROM became colloquialisms, with K referring to a person's intelligence or knowledge level. For example, if you have lots of K you're very smart. Or, did you get lots of K for the exam? That is, study for the exam? ROM, or ROM Head took on a negative connotation describing someone who was not too bright. In one class the use of this term almost started a fight.

As the students refined and developed their computer and technological skills, they began to understand how the machines they were training on might effect their lives in both an academic and vocational manner. Most realized that they would be encountering technology at higher educational levels. This might take the form of further computer course work or, for most, the acquisition of more technological skills to enhance their academic performance. When questioned about these academic ideals, students almost universally indicated that their parents were encouraging them in this direction. However, many of these same students whose sole interest was computers were also described by cohorts as probably nerds and having no friends.

There was a recognition in this community that technology is changing the workplace as well as the nature of work in our society. Parents of many of these students had jobs in either the civilian or military sector that directly tied them to the new technology. A high value on the use of technology in their parents' work was present among these children. Any negative aspects of technology and work such as technological displacement in a factory due to the introduction of robots or mechanized assembly lines had not yet been felt by this populace due to its physical locale.

Social interactions. In all school settings students quickly learn how to establish their own communications network. The microcomputers that were available to these students provided another avenue in this process. Despite careful monitoring of the computer laboratory by both teachers and staff, students were able to leave messages on the screens of the computers and print out special notes for selected class members. In fact several students who had learned how to produce graphics and illustrations did so in both positive as well as negative manners.

The messages and graphics took all forms. Love notes, derogatory comments, pictures of semiclad individuals, and homework all were seen throughout the study. Although no formal electronic bulletin board had been established, there was one in active use for many months. The teachers tried, on several occasions, to either curb or halt these practices by issuing threats or by institutionalizing them in a formal message center. Both of these efforts failed as the students avoided a formal communications structure. The only success at halting these efforts was when the laboratory was closed.

Social reward structure. The instructors observed in this study had not fully incorporated the technology they were using as part of their instructional delivery mechanisms. All of them indicated that it had taken a period of time before they were comfortable with the technology. Even though they knew how to use it, they had difficulty in picturing how students at the upper elementary level could acquire enough knowledge in a short period of time to develop any real expertise in using the computer. They rewarded any student who showed technological promise. Rewards were both overt as well as covert in nature. For example, some students were allowed to work ahead, given special assignments, or were dismissed from tedious or repetitious tasks.

The teachers also used withdrawal of technology for punishment. Some student were not allowed to use the equipment if they were disruptive. Others, however, were given before or after school makeup assignments in the lab if they misbehaved. These sanctions varied according to the offense that was committed.

Social and economic value of information. The students were told repeatedly that both the hardware and software they were using had cost their school district many tax dollars. This message was reinforced many times but came across in an especially forceful manner when the students were working with a costly piece of software.

Some of the students decided that if the software cost so much they might develop their own, sell it and become rich. Most did not realize the time that was needed to accomplish this task, but all could see the possible economic benefits. This effort related to some of the class discussions on the economic and social implications of technology that had been previously presented to them.

As part of class work, a unit on the use and possible misuse of information in our society was taught. This took the form of case studies on the abuse of information acquisition and the right to privacy, the study of laws relating to the stealing of computer files, as well as discussion of white collar crime as related to technology. The students understood the nature of these criminal acts, but failed to realize the ethical and moral dilemmas that might arise as one segment of a society controls information flow from another. Information censorship, and its possible societal ramifications, were not discussed with these students in either a formal or an informal setting.

Summary

Most school and educational systems now recognize that a basic understanding of computer technology is a necessary attribute for today's elementary and secondary student. The short and long range socialization effects of the introduction of this powerful medium on classroom instruction and group interaction have yet to be fully understood.

Within this study the introduction of technology seemed to reinforce the stereotypical roles of females and males as they used computers in group problem-solving activities. Males dominated as both leaders and experimenters in devising ways to harness technology for their own purposes. Interestingly the instructors, all female, did not try to augment or change this situation. If this attitude is prevalent within a broad spectrum of computer literacy instructors, stereotypes regarding females and computer usage will be reinforced and increasingly difficult to overcome.

It is also noteworthy that teachers, in this study, stratified their classes through the recognition of computer expertise among their students. While all classes were a mixture of high and low academic achievers, teachers developed an in-class tracking system that allowed advanced students to either work ahead or to be an assistant teacher or grader for the instructor. As more students acquire their own computer hardware this process may continue. In communities not as affluent as this one, economic barriers that prevent the self-purchase of technology may exacerbate this situation and put economically disadvantaged children at a technological disadvantage at a very early age.

Although the observed classes were in computer literacy, most teaching was done through traditional methods. Lecture, discussion, reading textual material and answering questions dominated as the major modes of instructional interaction. This was not surprising in view of the fact that most teachers, in any field, have not been taught how to use technology as a pedagogical mechanism. Many pre-and inservice teachers are being instructed in computer literacy but few are being taught computer and technological applications. Until this is done teaching styles will not change.

The information presented in these classes was current and very elementary. If these students do not take any other formal classes in computers or technology throughout their secondary school experience, will they have gained enough information to assist them in either the academic or vocational workplace? Better still, will they understand how their lives may change due to the introduction of new and potentially even more powerful home technology? Other studies will have to be developed to judge any carry-over effects in either the classroom or other social environments of these early computer literacy classes. What was evident in this study, however, is that for students to understand the social effects of technology they must discuss and gain information on this topic from the entire curriculum rather than in any one class. A compartmentalizing of technology will only reinforce the notion that total specialization is the only way one can gain access and expertise in this area.

The next few years will be crucial for computer and technological innovation in school settings. There are those who are ready to abandon the new technology by equating it with workbook style teaching and the purchase of televisions in the late 50s. The current generation of elementary and secondary students will have to face the impact of this technology in both aca-

demic and social terms. Providing adequate instruction and an understanding of the social consequences of technology is an essential element for our schools both for today and for the future.¹

Endnote

1. Another version of this paper was presented at the 1985 College and University Faculty Meeting of the National Council for the Social Studies in Chicago, Illinois.

References

- Alderman, D. L. (1978). *Evaluation of the TICCIT computer assisted instructional system in the community college* (Final report, Vol. 1, PR-78-10). Princeton, NJ: Educational Testing Service.
- Bolter, J. (1984). *Turing's man-western culture in the computer age*. Chapel Hill, NC: The University of North Carolina Press.
- Burns, P. K., & Bozeman, W. C. (1981). Computer assisted instruction and mathematics achievement: Is there a relationship? *Educational Technology*, 21(10), 32-39.
- Diem, R. A. (1985). A study of children's attitudes and reactions to the new technology. *Social Education*, 49, 318-320.
- Diem, R. A., & Searles, J. E. (1985). Social responsibility in technological environments. *International Journal of Computers and Education*, 10(1), 1-5.
- Edwards, J. (1982). Computer software in education. In H. S. Kepner (Ed.), *Computers in the classroom*. Washington, DC: National Education Association.
- Edwards, J., Norton, S., Taylor, S., Van Dusseldorp, R., & Weiss, M. (1974). Is CAI effective? *Association for Educational Data Systems Journal*, 7, 122-126.
- Hess, R. D., & Tenezakis, M. D. (1970). *The computer as a socializing agent: Some socioaffective outcomes of CAI*. (Tech. Rep. No. 13). Palo Alto, CA: Stanford Center for Research and Development in Teaching.
- Holmes, G. (1984). The computer and limitations. *Foreign Language Annals*, 4, 413-415.
- Hunter, B. (1983). *My students use computers*. Reston, VA: Reston Publishing Company.
- Pappert, S. (1980). *Mindstorms*. New York: Harper & Row.
- Poulsen, G., & Macken, E. (1978). *Evaluation studies of CCC elementary school curriculum, 1975-1977*. Palo Alto, CA: Computer Curriculum Corporation.
- Ragosta, M., Holland, P. W., & Jamison, D. T. (1982). *Computer-assisted instruction and compensatory education: The ETS/LAUSD study*. Princeton, NJ: Educational Testing Service.
- Saracho, O. N. (1982). The effects of a computer-assisted instruction program on basic skills achievement and attitudes towards instruction of Spanish speaking migrant children. *American Educational Research Journal*, 19, 201-219.
- Smith, I. D. (1973). Impact of computer-assisted instruction on student attitudes. *Journal of Educational Psychology*, 64, 366-372.
- Thomas, D. B. (1979). The effectiveness of computer-assisted instruction: Secondary schools. *Association for Educational Data Systems Journal*, 12, 103-116.
- Turkle, S. (1984). *The second self*. New York: Simon & Schuster.

Toward More Adequate Quantitative Instructional Research

Ronald L. VanSickle
University of Georgia

Abstract

Quantitative instructional research is stronger when research questions are investigated in the context of well-defined lines of inquiry. Lines of inquiry allow the application of alternative perspectives, such as a social perspective, the refinement of the operationalizations of independent and dependent variables, and the development of high quality measurement instruments. They help manage interpretation problems caused by threats to internal and external validity. They promote the generation of cumulative knowledge about the design and implementation of instructional methods. They produce patterns of findings that can reduce researchers' dependence on inferential statistics. These characteristics enhance the scientific and practical meaning of quantitative instructional research.

Quantitative instructional research can be an effective way to investigate and understand teaching and learning. However, the quantitative empirical-analytic approach to studying instruction has been implemented incompletely and inadequately so often that some critics regard its utility as an open question and others regard it as a dead end. Critics claim that methodological limitations, inappropriate statistical procedures, and intractable validity problems produce research findings that have little practical or scientific meaning. Nevertheless, there are examples of success in understanding and promoting instructional effectiveness through quantitative research. Four general questions are addressed in this paper. First, how can quantitative instructional research be conceptualized adequately? Second, how can threats to internal and external validity be managed? Third, what strategy will maximize the usefulness of instructional research for generating scientific knowledge about instruction and for generating knowledge relevant to practice? Fourth, how can statistical analyses be utilized productively? Each question is answered and illustrated with examples of research relevant to social education.

Research Conceptualization

Quantitative instructional research which is adequately conceptualized is characterized by a social perspective, systematic conceptual and operational

definition of concepts, and high quality measurement instruments. Social-psychological and social-structural classroom variables are investigated along with the more common mental and behavioral variables. The research questions include explicit and logically rigorous definitions of all independent and dependent variables. Measurement instruments are developed from the operational definitions and have strong reliability and validity, especially construct validity. These are key aspects of adequately conceptualized quantitative instructional research.

Social Perspective

A social perspective views learning as a function of relationships between individuals and groups within a classroom context. An individual student's academic motivation and expectations are shaped by social structures, sometimes overlapping, composed of status systems and their associated reward structures and roles. The characteristics and relationships of students as a group or groups in a classroom can structure the quantity and quality of teacher/student, student/student, and student/task interaction (VanSickle, 1982).

A social perspective might involve the treatment of the classroom group conceived as a whole instead of a collection of individuals. Social variables which characterize relationships within the group, such as status and reward structures, are manipulated in an effort to affect individual learning performance. In contrast, much experimental instructional research focuses on attempts to change individual students' mental characteristics (e.g., conceptual structures) or behavioral characteristics (e.g., mathemagenic behavior). Changes in these characteristics, it is hypothesized, result in changes in learning performance.

Allen and VanSickle (1984) utilized a social perspective and manipulated classroom reward structure in low achieving, homogeneously grouped, ninth-grade, world geography classes. The reward structure was altered experimentally by means of a team learning technique called Student Teams-Achievement Divisions (Slavin, 1980). Students were grouped in teams and they helped each other achieve instructional objectives. Students responded to quizzes individually and the teacher compared scores of students of past similar performance across teams. Points were assigned to each score depending on relative performance. Points were totalled to produce team scores. When the team learning/competition treatment was compared to an individual study/whole class discussion comparison group, a substantial achievement difference in favor of the experimental group was observed. This research illustrates how a socially conceived group treatment can affect individual student performances.

A social perspective should also be applied to instructional strategies in social studies education which have never worked as well as they should have, for example, inquiry teaching strategies. Due to the operation of social psychological or structural variables affecting student participation

many students might not have participated in the instructional interaction of classrooms being studied (VanSickle, 1979). Consequently, no instructional effects might have been observed for various instructional strategies because the treatments occurred sporadically, if at all, for many students. A generally unanswered question is: What social conditions must exist for a given instructional treatment to be implemented effectively? Perhaps we unknowingly assess instructional nonevents more often than we might expect.

Quantitative research focuses on a relatively small set of variables which are measured precisely and related; however, this limitation can be largely overcome by attending to a broader range of variables, such as social psychological and social structural variables. Also, the limitation of narrow scope can be overcome by studying instructional phenomena over a series of studies. In a set of related studies, the research problem can be explored in a variety of ways and from a variety of perspectives. Gradually a comprehensive understanding of the complex phenomena of interest can be developed. Pursuing lines of inquiry is of critical importance in realizing the potential of quantitative instructional research. The importance of well articulated studies in series will be emphasized repeatedly in this paper.

Conceptual and Operational Definition

Social studies instructional research will be conceptualized more adequately to the extent that researchers systematically define all variables conceptually and operationally so that they actually can vary. Categorical variables, such as inquiry strategy or democratic classroom, are the most common types of concepts found in instructional research. Hage (1972) recommended that dimensions of continual underlying categorical phenomena be identified for more comprehensive, generalizable understandings. Comparing categorical independent variables, such as simulation gaming and lecture/case study discussion, provides few insights into the nature of effective instruction. However, attention to variable instructional concepts, such as frequency of concept applications and frequency of interpersonal academic communications, could generate important clues. Even when no relationships are observed, a clearer idea emerges of what is not relevant than is the case with gross, categorical variables.

A study of simulation game design characteristics (VanSickle, 1977) illustrates this approach to systematic definition of variables. VanSickle believed participant attitudes were a result of whether participants were dependent on or independent of others for resources to play the game. In this case, the resource was money. Resource independence was originally defined as enough money to play the whole game. This was in contrast to no resource independence defined as not enough initial money to play the whole game. With a little more effort, the categorical notion of resource independence/dependence was converted to a variable concept. Degree of resource independence was then defined as the ratio of the number of decisions a participant could afford in a game to the number of decisions

required by the game. The number of decisions a participant could afford was found by dividing the resource units which a participant possessed at the beginning of the game by the cost of the participant's most expensive decision alternative. This produced a range from 0.0 to 1.0. This variable redefinition allowed finer analysis and possible application to other simulation-gaming situations.

Another example of the utility of variable concepts is based on the effort to clarify the nature of a democratic classroom as opposed to an autocratic classroom (VanSickle, 1983a). Several variables were identified underlying the democratic/autocratic dichotomy which were more informative than the categorical distinction. Among these variables were initiation rate, response rate, performance opportunity distribution, influence attempt rate, cross-status interaction rate, degree of consistency in student status rankings, proportion of instructional decisions made with student input, and number of within-class friendship choices. Attention to these variable concepts revealed the relevance of research related to them but not commonly considered in relation to democratic social relationships and citizenship education. The qualitative distinctions were reconceptualized in a potentially quantifiable form which clarified the original qualitative concepts.

Measurement

High quality measurement of instructional outcomes is critical for making practical instructional decisions and for understanding teaching and learning phenomena. Producing high quality measurement instruments is a difficult task. Unfortunately, it is a task that is not only difficult but also frequently slighted. Low instrument reliabilities are frequently reported with little or no comment by the authors in journal articles. Evidence for instrument validity is often omitted entirely. In a review of dissertations in social studies education, Larkins and McKinney (1983) observed that a majority of the dissertations, for which instrument reliability and validity were issues, did not report one or both. This lack of attention to instrument quality has serious implications for social studies education.

Inadequate instrument reliability can obscure relationships between variables; generally, relationships will be attenuated (Bohrnstedt, 1970). Studies reporting observations of no instructional effects or trivial effects for treatments might simply have missed observing effects that were there. A patient can have a fever even if an observer cannot detect it with his or her hand. If measurement instruments in a study have low reliability, then little can be concluded if no or trivial effects are observed. If an important effect was observed with an unreliable instrument, then there probably was an effect and it might well have been stronger than the observed effect. Since tests of statistical significance are a function, in part, of effect sizes, unreliable instruments increase the probability of accepting a false null hypothesis.

Inadequate validity of measurement instruments also confuses scientific and practical efforts. Does a low score on a test indicate a low level of a construct (e.g., geographic knowledge, decision-making skill) or does it indicate anxiety, low motivation, or social expectations for performance (Messick, 1981)? Without evidence that measurement scores only correspond to levels of a particular construct, it is difficult to interpret scores meaningfully.

Another validity problem involves scales with multiple dimensions. An instrument might be developed to assess a single construct, such as political efficacy. Actually, it might assess several constructs relevant to political attitudes and beliefs. This is what Stentz and Lambert (1977) found when they empirically evaluated widely used political efficacy scales. Four factors emerged: (1) belief in the responsiveness of public officials; (2) belief in the existence of a means of influence besides voting; (3) feeling that the government is not too complicated to understand; and (4) belief that average citizens influence the course of government. They also observed that there were insufficient items to form adequately reliable scales. Their analysis provided the basis for a major improvement in instrument quality which would promote more meaningful research and evaluation in citizenship education. Stentz and Lambert's research is an excellent example which needs to be emulated.

Given that the reliability and validity of many measurement instruments used in instructional research are inadequate, what should be done? There are numerous technical ways of producing high quality measurement instruments and they can be found in conventional measurement sources. Improvement in measurement quality will come only when researchers and developers collectively expect it, and require it. Few researchers want to develop instruments for a living; most want to answer questions important to them and the profession. However, much greater emphasis must be placed on instrument development if quantitative instructional research in general is going to mean much either practically or scientifically. Low quality instruments produce unacceptable and largely unnecessary levels of confusion in the research and evaluation literature.

Threats to Internal and External Validity

Whenever a hypothesis is investigated, there are always rival hypotheses to explain the observations. If no effect was observed, those who truly believe in their research hypothesis are tempted to invoke alternative hypotheses to explain why the expected effect did not materialize. If an effect is observed, there are several factors which might have been operating systematically to produce it other than the treatments or a hypothesized but unmanipulated factor. Campbell and Stanley's *Experimental and Quasi-Experimental Designs for Research* (1963) is the conventional source for identifying rival hypotheses, such as history, maturation, and testing ef-

fects. Their charts of research designs and potential threats to internal validity are familiar to virtually everyone who attempts an educational research project.

If a finding survives scrutiny in terms of the inquiry's internal validity, the question arises: To whom and to what situations does the finding apply? That is, what is the external validity of the study? Do the subjects compose a sample which represents a larger population of interest (e.g., high school history students)? Alternatively, do they represent some small local population? Do they just represent themselves? Do the subjects operate under fairly common schooling conditions? Shaver and Norton (1980) observed that random selection from a population is seldom reported in social educational research and that descriptive data are usually inadequate to identify researchers' target populations. Most instructional research is conducted under such constraints that there generally are uncontrolled threats to internal validity and few or no plausible claims for much external validity.

The Allen and VanSickle (1984) study cited earlier is a useful example for considering these validity issues. Intact classes were used, although, a plausible case was made that students were assigned randomly to classes. However, internal validity was threatened because the histories of the two classes could have been different. There may have been a serious persistent discipline problem in the comparison group which depressed achievement. The teacher's enthusiasm may have differed from class to class depending on the treatment. The experimental group obviously was experiencing a new instructional process while the comparison group received the teacher's conventional teaching efforts. Also, the subjects were randomly selected at best from that high school's low achieving ninth grade students. Consequently, some threats to internal validity are not ruled out and there is nothing about this study to warrant generalizing the findings very far. After these observations, the conclusion seems reasonable that the findings of an achievement effect for the experimental treatment could be regarded as simply a curiosity except by those who already believe in the hypothesis.

This study has a quality that most instructional research inquiries do not have; it is part of a well-defined line of inquiry. The Allen and VanSickle study was a systematic, not an exact or direct, replication of several other studies. They used a standardized instructional treatment under somewhat different conditions than previous studies to investigate an ambiguous area in the research literature. The previously identified achievement effect was not consistently observed in social studies classes as it was in other subject areas. Allen and VanSickle suspected, based on the literature, that in previous social studies inquiries, the cognitive level of the instructional objectives was higher than in the mathematics and language arts studies. Consequently, they kept their objectives at the knowledge level. Also, the team learning treatment was designed originally for use in classes composed of students of widely varying ability unlike the treatment groups in the Allen and VanSickle study.

Slavin (1983) reviewed 28 experiments that studied team learning conducted in elementary and secondary schools. The type of team learning studied was similar to the Allen and VanSickle experimental treatment in that group rewards were based on team members' individual learning performances. Superior achievement for the team learning treatments were reported in 89% of the studies. The Allen and VanSickle achievement finding fits into a pattern of similar findings and has considerable meaning and credibility because it does not stand alone.

The plausibility of the research hypothesis and of the generalizability of the finding both were boosted due to the context of a set of inquiries. The line of inquiry was also strengthened by the additional finding and by another test of the limits of the finding's generalizability. While nothing is ultimately proven, it is increasingly difficult to believe that team learning, as defined in these studies, is not the cause of strong achievement effects across a broadly defined population of elementary and secondary school students.

This example illustrates one major value of a line of inquiry. Common and often unavoidable threats to internal validity can lose their plausibility over a series of studies. The lack of external validity in any particular study can become a diminishing concern as succeeding studies are conducted with subjects from various parts of a larger population of interest. Unfortunately, direct and systematic replication is a rare research strategy in social educational research and educational research generally (Shaver & Norton, 1980). If most studies in social education were replications and extensions in well defined lines of inquiry, the interpretive confusion in the research literature caused by weak internal and external validity could be reduced.

Instructional Research Strategy

A strategy to maximize the usefulness of instructional research for generating scientific and practical knowledge about instruction is to focus research on the effects of a range of variables which compose an instructional method. Over the course of a series of studies, variations in the design of a method are assessed in terms of important educational outcomes. Knowledge is produced which defines operationally what well designed and effective mean for a particular method. This kind of instructional research is rare. Nevertheless, it is a prerequisite for designing effective instruction and for making absolute and comparative assessments of instructional effectiveness.

One approach to constructing empirically grounded operational definitions of effective design in relation to a particular instructional method is: (1) to identify key features in an instructional method; (2) to conduct experiments in which those features are varied; and (3) to assess the relative effectiveness of those variations in terms of learning outcomes. Research on concept teaching illustrates this approach. For example, McKinney (1985) designed a set of treatments using knowledge generated in a line of inquiry

oriented in terms of Merrill and Tennyson's concept teaching model (1977). Presentation of definitions, presentations of examples and nonexamples, and questioning regarding the classification of instances as examples and nonexamples were varied in the treatments. McKinney observed that the full model was no more effective than the treatments using parts of the model with the exception of the definition only treatment. As a result of research like McKinney's, knowledge regarding the design of effective concept instruction is becoming increasingly refined. This approach, exemplified in research on concept teaching, can be applied to other instructional methods.

Another way to construct empirically grounded operational definitions involves comparing a particular instructional method with other methods of instruction that differ in key ways. For example, mastery learning (Bloom, 1976) was developed and refined in experimental contrasts with non-mastery instruction that did not include certain features of mastery learning (e.g., formative tests). Team learning techniques, such as Teams-Games-Tournament and Student Teams-Achievement Divisions, were similarly improved through comparisons (Slavin, 1983). Slavin and Karweit (1984) further illustrated this approach when they experimentally compared mastery learning and team learning. Both instructional methods utilize focused instruction (i.e., regular schedules of teaching, worksheet completion, and quizzes). They differ in terms of team work, team rewards, formative tests, and correctional instruction. Slavin and Karweit reported that team learning was more effective than mastery learning instruction, and focused instruction and mastery learning were equally effective. The effectiveness of components in both instructional methods was probed and clarified further in this research. The comparative approach, while less direct than the single method approach, can also be an effective method of generating knowledge about high quality instructional design.

Shaver (1979) observed that interactions between instructional method, teacher characteristics, student traits, and situational factors are rarely considered. A set of studies on lecturing by Berliner (cited in Gage & Berliner, 1974) illustrates the importance of attending to interactions. Berliner assessed student acquisition of knowledge regarding Chinese history from lecture in relation to students answering questions periodically, taking notes, or simply paying attention. Questioning produced the highest achievement; however, unexpected patterns of achievement within the note-taking and non-note-taking groups were observed. Further research revealed that students characterized by relatively low short-term memory capacity achieved at higher levels if they did not take notes. The opposite was true for students with relatively high short-term memory capacity. Without explicit attention to such method/student trait interactions and sufficient replications to assess the consistency of such interactions, instructional research will not be very productive for either scientific or practical purposes. Interactions can

be explored by using the two previously discussed approaches to constructing empirically grounded operational definitions of effective instructional design.

Another less powerful strategy to produce instructional knowledge is to compare instructional methods with the goal of stating a general conclusion about the methods' effectiveness. This commonly used strategy is intended to be relevant to practice. However, it is simplistic because it is based on the assumption that the methods compared are well designed. What are the characteristics of a well designed lecture, inquiry lesson, or simulation game? Sufficient knowledge exists to operationalize only a few instructional methods effectively, such as mastery learning, concept teaching, and team learning. However, even these methods have areas of ambiguity regarding key design characteristics. Most methods, however, lack the empirically grounded knowledge base needed to operationalize the meaning of well designed. This strategy can produce data useful for evaluating components of an ongoing instructional program in a specific situation. However, the strategy of comparing instructional methods with the goal of drawing general conclusions about effectiveness is not likely to be effective.

Statistical Analysis

Statistical significance tests have limited interpretive value, but they are often used as grounds for extensive claims. Their meaning needs to be understood more widely so they will be used appropriately. Also, focusing on the presence or absence of a level of statistical significance (e.g., $\alpha = .05$) distracts attention, sometimes completely, from what actually was observed. It is useful to place more emphasis on generating and interpreting descriptive statistics. Effect size indexes are needed especially to interpret the scientific and practical meanings of observations.

Tests of Statistical Significance

When a researcher observes a relationship between two or more variables in a study, questions arise immediately about why the relationship was observed. Is the observation in this particular situation an example of a common phenomenon? Would the relationship be observed in other situations similar to this situation? Was this observation produced by a systematic process or a randomly occurring set of conditions? Tests of statistical significance were invented to probe these questions and help researchers avoid claiming that effects regularly and systematically exist which actually are due to chance conditions. In order to emphasize caution, the null hypothesis of no relationship between variables was selected as the focus of the tests. If a test of statistical significance indicates that an observation is very improbable assuming the null hypothesis is true, then the logic of the testing process leads to the rejection of the null hypothesis and to the consideration of systematic reasons for the observation.

A test of statistical significance produces a statement of probability which is the ratio of (a) the number of scores in a set equal to or greater than a certain magnitude to (b) the total number of scores in the set (Glass & Stanley, 1970). For example, a score equal to or greater than X occurs 5 times in a set of 50 scores (i.e., probability = $5/50$ or $1/10$ or $.10$). A statistical significance test *assumes* that the null hypothesis is true; that is, there is no systematic relationship between two or more variables. Therefore, if a score (e.g., a correlation coefficient or the difference between group means) equal to X is observed in the example above, a researcher knows that observing that score is not very probable if the null hypothesis is true. If the null hypothesis is not true then the probability of observing a score equal to or greater than X might be much higher; however, a statistical significance test does not address this situation. A statistical significance test indicates the probability that an obtained effect will be observed assuming the null hypothesis is true.

Statistical significance is a criterion set by researchers to indicate that a score would be sampled rarely from a population (e.g., 1 in 20 or $.05$). Assuming the null hypothesis is true, if the probability of observing an obtained score is low (e.g., $p < .05$), then a researcher makes a conventional judgment and rejects the null hypothesis. This is an inferential leap based on the plausibility of alternative interpretations of the data. The plausibility that the obtained score was simply a sampling error and does not fairly represent the population is weakened by the test. The plausibility that the score was produced systematically is strengthened by the test. Nothing is proven by a test of statistical significance; it is simply a constraint on the tendency of researchers to perceive order they expect or want in data that are produced randomly.

Tests of statistical significance are often misinterpreted. A test of statistical significance indicates the probability (p) that a given observed effect (D) will be observed assuming ($/$) the null hypothesis (H_0) is in fact true. It can be stated symbolically, $p(D/H_0)$. That is all that a test of statistical significance can indicate. Carver (1978) identified four misinterpretations. One misinterpretation is the belief that the probability value obtained from a test is the probability that the null hypothesis is true assuming an effect as large as the one observed; that is, $p(H_0/D)$ rather than $p(D/H_0)$. A second misinterpretation holds that a statistical significance test indicates the probability that the finding will be observed again (R) in a replication of the study; that is, $p(R/D)$ rather than $p(D/H_0)$. A third misinterpretation uses the results of a statistical test to determine erroneously the probability that the research hypothesis (H_1) is true given an effect as large as the one observed; that is $p(H_1/D)$ rather than $p(D/H_0)$. A fourth interpretation error involves equating statistical significance with scientific or practical importance. However, a probability value does not indicate the magnitude of an effect; $p = .001$ does not refer necessarily to an observed effect that is larger than one associated with $p = .05$.

One application of statistical significance tests is based on the assumption that a sample was selected at random from a population of interest and that replications of the study will also use samples randomly selected from the same population. In this way, there can be, for example, a ratio of 1 in 20 ($p = .05$) that an obtained effect of a certain size will be observed if the null hypothesis is true. Note that the ratio applies to a set of replicated studies. Without the assumption of random selection from a given population, a ratio of occurrence for a score of a certain magnitude is meaningless. Without random selection, a probability value can not be determined because the theoretical distributions underlying the test statistics assume randomness (Shaver, 1980). Since few research samples are selected randomly from known populations, there is no point in conducting statistical significance tests on data from nonrandom samples if one's purpose is to estimate the probability that a given observation would have been obtained due to an atypical sample selected randomly from a population.

Random selection is only one aspect of randomizing in a study; there is also random assignment to treatment groups or categories. If random assignment is performed, a test of statistical significance can provide useful information. In an experiment using a nonrandomly selected sample, there is an important question regarding the effectiveness of a treatment in terms of the dependent variables. Did the observed effect occur as a random result of assigning subjects to treatment groups or as a consequence of the treatment or some other systematic factor (e.g., history, maturation, mortality). This is a question of internal validity in the study (Campbell & Stanley, 1963). If 50 subjects are assigned randomly to two groups, there are many other possible ways the 50 subjects could have been assigned to two groups. The assignment used in the study is only one of a population of possible assignments. A test of statistical significance can provide an indication of the probability of an observed effect's occurrence in the population of all possible assignments of subjects to treatment groups assuming the null hypothesis is true (Winch & Campbell, 1969).

Winch and Campbell (1969) outlined the logic of this approach to statistical significance testing. How plausible is it that two sets of subjects in an experiment are really a homogeneous group? That is, if the null hypothesis of no treatment effect is true, what is the probability that an effect equal to or greater than the observed effect would be obtained? (1) Divide the subjects and their scores into all the pairs of subsets possible or a specified number of randomizations. (2) Construct a sampling distribution of the equally probable subset pairs. (3) Determine whether the observed mean difference of the particular subset of pairs in the experiment is well within the sampling distribution or toward one of the tails of the distribution. If the probability of occurrence of the effect is very low, then the plausibility of the null hypothesis is so weak that it is rejected conventionally. The explanation for the observed effect can be sought in terms of the treatment or some threat to internal validity. Typical tests of statistical significance are

good approximations of the literal randomization described above (Winch & Campbell, 1969; Edington, 1966).

From the perspective explicated above, the question addressed by a statistical significance test is: If there is no true effect, then what is the probability that subjects randomly assigned to the treatment conditions or categories will evidence an effect at least as large as the observed effect? If the probability is very low, then the null hypothesis can be rejected and interpretation of the results can proceed. This is the view of statistical significance testing that appears most widely appropriate in social educational research and educational research generally. Random assignment of subjects to treatment groups or categories is critical. Fortunately, some intact school groups are assigned in essentially random ways. Without random assignment, the results of statistical significance tests can not be meaningfully interpreted.

A further complication emerges because statistical significance might not be obtained because the research design is characterized by insufficient statistical power. Statistical power is the probability that a null hypothesis will be rejected ($\text{power} = 1 - \text{probability of a Type II error}$). In a survey of statistical power in published social educational research reports, VanSickle (1983b) observed that the studies were collectively very weak for small effects and weak for medium effects. Consequently, the absence of statistical significance could be a result of a small observed effect or a small sample size. Statistical significance tests must also be interpreted in light of a study's statistical power.

Descriptive Statistics

If greater than usual emphasis is placed on generating descriptive statistics, then researchers and research consumers are likely to interpret the results of studies more effectively. Score distributions, means and standard deviations, and correlation coefficients enable readers to obtain a clearer idea of what actually happened in a study than analysis of variance tables. Descriptive statistics also allow the computation of effect size indexes.

Effect size indexes can be useful in interpreting the scientific or practical meaning of a study's findings.

An effect size is the degree of departure from the null hypothesis or, in other words, the degree to which a phenomenon is manifested (Cohen, 1977). Cohen devised a set of effect size indexes for various inferential statistics. Shaver (1979) recommended the use of Eta^2 to estimate the proportion of the variance in the dependent variable accounted for by group membership on the independent variable. Glass (1977) recommended as an effect size index the difference between the treatment and control group means divided by the control group standard deviation. Over a series of studies, a pattern of descriptive statistics, especially effect size indexes, will be more important for interpretation than the statistical significance test of a single study.

Conclusion

Social educational instructional research is stronger when research questions are investigated in the context of well defined lines of inquiry, even if that means fewer questions are studied. Lines of inquiry allow the use of alternative perspectives, refinement of the operationalizations of independent and dependent variables, and development of high quality measurement instruments. They help manage the interpretation problems caused by threats to internal validity and limited external validity in single studies. They also promote the generation of cumulative knowledge about the design and implementation of instructional methods. Also, they produce patterns of findings that can reduce researchers' dependence on inferential statistics when they make judgments about the meaning of observations. Quantitative instructional research as a means to scientific and practical knowledge will be more productive when lines of inquiry with the characteristics discussed here are the rule rather than the exception.

References

- Allen, W. H., & VanSickle, R. L. (1984). Learning teams and low achievers. *Social Education, 48*, 60-64.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.
- Bohrnstedt, G. W. (1970). Reliability and validity assessment in attitude measurement. In G. F. Summers (Ed.), *Attitude measurement* (pp. 80-99). Chicago: Rand McNally.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Carver, R. P. (1978). The case against statistical significance testing. *Harvard Education Review, 48*, 378-399.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Edgington, E. S. (1966). Statistical inference and nonrandom samples. *Psychological Bulletin, 66*, 485-487.
- Gage, N. L., & Berliner, D. C. (1974). *Educational psychology*. Chicago: Rand McNally College Publishing.
- Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education* (Vol. 5). Itasca, IL: F. E. Peacock.
- Glass, G. V., & Stanley, J. C. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Hage, J. (1972). *Techniques and problems of theory construction in sociology*. New York: Wiley.

- Larkins, A. G., & McKinney, C. W. (1983). Reviews of dissertations in social studies: 1980. *Journal of Social Studies Research*, 7 (Monograph No. 1).
- McKinney, C. W. (1985). A comparison of the effects of a definition, examples and nonexamples on student acquisition of the concept of "transfer propaganda." *Social Education*, 49, 66, 68-70.
- Merrill, M. D., & Tennyson, R. D. (1977). *Teaching concepts: An instructional design guide*. Englewood Cliffs, NJ: Educational Technology Publications.
- Messick, S. (1981). Evidence and ethics in the evaluation of tests. *Educational Researcher*, 10(9), 9-20.
- Shaver, J. P. (1979). The usefulness of educational research in curricular/instructional decision-making in social studies. *Theory and Research in Social Education*, 7(3), 21-46.
- Shaver, J. P. (1980, April). *Readdressing the role of statistical tests of significance*. Paper presented at the meeting of the American Educational Research Association, Boston.
- Shaver, J. P., & Norton, R. S. (1980). Populations, samples, randomness, and replication in two social studies journals. *Theory and Research on Social Education*, 8, 1-10.
- Slavin, R. E. (1980). *Using student team learning*. Baltimore: Center for Social Organization of Schools, The Johns Hopkins University.
- Slavin, R. E. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin*, 94, 429-445.
- Slavin, R. E., & Karweit, N. L. (1984). Mastery learning and student teams: A factorial experiment in urban general mathematics classes. *American Educational Research Journal*, 21, 725-736.
- Stentz, M. C., & Lambert, H. D. (1977). An empirical reformulation of political efficacy. *Theory and Research in Social Education*, 5(1), 61-85.
- VanSickle, R. L. (1977). Resource independence and decision making procedure specificity in simulation games. *Journal of Experimental Education*, 46, 47-55.
- VanSickle, R. L. (1979). Neutralizing status constraints on student performance in small group activities. *Theory and Research in Social Education*, 7(2), 1-33.
- VanSickle, R. L. (1982). A social perspective on student learning. *Theory and Research in Social Education*, 10(2), 21-31.
- VanSickle, R. L. (1983a). Practicing what we teach: Promoting democratic experiences in the classroom. In M. A. Hepburn (Ed.), *Democratic education in schools and classrooms* (pp. 49-66). Washington, DC: National Council for the Social Studies.
- VanSickle, R. L. (1983b). Statistical power and effect size in social educational research. *Journal of Social Studies Research*, 7(2), 1-17.
- Winch, R. F., & Campbell, D. T. (1969). Proof? No. Evidence? Yes. The significance of tests of significance. *The American Sociologist*, 4, 140-143.

Journal Information

Theory and Research in Social Education is designed to stimulate and communicate systematic research and thinking in social education. The purpose is to foster the creation and exchange of ideas and research findings that will expand knowledge about purposes, conditions, and effects of schooling and education about society and social relations.

We welcome manuscripts on a variety of topics including:

Purposes of social education;

Models, theories, and related frameworks concerning the development, diffusion, and adoption of curricular materials;

Instructional strategies;

The relation of the social sciences, philosophy, history and/or the arts to social education;

The politics, economics, sociology, social psychology, psychology, anthropology, philosophy, and/or the history of social education;

Alternative social organizations and utilizations of the school for social education;

Comparative studies of alternative models of social education;

Models of and research on alternative schemas for student participation and social action;

Relationship of different pre- and in-service patterns of teacher training to social education;

Models of the utilization of objectives in social education and related research findings;

Implications of learning theory, child development research, socialization and political socialization research for the purposes and practice of social education;

The relationship of different independent, explanatory variables to educational achievements in the area of learning about society and social relations;

The social climate and cohesion of schools and other school characteristics as independent, explanatory variables predicting general achievement.

Manuscripts

Manuscripts submitted to *TRSE* should conform to *Publication Manual of the American Psychological Association: Third Edition*. Five copies of each manuscript should be submitted. Send manuscripts to:

A. Guy Larkins, Editor *TRSE*
Department of Social Science Education
University of Georgia
Athens, Georgia 30602

Subscription Information

A subscription to *Theory and Research in Social Education* may be obtained by membership in the College & University Faculty Assembly of the National Council for the Social Studies. Institutional and non-CUFA subscriptions are \$35.00 per year. Membership and subscription information are available from the Membership Department, NCSS, 3501 Newark St. NW, Washington, DC 20016.

Back Issues Reprints

Back issues may be obtained for \$10.00 each when available. Write to: NCSS, 3501 Newark St. NW, Washington, DC 20016.

Change of Address/Missing Issues

Send change of address notices and a recent mailing label to: Membership, NCSS, 3501 Newark St. NW, Washington, DC 20016. Missing issues are also available from this address.

Advertising

Advertising information may be obtained from: Peter Stavros, Director of Meetings and Marketing, NCSS, 3501 Newark St. NW, Washington, DC 20016.

Theory and Research in Social Education

NCSS

**3501 Newark St., NW
Washington, DC 20016**

Second Class Postage
Paid at Washington, D.C.
and additional
mailing offices