THEORY AND RESEARCH
in Social Education

Vol. X No. 3 Fall 1982

Martorella  Cognition Research: Some Implications for the Design of Social Studies Instructional Materials

King  School Uses of Materials Traditionally Associated with Children's Play

Peterson and Nett  Sexuality in the Classrooms of Teachers with Various Sex Role Orientations

Napier and Hepburn  Social Studies Project Evaluation: Case Study and Recommendations

Cornbleth  Book Review: The Myth of Educational Reform
Reviewers for this Issue of TRSE

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

James Akenson, Tennessee Technological University
Jean Anyon, Rutgers University
Beverly Armento, Georgia State University
Buckley Barnes, Georgia State University
Sáumel Bell, Bradley University
Catherine Cornbleth, University of Pittsburgh
Lee Ehman, Indiana University
Janet Eyler, Vanderbilt University
Jack Fraenkel, San Francisco State University
Judith Gillespie, Indiana University
Carole Hahn, Emory University
Michael Hartoonian, Wisconsin State Education Department
Mary Hepburn, University of Georgia
Wayne Herman, University of Maryland
Joseph Hurst, University of Toledo
A. Guy Larkins, University of Georgia
William T. Lowe, University of Rochester
Lawrence Metcalf, University of Illinois
Sharon Muir, Oklahoma State University
John Napier, University of Georgia
Murry Nelson, Pennsylvania State University
Thomas Popkewitz, University of Wisconsin
Howard Schober, Louisiana State University
Lynn Schwab, University of North Florida
James Shaver, Utah State University
Judith Torney-Purta, University of Maryland
Ronald Van Sickle, University of Georgia
Jane White, University of Maryland
## CONTENTS

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peter H. Martorella</td>
<td>Cognition Research: Some Implications for the Design of Social Studies Instructional Materials</td>
<td>1</td>
</tr>
<tr>
<td>Nancy R. King</td>
<td>School Uses of Materials Traditionally Associated with Children's Play</td>
<td>17</td>
</tr>
<tr>
<td>Elizabeth L. Peterson and Emily M. Nett</td>
<td>Sexuality in the Classrooms of Teachers with Various Sex Role Orientations</td>
<td>29</td>
</tr>
<tr>
<td>John Napier and Mary Hepburn</td>
<td>Social Studies Project Evaluation: Case Study and Recommendations</td>
<td>41</td>
</tr>
<tr>
<td>Catherine Cornbleth</td>
<td>The Myth of Educational Reform: A Study of School Responses to a Program of Change</td>
<td>63</td>
</tr>
<tr>
<td>Abstracts</td>
<td></td>
<td>67</td>
</tr>
<tr>
<td>Dissertations in Progress</td>
<td></td>
<td>69</td>
</tr>
</tbody>
</table>
From the Editor

Unfortunately for TRSE, Stuart Palonsky will be unable to continue as Associate Editor. Stu had been among the few involved in the original issues of the journal while he was a graduate student at Michigan State, and he served CUFA and TRSE with distinction as Associate Editor during the 1981–82 year. Stu's current activities preclude his continuation in editing, and he has asked for some relief from these duties. His efforts, pithy comments and wit will be sorely missed when we are thrashing over manuscript and printing decisions. More importantly, his intellect and keen insights will be missed in editorial judgments. We hope to keep him involved in TRSE work and look forward to his friendly criticism.

Fortunately for TRSE, we have been able to convince Kenneth Carlson to become Associate Editor. The editorial "we" is used in the preceding sentence to indicate that I had help from Stu, my wife and others to cajole Ken into this work. Ken has been Associate Editor of Social Science Record, the New York State Social Studies Council journal, and continues contributions to the field through publications in a variety of journals and multiple presentations at NCSS, CUFA, and the AERA SIG in Social Studies conferences.

An additional good fortune for TRSE is the willingness of William Fernekes to become Editorial Assistant for 1982–3. Bill is a doctoral student at Rutgers and a teacher at Hunterdon NJ Central High School.

As a final note, TRSE published an index to articles published during the period from Volume I (1973) to issue 2 of Volume IX (1981) in the Spring, 1982 issue. (TRSE, Vol. X, No. 1) We have heard of one error in that index and would ask that you check the index. If you find any errors, please let us know. We also seek your comments on the journal itself.

Jack Nelson
Cognition Research: Some Implications for the Design of Social Studies Instructional Materials

Peter H. Martorella
Temple University

Introduction

A basic purpose of this paper is to target a significant group of policy makers whom research appears to have little influenced, namely, those who design instructional materials (Schramm, 1955, p. 158). More narrowly, the focus of my analysis will be confined to one kind of instructional material, textbooks, one aspect of research, cognition, and one subject matter area, social studies. If there is a dominant, overriding characteristic of instruction in American schools that transcends subject areas, teacher orientations and regional school variations, it is our obeisance to textbooks as instruments of basic instruction (Shaver and others, 1979; Clifford, 1978; Weiss and Place, 1978). They have been scorned, reviled and ridiculed—often unfairly—scrutinized endlessly in the field of social studies and almost always found to be deficient in some important respect, but they endure (Fitzgerald, 1979; Mayer, 1962; Cronback, 1955) and continue to serve a major role in instruction, as teachers themselves testify (Weiss, 1979). Their role takes on even greater significance at the elementary level if one draws certain inferences from the fact that among all required subjects taught at that level, social studies is near the bottom in terms of the amount of the teachers' time that is spent on instruction (Weiss, 1979). By inference, the text may carry a disproportionate responsibility for instruction when compared with its secondary counterparts.
It is important to emphasize at the outset that the analysis is not intended as another attack upon insensitive Philistine publishers and their author collaborators. There are constraints of size, cost and media format under which texts are produced which are enormous and which will likely continue into the 21st Century or until alternative technological developments render texts financially infeasible.

Furthermore, our concerns do not lie with the extent to which descriptive analyses of texts have shaped or modified their present character. Historically, the careful monitoring by researchers of preceding generations of texts has often alerted publishers to distortions, inaccuracies or oversights and consequently often contributed to needed revisions. Such descriptive research generally is concerned with objectives, political exigencies, and other such standards, but has little to do with clearly verifiable effects of texts. For example, it may be argued plausibly that a text which outlines female stereotyped career roles (as one popular contemporary first-grade social studies text does) should be modified. It is a different issue, however, to argue from research what the effects of such material on cognitive learning are or how such material should be designed.

For purposes of our analysis, we will consider texts—for good or for bad—as instructional givens in thousands of classrooms across the United States for many years to come. Our task will be to consider several major areas in which cognition research would seem to have pregnant implications for the conscious design of social studies texts and to sketch the ways in which the structure of texts might be shaped if research were applied.

Books Compared with Textbooks

One set of assumptions upon which this analysis is based concerns the instructional distinctions perceived between textbooks and books in general. One of the more obvious ways that a textbook differs from its more conventional book counterparts is that it normally is designed to be a basic, complete and balanced survey of the topic it covers, appropriate for the audience it intends to reach. As a basic medium of instruction, a textbook carries a heavier burden than other resources which are intended to be just one source of information among many. Since learning is the primary intended purpose of a textbook, a more precise definition of the potential users is important as a basis for its construction. Typically, characteristics of users that have been used as bases for design have been prior experience with other correlated texts in a series, grade level, and reading levels which have been defined by mechanical formulae.

Perhaps in most striking contrast to conventional books, texts incorporate a variety of instructional aids, ranging from statements of objectives, summaries of chapters, questions inserted within text or at the ends of chapters, activities or projects to be done in conjunction with text, and tests—self or externally administered. The integral function of the text is to
instruct consciously in concert with some cognitive, affective or psychomotor objectives that have been determined for the users.

As texts are primarily instructional resources, issues concerning the instrumental efficiency of the design characteristics of a textbook in producing learning are paramount. In developing a conventional book, issues of cosmetic appeal, narrative plot developments, the fluid chronological flow of events, and the potential for interesting serendipitous insights may be overriding ones. The acid test of the validity of the features and organization of a textbook is the extent to which it efficiently and effectively achieves intended learning objectives. Gagné (1968) stated the matter quite succinctly, “Literary standards are not the same as pedagogical or instructional standards.” Other researchers have made similar observations (Macdonald-Ross, 1977, 1978; McConkie, 1977; Rothkopf, 1972, 1976).

Cognition Research and Text Design

All designers of instructional materials make some assumptions, explicitly or implicitly, concerning the cognitive capabilities of the intended audience. Similarly, a set of stated or implied objectives frame the design of a text. Since students must act upon the data provided in texts in order to learn it, the issue of which learner cognitive characteristics will be considered in the design of instructional materials would seem to be a critical one (Wittrock, 1979). To discover meaning in subject matter, students must use information processing strategies, attention and motivation mechanisms and memory devices, to suggest but a few of the cognitive processes involved. Which of these processes the text author considers and makes provision for in the design of the text and the extent to which alternatives are offered for learners of equal ability but with differing cognitive processing strategies are likely to influence the effectiveness of the instruction for a given student (Wittrock, 1979).

Along with the challenge of matching text design and cognitive characteristics of users, there is the issue of whether authors apply established instructional principles to teach certain classes of typical text objectives which they state, such as the learning of facts and concepts. A variety of such principles exist and the extent to which an author employs them in shaping the design of a text should affect in some way the cognitive learning which results from them.

Research into the fundamental questions of how individuals think generally, typically approach a learning task, process data, retain information, and most efficiently achieve a given learning objective, and how they often differ in performing many of these cognitive tasks has yielded a number of well grounded generalizations and hypotheses for investigation. Many of these appear to have implications for how social studies textbooks might be designed more effectively or for showing in which directions research on optimal design characteristics might proceed.
Some of the more salient findings from cognition research which seem fruitful to pursue for implications have been lumped arbitrarily under four headings: information processing styles, developmental differences, retention of key material, and instructional principles. The literature in the fields of cognition research which fall under these four headings is extensive and accumulating rapidly, and a comprehensive review is well beyond the scope of this article. The intent, rather, is to draw from a large and growing body of established findings and to suggest selective implications and some further productive research avenues concerning the design of social studies text materials.

**Information Processing Styles.** There are a variety of learner characteristics or styles which affect how an individual perceives data, remembers subject matter and solves problems (Messick, 1970; Robinson and Gray, 1974; Thornell, 1976; Witkin and others, 1977). While research into several areas subsumed under this broad topic is still in an embryonic stage—for example, the topic of individual styles of processing and answering questions (Lehnert, 1978; Norman, 1973; Schank, 1975), two closely related areas have an established research literature: field independence/dependence and conceptual tempo.

**Field independence/dependence.** Using measures developed by Witkin and his associates (1977), an individual's dominant cognitive style may be classified as being either field-independent or field-dependent. The former style is characteristic of those who are able to abstract parts of the whole (or text) as separate from the whole and who can impose their own structure into a situation or problem, if one is lacking (Witkin and Moore, 1974). By contrast, a field-dependent individual looks to the whole for guidance in how to process parts of the whole, and is unable to isolate material without help from the context of the surrounding environment (Tyler, 1974). Field-independent are more task oriented, more likely to require less structure in solving a problem, and better able to break a larger problem down into its component parts than their field-dependent counterparts (Witkin and others, 1977). Field-dependents, on the other hand, are more likely to have a "with people" orientation, to be more interested in what others say or do, to enjoy being physically close to others and to like to work with them in groups (Witkin and Moore, 1974; Holly, 1972). When material to be learned lacks structure and requires the learner to provide organization for mastery of the subject matter, field-dependent students are likely to have more difficulty than those who are field-independent (Witkin and Moore, 1974).

**Conceptual tempo.** Conceptual tempo measures tap a different dimension of information processing, indicating the basic way in which an individual approaches the resolution of a problem and the relative degree of success consistently experienced in generating solutions (Greer and Blank, 1977). A
student who typically takes time to ponder the answer to a problem and who also usually gets a high percentage of correct answers would be designated as having a *reflective* conceptual tempo (Kogan, 1976; Haskins and McKinney, 1976). Those who respond quickly and typically with a high error rate are characterized as *impulsive*. They often do not take sufficient time to analyze a problem or question carefully and consider all the information available or consider which are the alternative answers possible (Messer, 1970). Among the characteristics that reflective students seem to exhibit are better attention spans and better toleration of frustration than impulsive students (Block, Block and Harrington, 1974; Kogan, 1976). While reflectivity in students seems to increase with age generally, the issue of whether the characteristic of impulsivity will be retained in later years, with or without further intervention, awaits more study (Tyler, 1974).

**Developmental Differences.** Matching text design with the developmental characteristics of learners requires a complementary but different dimension of organization. Among other things, the findings on developmental differences in students reveal in which ways we can expect students to vary in their ability to handle certain cognitive tasks as a function of their maturational level. Over the span of the elementary and secondary years, students make transitions from one mode of learning to another enroute to adulthood. Whether the passage is characterized in stages or in some other form, youngsters exhibit qualitative shifts in the way in which they approach and transact thinking generally, the processing of information, and the resolution of problems (Kohlberg and Turiel, 1972; Lickona, 1976; Erikson, 1964; Piaget, 1926, 1950, 1971; Sigel and Cocking, 1974). As children mature, they move from a preference for kinesthetic or motor modes of learning during preschool years, to preferences for visual or imagery modes, and then for verbal modes of learning (Bissel and others, 1971; Bruner, 1973). Symbolic modes, which texts and printed materials represent, are the most complex forms, and for some students present difficulties well into adulthood. Bruner (1973) has characterized these three distinct sequential ways through which learning can occur and information can be represented as enactive (learning through doing), ikonic (learning through imagery) and symbolic (learning through abstract symbols or words).

A study involving social studies material (Meinke and others, 1975) offers some support for the theory of transitional preferences in modes of learning. It was found with fourth, sixth, and eighth graders as subjects that abstract thinkers did significantly better than concrete thinkers on criterion measures of learning of the concepts freedom, justice, nonfreedom, and nonjustice. Perhaps most significantly, scores improved as a function of grade level. Other investigators also have demonstrated that children were better able to remember pictures than the names of objects (Sommer, 1978; Levin, 1976).
Retention of Key Material. There are many facets to memory and retention of key material, and the need to retrieve and reproduce subject matter with a low error rate is obviously one important component of social studies instruction. While instruction which emphasizes memory to no significant learning purpose is intellectually bankrupt, that which ignores the utilitarian value of selective data retention is impractical. Strategies which facilitate the retention and recall of information expedite a learner's task and provide useful tools for work in social studies.

The simple process of cueing students in the form of preliminary questions or questions embedded in texts or other assortment of props has been shown to improve students' performance in recalling cue-related material (Anderson and Biddle, 1975; Frase, 1972, 1978; Kaplan, 1976).

A number of investigators have underscored the value of the systematic use of imagery in enhancing the retention of information (Attwood, 1971; Bellezza, 1981; Pavio, 1971; Perlmutter and Myers, 1975; Sommer, 1978; Wittrock, 1979; Levin, 1981; Levin and others, 1982; Pressley and others, 1982). Students, for example, who wished to remember various battles of the Civil War might visually imagine Bull Run as a bull, Gettysburg as a Getty gasoline station, and so on.

Techniques which when used enhance the recall and retention of subject matter are referred to as mnemonic devices. One mnemonic technique known as the "keyword method" employs both imagery and an associated linkage of images. Using this technique, fourth- and fifth-grade children in one study were taught states and capitals (Levin and others, 1980, 1982). Students learned an imagery referrent (for instance, "apple" for "Annapolis," "marry" for "Maryland") and were shown illustrations of the two image referrents interacting (as an example, apples getting married), and learned more state capitals than control subjects (Levin and others, 1980). Similar findings have been reported for related subject matter dates (Pressley and others, 1982).

Another more traditional mnemonic device is the "first-letter" mnemonic, wherein words are remembered either by arranging their final letters in some form of alphabetical order or by creating a new word with the list of first letters (Blick and Boltwood, 1972; Gruneberg, 1973; Bellezza and others, 1977). Often several different specific mnemonic devices may be combined in one technique. For example, to learn the list of American presidents in correct chronological order one may remember the rhyme, "Watch a jolly man make . . ." (Washington, Adams, Jefferson, Madison, Monroe).

Instructional Principles. All instruction, regardless of its objectives, implicitly reflects some organizational scheme and pattern of how information is to be sequenced. Point B follows point A; element X is presented in one fashion and not another; some points that the author or instructor might
have included were ignored while others were chosen; a great deal of time is spent on some aspects of a topic, very little on others. Such designs may be consciously constructed or may be injected haphazardly.

Social studies texts generally incorporate a variety of learning objectives, among them the learning of facts and concepts. For these two typical classes of objectives, some instructional guidelines exist and each implies a different pattern of how text material should be structured.

Although the results of studies are checkered (Hartley and Davies, 1976), some evidence exists that explanatory overviews or "anchoring ideas" which summarize in advance how factual material is related can facilitate learning (Ausubel, 1968; Steinbrink, 1970; Allen, 1970; Mayer, 1976; White and Mayer, 1980; Beeson, 1981). General principles for organizing subject matter to facilitate the learning of facts have been widely reported and summarized (Dixon and Horton, 1968; Eggen, Kauchauk and Harder, 1979; Gagné, 1977, 1978; Klatzky, 1975; Rothkopf, 1970; Underwood, 1959, 1964).

A set of principles for efficient fact learning that may be inferred from these investigations is summarized as follows. In factual learning, a learner first discriminates what is to be learned among related stimuli (that is, the key data are isolated or objectives are stated). Often, some mediator or structuring scheme is provided to help organize the material to be learned (for example, a mnemonic device for remembering the list of Presidents in correct order). Practice in learning the material occurs next (for instance, students try to recite or write correctly the list of Presidents). Feedback is provided (that is, correct responses are acknowledged or provided, as required). Reinforcement for learning is provided through devices such as praise, tests, and structured tasks. Some overlearning is encouraged through further practice to maintain retention of the factual material.

There are a number of extensive reviews of research related to instructional patterns for concept learning (Bank and others, 1982; Fowler, 1980; Glaser, 1968; Sax, 1969; Gagné, 1977; Klausmeier, 1976; Martorella, 1971a, 1971b, 1982; Levin, 1974; Markle, 1975; Markle and Tiemann, 1976). The net result is a rather extensive set of prescriptions concerning concept instruction. If findings from the various investigations were incorporated into an instructional design, it might be summarized as follows. The author or instructor must begin with a clear and conventionally accepted definition of the concept. At some point, instruction should include the concept definition and/or its primary attributes (for example, polygamy: the attributes, a custom and the having of two or more husbands or wives). Varied examples and nonexamples of the concept should be provided and some strategy should be included to help the learner determine which discussions pertain to examples and which to nonexamples. Unless students know the critical attributes of a concept (for instance, custom as an attribute of polygamy), they should be taught them prior to or concurrently with instruction. Some
opportunities to experiment with identifying examples and nonexamples should be incorporated, along with feedback on the correctness or incorrectness of the responses. Finally, where young children are involved, pictures or other visual stimuli, rather than strictly verbal stimuli, may facilitate concept learning.

Implications of Research for the Design of Social Studies Texts

Reviewing some of the salient research findings under the four general headings, information processing styles, developmental differences, retention of key material, and instructional principles, a number of possible implications for the design of social studies texts arise. Some implications are more specific than others, while still others are very general and will require research for more detailed guidelines. Several topics also suggest further research possibilities for validation by investigators dealing with social studies text materials, since many of the existing findings were generated in other subject matter fields and with nontext materials.

Information Processing Styles. Given the fact that individuals vary along field-dependent/field-independent and reflective/impulsive dimensions, and that, while no particular field dimension is “better”, the characteristic of reflectivity is preferable to impulsivity, two general guidelines for text designers emerge. Problems, tasks and projects could be alternately structured to place field-dependent and field-independent students on an equal footing, and along with questions and assignments, steps could be taken to encourage and reward reflectivity and penalize impulsivity.

Addressing the issue of field-dependence/independence in text design requires an author to accept the assumption that alternate ways or processes for solving a problem are acceptable and that one solution mode is not better than another. For some author objectives, this assumption very well may not hold: for instance, when a student is being taught the importance of being able to tolerate ambiguities in social situations and look for many different shades of opinion rather than “right” or “wrong” ones. On the other hand, a clear implication of the condition of field-dependent or field-independent dominance among individuals is that we process data, including those in texts, more efficiently when it has been structured to accommodate our preferred style. Moreover, we are more likely to feel comfortable with a learning task when it has guidelines which match with our style. Finally, we probably will be happier (and possibly more productive) when we work alone or in a group, depending upon whether we are, respectively, field-independent or field-dependent.

Put another way, the author who constructs a text exclusively with an open-ended structure or with a heavy demand upon the reader to organize the subject-matter places field-dependent students at a disadvantage.
This type of student may need more explicit instructions with problems in assignments and may require more specific statements about what is expected. Field-independents, however, will be better able to handle and may even prefer loose or nonexistent guidelines.

Some of these issues may be assessed directly in providing alternate structures within the design of the text material to accommodate both types of students. Others obviously must be incorporated into the teacher's manual or into the sets of questions, projects and assignments typically found at the ends of chapters or in supplementary workbooks.

Promoting reflectivity and discouraging impulsivity in text design, at a minimum, suggests periodically reinforcing directly through prompts, reminders and text insertions the perspective that one should give a serious question some careful thought before responding to it. In addition, models of deliberate and analytical question answering and problem solving can be provided to demonstrate reflective behavior. Sample problems with specific steps to reaching a reflective, compared with an impulsive, answer to a problem could be included in each chapter. In discussing the general question in a fifth grade text "What was/were the cause(s) of the Civil War?", for example, students might be prompted with a series of sub-questions such as, “Have you considered social, political and economic reasons?”

The basic issue for the text designer to confront is that one type of reader is, as a general rule regardless of topic or interest level, likely not to take sufficient time to think carefully about the nature of the information or the alternative responses that he/she might make to questions. Further, since reflectivity increases in general with age, the problem is likely to be most pronounced at the elementary level. Given that most social issues and questions of importance require reflective analysis, the social studies text designer's challenge is to build into the questioning material a pattern of responding which forces impulsive students to override their normal behavior.

Developmental Differences. Ideally, a sensitivity to the developmental characteristics of learners would provide suitable learning modalities, whether they conformed to textbook constraints or not. The basic instructional implication of modality shifts is that as children develop, they should be phased primarily through activities that emphasize learning through acting on phenomena, learning through visual forms of material, and, finally, learning through abstract or symbolic forms. Currently, such considerations do not appear to be consistently reflected in typical social studies classrooms, witness the dominance of basal texts as a medium of instruction. Perhaps the most important implication of developmental research for the social studies lies in its recognition of the differential capabilities of students at varying age levels to process highly symbolic or abstract data of the type customarily found in written texts. Developmental research sug-
gests that learning problems will occur as long as basal texts, which by their nature emphasize symbolic forms, continue to dominate and remain the nearly exclusive medium of social studies instruction through the elementary and middle-school grades.

From a developmental perspective, texts would rely heavily upon pictorial content throughout the elementary grades. The visual data would be used primarily to represent the basic narrative, rather than for cosmetic purposes or to augment or reinforce symbolic text. Imagery material in texts, in concert with three-dimensional materials and experiential activities, would be augmented by symbolic, printed text as a tertiary source of subject matter. As students progressed through the middle and secondary school years, the ratios of the three modes of representing subject matter would gradually reverse the orders of priority and text design would reflect this shift.

For secondary students, where subject matter was particularly complex or abstract—as an example, the economic principles of inflation and deflation or supply and demand—enactive and ikonic modes of instruction would be provided as alternatives, along with printed explanations. Where secondary students appeared to lag developmentally, visual text materials which were equivalent to their heavily symbolic text counterparts would be available.

**Retention of Key Material.** Various popular social studies textbooks frequently include a variety of cueing devices in student texts and teacher manuals to highlight some of the material considered by the author to be important to remember. Where it is not practiced, its use on a systematic basis could facilitate greater retention of subject matter.

Author-generated mnemonic devices throughout each chapter, such as the keyword method, the first-letter technique and others, would assist students in identifying important material, as well as aid them in remembering it. Since devices such as the keyword method do not often mesh well with conventional narrative discussions, they could be included in the review/study segments often found at the conclusion of chapters in texts. Similarly, some mnemonic techniques could be reserved for the teacher's manual for review sessions. Since the specific associations in effective mnemonic devices often are best constructed by those who are to remember the material, some provisions in text or teacher manual could be made to encourage students to apply the techniques themselves.

A variety of mnemonic techniques exist for facilitating retention of key material, and they are easily adapted to any topic in social studies. Since authors (and teachers) often wish students to remember a vast array of often isolated social studies items, they can better achieve this objective by including in text periodic memory aids.

**Instructional Principles.** There are many sets of instructional principles which relate to how specific classes of objectives might best be achieved. We have considered two such classes commonly included by all social studies text authors: facts and concepts.
The most direct and obvious implication for authors is that matches between instructional principles or guidelines and objectives is desirable to make learning more efficient. A corollary is that different classes of objectives require or draw upon different instructional principles for their achievement. Put another way, designing instruction in text to teach facts requires application of a different set of principles than if the objective is to teach a concept. Overriding these principles and attempting to teach in the same way both facts and concepts as instructional objectives within the general design of prose narratives in a typical social studies text is, of course, possible. The issue, however, is how efficient in terms of student learning time is such an approach.

The two sets of instructional principles presented earlier offer different guidelines for how text passages to teach facts and to teach concepts may be structured by an author. They require authors to think carefully about which actually are the salient facts and concepts to be taught within each chapter. They also present some new literary problems of how to reconcile the requirements of designing an effective instructional text with the literary expectations of a standard prose book. Applying instructional principles to text design requires some trade-offs in terms of narrative continuity and chronological development.

Conclusion

If the textbook remains—as it seems likely to—the primary instrument of social studies instruction, its producers will continue to bear a heavy instructional responsibility. Perceived as an instructional program rather than merely as a book, it has many objectives to achieve.

Research can play a much broader role in the design of social studies texts than its traditional important one of providing acceptable standards of scholarship. Both a sense of instructional principles and attention to important characteristics of potential readers should shape the structure of our texts. Existing cognition research provides some implications for how our texts might alternately be shaped to achieve more effectively some of the objectives which authors and teachers state. It also suggests further areas of needed research for social studies educators to undertake in order to verify the specific effects of various text designs on learning outcomes.

A selective sample of some of the more salient findings from cognition research was reviewed under four headings: information processing styles, developmental differences, retention of key material, and instructional principles. Illustrative implications of these findings for the design of social studies textbooks were presented. There is much that can be done to design social studies texts to achieve more consistently what we say we wish to accomplish, given the cognitive characteristics of the students who must use them. Cognition research offers some new directions in which future studies of social studies texts might proceed.
References


Fitzgerald, F. Onward and upward with the arts: Rewriting American history, I, II and III. The New Yorker, 1979, 55.


Klausmeier, H. J. Instructional design and the teaching of concepts. In J. R. Levin


Levin, J. R. *What have we learned about maximizing what children learn?* Unpublished paper, University of Wisconsin, Development Center for Cognitive Learning, Madison, Wis. ED 101 318.


Markle, S. M. *They teach concepts don't they?* *Educational Researcher*, 1975, 4, 3-9.


Mayer, R. E. *Some conditions for meaningful learning for computer programming: 14*


School Uses of Materials Traditionally Associated with Children's Play

Nancy R. King
University of Maryland

Introduction

Recent work in curriculum theory emphasizes the need to understand schooling in the context of the larger society and to explore the ways in which societal forces are manifest in schools. While a mechanical or strictly reproductive role for schools is inadequate to explain the relationship of schools to other social institutions, it is clear that schools do function, in part, to prepare children for their future participation as adults in a variety of social institutions. In addition to the knowledge and skills necessary for successful participation in adult life, schools also impart values, norms, and attitudes which shape children's world-views.

It is essential that curriculum designers and teachers, particularly in the area of social education, understand the impact of these socializing messages. The social organization of the classroom contributes to the curricular meanings found in academic experiences, and the relationship between this "hidden" curriculum and the academic curriculum must be explored for an adequate understanding of either. Further, the socializing messages embedded in classroom interaction often concern issues, such as the nature of work and the meaning of rules, which are also the focus of academic instruction in social education. It is all the more important to understand the impact of the social setting in which lessons such as these are shared.

The socialization of school children is not accomplished without negotia-
tion and compromise among the classroom participants. As anybody who has taught knows, teachers do not simply impose a new set of beliefs on acquiescent, passive children. Rather, various aspects of the children's culture are incorporated and transformed in negotiating a viable procedure for conducting classroom events. In the process, the curriculum and the teacher's instructional strategies are also shaped. The reciprocal outcomes of these negotiations are often overlooked because the consequences are not equally profound. That is, the children's perspectives on classroom events and behaviors in classroom settings are ordinarily more obviously changed by the socializing process than either the curriculum or the teacher.

The socialization of school children is achieved at the level of practice. The structure of daily classroom interaction produces agreed upon, socially defined ways of thinking and acting. Reacting to the demands of classroom life in a socially acceptable manner becomes second-nature to most children, and a common definition of the situation eventually is taken for granted by all classroom participants. At the moment when most aspects of the structure of classroom life are no longer recognized as choices from a range of alternatives, but are acknowledged to be both appropriate and inevitable consequences of school life, the socialization of the children can be said to be more or less complete.

The process of socialization and the imposition of adult frames of reference are never absolutely complete, however. There are always difficult moments and pockets of resistance which teachers must defuse daily. In addition, there are contradictions in the socializing messages themselves which teachers must resolve. It is at these moments of conflict and/or contradiction that both the dominant messages and the manner in which they are transmitted are most obviously revealed.

Socialization in Kindergarten

Kindergarten classrooms are particularly suitable settings for the study of the socialization of school children. Kindergarten attendance is widely accepted as a valuable experience for children because it is presumed to establish the foundation for the years of schooling that follow. Kindergarten training appears to exert its most powerful and lasting influence on the attitudes and behavior of the children by acclimating them to the classroom environment. The socializing aspects of the curriculum, then, are likely to be highlighted in kindergarten classrooms.

The teachers' messages about appropriate classroom behavior are particularly evident in kindergarten classrooms for two other reasons. First, the kindergartners are new participants in public school settings. Since this is their first experience, it is necessary for the teacher to make norms of appropriate classroom behavior explicit, using clear, straight-forward, repeated instructions. The fact that the kindergartners are new participants is compounded by the fact that they are five years old. This combination of
factors influences teachers to be direct and consistent in sharing messages concerning appropriate classroom conduct and the proper relationship of the children to the teacher, to the curriculum, and to each other.

Second, kindergarten classrooms are particularly relevant settings for studying the negotiation of a common definition of a social situation because they ordinarily include a potent contradiction in the curriculum. Schools are organized as workplaces, and the children are evaluated largely in terms of how well they do their work. Behaviors such as achievement for the sake of extrinsic reward, structures such as evaluative standards external to the project and the participants, and values such as punctuality and perseverance, are indicative of the work orientation of classroom settings.

Kindergartens, however, traditionally include an element of play activity in the daily schedule, and although the concept of appropriate and educative kindergarten play has changed since the first kindergartens, the belief that play is both appropriate and educative for kindergarten children has not. One clear indication that play is both acceptable and expected in kindergarten is the presence of play equipment and materials. Most kindergartens include doll corners, easels and paints, blocks, crayons, and other materials traditionally associated with children's play. Further, most early childhood educators believe that play is a developmental characteristic particularly relevant to the ages of the children in the room. Play is said to promote physical coordination, to help develop intellectual skills, and to provide experiences in social living.

Since play is considered necessary for the healthy growth and development of young children, it cannot be excluded from kindergarten classrooms. However, the presence of play materials in the classroom and of play activities in the daily schedule contradict the work orientation of the school setting. Teachers are left to reconcile this contradiction at the level of daily practice.

Incorporating Play into the Kindergarten Curriculum

There are a variety of strategies teachers use to resolve the conflicts engendered by the resistance of the children to the teacher's control and by contradictions within the curriculum. In the case of the contradictions play in the kindergarten curriculum exposes, the manner in which teachers must respond is dictated by the dilemma itself.

Play cannot be banned from the classroom, and expressions of play cannot be suppressed entirely because they are considered to be a necessary activity of young children. Consequently teachers must find a way to incorporate play into the curriculum without damaging the work atmosphere of the classroom. While incorporating play into the curriculum may be difficult for classroom teachers, it also appears to hold the potential for precipitating important changes in the children's perspectives concerning school behavior.
Teachers reconstitute significant aspects and artifacts of the children’s lives when they fit play into the school environment. While play activity may seem trivial and play materials insignificant to many adults, play experiences are an important part of most children’s daily lives. Children involved in play expect to organize their time, to select their materials and mates, and to direct their own activities. Play, from the children’s point of view, permits them a wide range of opportunities and options for self-expression.

Play objects in the classroom provide the teacher with an opportunity to shape the socialization of kindergartners using materials with which they are already familiar. Paul Willis points out that “Big change is no change unless it changes the small: our commonsense beings, commonplace habits, and accepted use of everyday objects.” Profound changes in the children’s understanding of their social reality may be accomplished best, then, by using the common objects of their play.

Incorporating play in the curriculum poses difficulties as well as opportunities. Play activities must not erode the capacity of the teacher to organize the classroom as a workplace; they must not permit children an autonomy which threatens the teacher’s authority and control, and they must not eclipse the accomplishing of assigned academic tasks. Yet, they must be recognizable as play to the children. Play in the curriculum must simultaneously serve the needs of the children and the goals of the teacher.

The Study

The study of the contradiction that play activities introduce into the kindergarten curriculum necessarily involves the observation of classroom interaction and the interviewing of classroom participants. The study reported here included the observation and interviewing of kindergarten children in four public school classrooms. The schools participating in the study served predominantly white middle-class families in New England and the Midwest. The classroom teachers were all women who had been teaching for many years and who were willing to permit a researcher to observe classroom activity and to interview children.

Each classroom contained materials traditionally associated with children’s play, such as dolls, trucks, blocks, paints, crayons, beads, and stuffed animals. One of the classrooms also had a tricycle and a wagon. These materials were invitingly displayed on the floor or on shelves, easels, and tables conveniently arranged for easy access by the children. While teachers occasionally taught one or more children the rules of a game or how to use some of the small manipulatives, for the most part the children seemed familiar with the materials in the classroom, eager to use them and knowledgeable in their use.

Data collection concerning the children’s use of play materials was restricted to those times when the children defined their interaction with
materials as play. Activities traditionally thought of as play (coloring a picture, for example) were not always labeled play by the children. Kindergartners labeled their classroom activities play only when they perceived themselves to be free of adult direction. When they were permitted to choose their activities, their materials, and their mates, the children said they were playing. When they were doing an assigned task, the children said they were working. The activities, themselves, did not carry an intrinsic meaning; the meaning was dependent on the social context. A series of entire morning sessions was observed in each classroom, and a narrative of events was recorded. The running record of children’s play activities was interrupted periodically by timed activity-samples. The teachers selected a boy and a girl who were poorly socialized, a boy and a girl who were adequately socialized, and a boy and a girl who were well-socialized. These children were observed in random order for a period of two minutes each and their activities were noted in detail. In addition, selected play materials were observed in random order, and their use was noted in detail for a period of two minutes each. These materials included the blocks, the paints and easels, the doll corner, the small manipulatives, the puzzles, the crayons and other drawing implements and materials, the books, the dress-up clothing, and the games. The data for this study, then, include a series of detailed accounts of the activities of selected children and interactions with selected materials embedded in an ongoing narrative of kindergartners’ classroom play.

The researcher was a non-participant observer and remained seated at a desk apart from the classroom activity most of the time. After the first one or two days of observation, neither the children nor the teachers seemed distracted by the presence of the observer. Some children initially expressed interest in what the observer was writing, but were satisfied with the explanation that the researcher was learning about “what children do in kindergarten.” Their interest quickly waned and, for the most part, the observer was ignored.

The children selected for observation during the timed behavior samples were interviewed and asked to respond to a series of open-ended questions concerning their kindergarten experiences. Interviews were conducted in the hallway outside of each classroom; no child was reluctant to participate and most of them were eager and gregarious respondents.

Findings

The data indicate that play periods are filled with conversation, planning, laughter, and busy activity on the part of the children. Although a small number of children occasionally chose to watch others, most children actively involved themselves with both materials and other children during play. The materials were central to the children’s experiences and the data reveal that, in addition to their usefulness in promoting interaction among
children, enhancing emotional development and stimulating cognitive
growth, play materials in kindergarten classrooms also embody institutional
messages and facilitate socialization. Materials found in kindergarten
classrooms are not simply physical objects; they are part of the curriculum
and, as such, they embody curricular meanings unlikely to be replicated
outside an educational setting. Play materials in kindergarten appear to
serve a number of functions more directly related to their classroom context
than to their traditional uses.

First, during the opening weeks of school, the presence of play equipment
in the classroom seemed to ease the children's introduction to kindergarten.
While children were observed to be wary of other children they did not
know and were somewhat reticent in their initial interactions with the teacher,
most children, after realizing they were permitted to use the objects in the
room, interacted readily with the play equipment. Involvement with
familiar materials seemed to be an entry into a somewhat strange and cer-
tainly novel environment.

During the first week of school one child, for example, began her morn-
ing by doing two jigsaw puzzles and coloring a picture. The speed with
which she was able to complete the puzzles indicated a familiarity with
them, and her coloring permitted her to sit with other children while in-
volved in a familiar activity. Her initial contacts with other children
developed at the "coloring table" while all were making pictures. During the
second week of school, she began to vary her routine and to use play equip-
ment with which she seemed somewhat less familiar.

The display of play materials, then, served to entice the new kindergart-
ners into the classroom and the school program. The play equipment was an
element of continuity for the children and demonstrated to them the rela-
tionship of their past experiences in informal non-school settings to ex-
periences in classroom settings. The children believed that the teacher
selected the equipment in the room; this belief was strengthened in one
classroom when the teacher told the class that she had brought some of the
"toys" to school for them when her own children outgrew them. The
relevance and familiarity of the materials may have shown the children that
the teacher understood their abilities and interests.

Second, the organization of classroom materials was a means by which
the pattern of authority in the classroom was demonstrated to the children.
The apparent availability of play materials in the physical space was con-
tradicted by the limited access the children had to these objects during the
kindergarten session. The kindergartners were not permitted to handle the
play materials during periods of teacher-directed instruction. For example,
when the teacher read a story, showed a movie, led the class in singing,
presented academic material, or required the children to complete assigned
individual tasks such as worksheets, the children were not allowed to handle
the "toys."
The availability of materials varied considerably in the kindergartens observed, ranging from no access whatsoever on a given morning in one classroom to complete access one morning in another classroom. In every case, however, it was the teacher who controlled access to classroom materials; she determined both the time and duration of the children’s play in school. The children seemed to learn this fact of classroom life quickly and were often observed to ask the teacher for guidance in planning their activities appropriately with questions such as, “Can we play now?” “If I finish can I play?” and “When will it be time to play?” Interaction which would have been spontaneous under other circumstances was regulated in school. The teacher’s control of the flow of events in the classroom was displayed, in part, by her ability to limit or permit the use of play materials. Since the children were eager to use these things, the opportunities for the teacher to demonstrate her control were numerous, and the pattern of authority with regard to play materials was quickly apparent to the class.

Third, the flow of the children’s play activity became part of the rhythm of classroom life, and the class learned to want to play when it was time to play. In classrooms where access to play materials was limited, no child was observed to refuse to play or to linger over assigned tasks when play was allowed. In some cases, the children rushed to the equipment and quickly involved themselves in play activities. The knowledge that they would be required to stop playing at a time designated by the teacher seemed to foster an almost feverish involvement at the beginning of the play period. This burst of energy settled quickly into participation at a less frenetic level.

The timing of play activity in the classrooms also demonstrated some aspects of the relationship between work and play in school. Undirected involvement with play materials ordinarily followed formal lessons or periods of directed activity, and teachers often emphasized this with comments such as, “After we talk about our lesson, you will be able to use the toys,” and “Finish your worksheets and you’ll have time to play.” Use of play materials, then, was often a reward for work completed or periods of participation in assigned tasks. Thus the pattern of play in the classroom served to complement and sustain the pattern of work. Children learned to defer their activity with play materials until their school work was finished.

Fourth, the children learned that particular patterns of play were appropriate in the classroom. They learned, for example, to play quietly. One kindergarten teacher held a lengthy discussion with her class on the second day of the school year about the use of “indoor voices” in the classroom. All of the teachers reminded children who were talking or laughing loudly to lower their voices and commended children who were playing quietly. As the children adjusted the volume of their play to an acceptable level, the manner in which they expressed surprise, disappointment, anger, or exuberance became less an outgrowth of their activity and more a response to the circumstances of their play.
Just as the children learned to exercise some restraint in their reactions to their play experiences, they also learned restraint in their use of some of the play materials. The tricycle and the wagon, in particular, seemed to encourage exuberant play, and it required quite a bit of self control to use this equipment in a manner suitable to the classroom setting. The teachers in all four classrooms praised children for their appropriate and careful use of materials and for the speed with which they put materials away at the end of a period of play.

The children were also praised for involving themselves with play materials promptly at the beginning of a period of play activity. The teachers encouraged children to participate in some recognizable activity, and acceptable play almost always included the use of materials in an active manner. In one of the classrooms, watching other children at play was not considered to be a viable play activity, and children who were observing others were persuaded to become directly involved in some activity themselves. In all four of the classrooms, children who were wandering or daydreaming were encouraged “to find something to do.”

Fifth, play materials in the classrooms did not belong to any of the children and kindergartners learned to play with equipment they did not and could not possess. The children's use of play materials was transitory, and the opportunity to play with any particular piece of equipment could not be taken for granted. In other words, all of the children had an equal claim to the play materials and use of the materials required sharing with others, including others in a game or project, and taking turns. Occasionally one or more children asserted ownership of some piece of equipment, and the teachers used such incidents as opportunities to make the institutional quality of classroom materials explicit. Teachers encouraged children to share and, if necessary, the teachers insisted. Sometimes simply saying, “Joan wants to play too,” was sufficient; in other cases, teachers were more direct, making comments such as, “Boys, we need to give the girls a turn with the blocks,” and “It’s Kelly’s turn to use the windmill.” If the children resisted, teachers explained the need to share in terms of fairness and in terms of the common ownership of the equipment. Children were told, “These toys belong to everybody,” and therefore no one child could exclude other children from their use.

Finally, the play materials became the focus around which the children’s culture in the classroom emerged and developed. During formal lessons and work periods, conversation between the children was limited by the teacher. Ordinarily, the children were not allowed to talk to each other during periods of instruction; at times quiet talking was permitted while children completed tasks at their seats. During play, however, the children sought each other out and established small groups of friends within the larger classroom population. There were far more spontaneous peer contacts during play than during work; there was more conversation, and more pur-
poseful selection of specific others to talk to. Teachers rarely imposed on spontaneously emerging groups during play unless activity became too noisy or too boisterous, or unless an excluded child complained.

The children learned to do their socializing during periods of play and to combine interaction with play materials and interaction with peers. Though some children chose to play alone on any given morning, most children balanced the independent activity required during work with peer interaction during play. Play times in the classrooms observed were times of talking, laughing, sharing ideas and secrets, making plans, and pursuing joint ventures. While work periods did not necessarily exclude such activity, it was more prevalent and spontaneous during play.

**Discussion**

Kindergarten marks a transition in the child's play life.º The natural, spontaneous play of young children is not the play one sees in the classroom. Such idealized notions of play are inadequate to describe play in school where the context adds a dimension of meaning which cannot be overlooked. The presence of play materials in their classrooms helped kindergartners become acclimated to school life and to recognize and join in socially appropriate behavior. The materials were familiar to the children, as shown by the fact that, in response to the question, “What is the same about kindergarten and home?” children referred to various pieces of classroom equipment and “toys.” The very familiarity of the materials themselves may have heightened the children's awareness of the novelty of the context in which they now encountered these things.

When asked to describe differences between kindergarten and home, most children responded by naming one or more restrictions on their use of play materials. Their answers included responses such as, “The teacher is in charge of the toys here.” “You have to talk low.” “You can't run in school.” and “You have to share.” Many children also pointed out that there were more children to play with in school. One girl responded, for example, “There's lots more kids here!”

The pattern of play in the classroom initiated children into the appropriate relationship between work and play in school. Children learned to ignore tempting materials, and to delay interaction with each other until the teacher allowed them to play. They learned to interact with restraint and to shift gears quickly from play responses to work responses. The children seemed able to accept the limitations on their activity during periods of work in part because they realized that a period of play would follow. Teachers were observed to encourage children to finish their work and to do their best work by pointing out that a play period would be available to those who concluded assigned tasks successfully.

Teachers also asked children to defer their interactions with each other until their work was completed. Children learned to place their expectation
for social interaction and self-directed activity on periods of play, and play materials thus provided not only satisfaction, but enjoyment and diversion as well. Periods of play served to relax and refresh the children so that they returned to their work with renewed energy and concentration.

This use of play in kindergarten is analogous to the role of play in the lives of adults. Adult play is also constrained by the demands of work, and is expected to provide a change of pace from work responsibilities. Adult play ordinarily differs from adult work in the activities, the materials, the interactions with others, and the location in both space and time. Kindergarten play is not yet entirely outside the sphere of work in that the play materials appear in the classroom itself and periods of play are permitted at designated times. Still, kindergarten marks the beginning of the transition from child to adult rhythms of play.

Classroom play provides the opportunity for children to practice recognizing and accepting the relationship between work and play found in many adult occupations. The distinction between adult work and play is anticipated by the classroom distinction between academic experiences and free time. As the timing and duration of play become determined by the demands of work, children learn both to accept constraints on their activities during periods of work and to accept constraints on their play imposed largely in service to the work orientation and requirements of the institution.

This distortion of the idealized notion of children's play can be explained by the school's need to function smoothly, and by the desire to prepare children to deal effectively with issues of work and play in settings devoted largely to work. Children's play is not banned entirely from kindergarten, but it is shaped by the classroom context. Play becomes instrumental, which is in contradication to its ordinary characteristics, as the child's natural play forms begin to be shaped into adult play forms in kindergarten.

Conclusion

The meaning of play materials in the classroom is part of the social definition of the situation. On the first day of kindergarten, the teacher has a more organized definition of the social situation than do the children. Further, the teacher is eager to influence the emerging understandings of the children and is in a position of considerable power to do so. Consequently, ordinarily, the teacher's definition prevails.

In the case of play materials familiar to the children, however, the children have well-developed individual attitudes and responses. There is, therefore, more negotiation required to reach a definition of play materials acceptable to all classroom participants. Within limits set by the traditional meanings of the play materials and each child's individual response to these materials, the teacher works to modify the children's orientation toward play materials in kindergarten. This is difficult because of the children's
past experiences with these objects, and significant because it highlights important differences between the kindergarten context and other contexts in which children encounter these materials.

Even as teachers use play materials to help the children accept the constraints of the kindergarten classroom, children also use play materials to achieve their own goals, and to develop personal meanings within the institutional context. The play materials hold, for the children, the richest possibilities for self-expression, and periods of play offer the best opportunities. Children exploit these opportunities even as they accept the institutional messages the play materials have come to include.

References


2. The socializing functions of the kindergarten curriculum have been widely acknowledged for a considerable period of time. As Harry Gracey, for example, notes, “Kindergarten is the place in which children begin to learn the pupil role. At the core of this learning is a set of classroom routines which the teacher introduces and then trains the children to follow.” Gracey, Harry, *Curriculum or craftsmanship*, Chicago: The University of Chicago Press, 1972, p. 163.


4. It is my responsibility to protect the anonymity of the children and teachers I observed and interviewed that prevents my thanking them by name. I am extremely grateful to all of them for their interest and encouragement throughout this research effort.


6. A study of four-year-olds in the United States found that they see themselves as people who should play. See: Goodman, Mary Ellen, *The Culture of Childhood*, New York: Teachers College Press, 1970, p. 72. The kindergartners interviewed in this study called themselves “workers.”

Sexuality in the Classrooms of Teachers with Various Sex Role Orientations

Elizabeth L. Peterson
Toronto, Ontario

Emily M. Nett
University of Manitoba

Acknowledgements

The first author would like to express her appreciation to Dr. Terry Morrison, Dean of Continuing Education, University of Manitoba, for the idea of ethnography. The second author is grateful to Rita Gunn, graduate student in the Department of Sociology, University of Manitoba, for her assistance in compiling the references on sexuality in privileged relationships other than teaching.

Introduction

The student-teacher relationship, like that of other professionals and their patients, clients, or parishioners, is ideally affectively neutral. Moreover, the societal constructs of childhood innocence and maternal asexuality infuse the traditionally feminine role of teacher with an element of the sacred trust.

Are teachers actually sexually neutral in their classrooms? In this paper we shall report on the unanticipated findings from an ethnographic study of sexuality as an apparently unconscious element in female teachers' classroom interactions, discuss relationships between teacher expressions of their own sexuality and sex role ideology subscribed to, and consider some implications for social studies of sexuality in learning situations.
Despite the fact that interaction in classrooms is a frequent topic of study, understanding the processes involved in teacher-pupil relationships is still an imposing problem (Jackson, 1968; Smith and Kleine, 1969; Dreeben, 1970; Stubbs and Delamont, 1978; and Martin, 1975). Furthermore, the existing research on sex in privileged situations is limited, since most expressions of sexuality are not overt and investigations have tended to stress exploitative sex contact and sex practices (see Stone, 1976; Taylor and Wagner, 1976; Greenberg, 1979; Pope, Levenson and Schover, 1979; and Gentry, 1980, on sex between therapists and patients, and see Ginsberg and Koreski, 1977; Farely, 1978; and McKinnon, 1979, on sexual harassment of women by employers).

The focus of this paper is not on "deviant" sexual behaviors, but on the normal presentation of the sexual self in a relationship where erotic cues are thought of as non-existent or "unthinkable" (Parsons, 1951). Awareness of the presence of sexual feelings in this type of situation has been acknowledged as both normal and controllable in the case of the counseling pastor (Stroup and Wood, 1974), but no reference to the potential in the classroom for the eruption of teacher sexuality has been found in the literature. For instance, an ERIC search using the descriptors "student-teacher relationship" and "sexuality" in combination yielded only three accession numbers of journals (Momentum, 1974; Educational Horizons, 1976; and Times Educational Supplement, 1978) and a single document (ERIC, 1966). Furthermore, of the four, two dealt exclusively with students' values and attitudes, one with helping students come to terms with matters such as sex, and the final one with situations involving overt sexual expression by children in school. Not one of them even considered overt behavior in the classroom let alone sexual interaction between teachers and students.

The purpose of this paper is to present the findings of a preliminary study of observed differences in sexuality exhibited in the classrooms of teachers with various sex-role orientations, and to explore further in view of the results the several aspects of sexuality as an interaction variable.

**Observed Sexuality in Teacher Interaction: An Ethnographic Study**

Based on the idea that the "hidden curriculum" — the repetitious and continual incidental learning which students in school experience in contact with one another, with the teachers, with the rituals and rules of the school and the subtle meanings in textbooks (the unwritten agenda of activities) — often has the greatest impact on the learning which takes place within the "culture" which the classroom represents, an ethnographic study was undertaken. The purpose was to investigate the role played by the teacher's sex-role ideology in classroom interaction.

Ethnography, or naturalistic observation, utilizes "trained observers to gather qualitative data by direct observation of human activity in an ongoing naturalistic fashion . . ." (Tikunoff, 1975: vi-vii). The aim of
ethnographic research is to develop hypotheses rather than test them. An observer usually enters a field situation with some tentative assumptions to organize what is viewed, but hypotheses to be verified or disproved are formulated only after the data collected from observation have been analyzed. That procedure was followed in this study.

Selecting the teachers. Using an attitude scale, a "dilemmas" test, and a follow-up interview (for details of the methods see Peterson, 1979), three female teachers from one school were found with different sex-role ideologies. Only women were chosen in order to control one variable. Furthermore, we were concerned with the idea that female high school teachers serve as role models for adolescent girls, and we wanted to investigate behavior being displayed by women in that role. One of the three teachers held rather traditional perceptual orientations toward females and males, masculinity and femininity. She took a dichotomous view of gender roles. For example, individuals could not be passive and agentic; they had to be one or the other. She tended to see roles as clearcut and appropriate to the inherent differences she perceived in the sexes. At the other extreme, another teacher subscribed to a more androgynous set of beliefs and ideas about role expectations, behavioral orientations and personality attributes. She subscribed to the flexibility of gender roles, believing persons capable of behaving in integrative feminine and masculine ways, such as being assertive and yielding, independent and dependent, expressive and instrumental. For her, roles were socially ascribed and thus more or less arbitrary. The third teacher could be described as neither "traditional" nor "androgynous"; she was classified as "mixed" since she appeared to hold some traditional beliefs about differentiated gender roles, but was particularly aware of and accepting of some of the financial benefits of more opportunities for women outside the home. The teachers were designated Teacher T (traditional), Teacher A (androgynous), and Teacher M (mixed).

Data collection and analysis. The classroom observation was done over a period of nine weeks in 1977 in a Canadian secondary school. The senior author sat in for approximately 75 hours on foreign language and typing lessons with the three teachers, mainly in their grade 11 classes. During the period of observation she was simply concerned to write down as much as she could of what was happening, concentrating on noting ways in which the teacher's sex role orientation may have affected the interaction. In order to provide as much context as possible for the learning mechanisms involved, her field notes included as many details of pupils' responses and other classroom events as she could write down in notes during class and elaborate upon immediately after the observation period. At the end of each day's observations of all three classes, comparison was made in order to monitor patterns of interaction and to sensitize the field worker for further observation.
Upon completion of the observation period, the observer wrote a description based on her field notes of the interaction in the behaviors of each of the teachers. These were then submitted to an educator and a sociologist specializing in women's studies who independently arrived at categories of analysis which seemed to them to represent the mechanisms of cultural learning in the classes. The concepts of kinesis, power, and sexuality were derived from this procedure. The categories emerged from the observations themselves since no preconceptions were hypothesized about the mechanisms by which teachers dealt with their various sex-role orientations in their classrooms. Using the three categories, the three descriptions were then compared and contrasted.

Findings.

The ethnography produced some surprising results. First, the obviously sexual content of some classroom interactions was striking and unanticipated. Second, the way in which the three teachers differently expressed their own sexuality in daily interactions with students had not been predicted.

Teacher T. The image this teacher presented was that of a woman who was extremely self-confident in her ability to perform the female role of teacher, one who displayed herself as being feminine in appearance without being overtly sexual, and yet one who behaved somewhat coquettishly and beguilingly. The distance and aloofness of Teacher T's formally-correct clothing emphasized her feminine gestures and body language, presenting an image to her students that was quite in line with the concept of a very traditional female person—one who knows she is being regarded by others (as an object) and who is dutifully meeting their expectations (to be pretty, but not sexy).

Teacher A. Stereotypically unlady-like, Teacher A conducted her classes in an informal, relaxed and yet controlled way. She dressed casually and in ways which neutralized her sexuality. Her behavior, too, was situationally appropriate.

Teacher M. It was obvious that Teacher M was ambivalent in the way she expressed her sexuality. On the one hand, she seemed totally unaware that she was conveying any messages about sexuality in her classroom; on the other hand, she, of the three teachers, acted in the most sexually provocative way. Unlike Teacher T whose sexual expression was peripheral and a subtle influence, or Teacher A whose sexuality was taken for granted, Teacher M had a much more physical presence and projected a rather alluring image. Particularly noticeable were Teacher M's clothes, and surely her students learned that to be female meant emphasizing one's best features in garments which were eye-catching and often sensual. Teacher M unconsciously reacted to her male students as tradition demands women react
to males—deferring to their demands, engaging in banter with them, and maintaining less of a personal space distance between herself and them than she did with female students. Her movements were typically feminine, not forceful or assertive. She rarely wrote on the chalkboard or gestured, but she moved about the room more than either Teacher T or A.

**Classroom Interaction**

In order to interpret the observed differences in the way the three teachers expressed themselves as sexual persons to their students, it will be necessary to examine other aspects of their classroom behavior. In this section the atmosphere each created and the procedure followed by each will be summarized.

*Teacher T.* The teacher with the traditional or conventional sex-role orientation ran an orderly classroom. She had the knowledge; the students were there to learn. It was her responsibility to teach them by means of a serious approach to cognitive learning with little time for the affective domain. Silence, order, and authority were the most obvious features of the classroom.

In establishing order and authority, time and space were important. Teacher T began to settle her class and do the routine tasks, such as attendance, prior to the start of class and as the students were entering the room so that she began to speak the moment the bell sounded. All of the lessons were teacher-directed; she asked questions, the students raised their hands, and she called on them. Teacher T also exercised strict control over space and movement within her room. Her territory, carefully demarcated from that of students and including the doorway which gave entrance and exit, was "center stage" or somewhere in front of her pupils, with distance between herself and students always maintained. Movements in the students' space was almost non-existent; in the teacher's space, almost perpetual.

With regard to appropriate roles for the sexes, there were no overt references made, but, on the other hand, there was a feeling, understood by all, that females and males acted in certain typical but very different ways and that the teacher's sex was an underlying factor in her teaching style and attitude toward education. Further, there were several incidences observed which seemed to contain differences based on gender. For example, in a class containing nine males and sixteen females, Teacher T continually forgot the names of females but never males. Furthermore, Teacher T singled out males to joke with more than females, injecting a different quality into these jocular interactions. The difference was that with male students there was verbal by-play between teacher and student, whereas with the females, the teacher carried on a monologue.

*Teacher A.* The teacher who subscribed to an androgynous sex-role ideology was personally involved with her students. Fairness and respect for
them were top priorities. She seemed able to pay adherence to the subject-matter at the same time as she continued to be aware of the students' and her own needs in the close environment of the classroom. In fact, balance characterized the interaction between teacher and students.

The atmosphere in the class was serious and at the same time friendly. The students knew and accepted the rules; the teacher had very good rapport with them. Her most effective disciplinary technique was silence. When she used it, the class stopped and the misbehaving person desisted. Her disapproval of mischievous behavior also was indicated effectively by a frown, which she managed in such a way that the lesson was not disrupted for those not involved in the misbehavior.

Space and movement in Teacher A's class was well-defined yet flexible. Her territory was in front of that of the students. Hers also included an unoccupied student desk at the front and center from which she often taught by sitting on top of it. From time to time she traded spaces with a designated student. In this exchange, however, there was no role reversal; students did not assume authority in the teacher's space and the teacher did not lose control while in theirs.

The result of Teacher A's fine sense of the temper of time and place was a calm environment where interactions were plentiful between students. Gender seemed irrelevant in Teacher A's behavior.

Teacher M. The teacher who held a mixed sex-role ideology allowed her students freer rein than either of the other two teachers, and her classroom reflected this in the auditory and spatial arrangements observed, as well as in the general climate of the room. The tone of Teacher M's approach was informal, but her kindness was often mistaken for leniency. All 27 students talked freely throughout the hour, but it was five of the 22 male students who sought constant attention either from the teacher or their classmates.

Teacher M had attempted to establish control by designating the front of the room with a raised platform and long table ("my throne") as her area, and the six rows of student desks as student territory, but she was unable to maintain the exclusiveness of each. There was indeed much movement in the room, with the teacher frequently walking up and down the rows answering the flow of questions and the students going from desk to desk to speak to each other.

Sex Role Orientation and Sexuality

The major hypothesis formulated from the results of the ethnography, as described in the section above, is that a link exists between a teacher's sex role ideology and the expression of her sexuality in the classroom.

Teacher T projected an image of feminine competence in the classroom, aware of her appearance without overtly stressing the sexual, and yet one who, within the boundaries of respectability, might captivate an audience, including males. On the other hand, Teacher A presented herself as a com-
petent teacher who happened to be a woman. Her style of dress, her mannerisms, her relationship to the males in her classroom seemed to take second place to the display of respect and appreciation for her students as learning individuals. Teacher M, with the mixed sex role ideology, was the seductive female in appearance and action. She appeared confused about her feminine roles; that is, she alternated between the female as sex object and temptress and the female school teacher as erotically-neutral intermediary between the legitimate male authority of the principal and the relatively powerless young persons in her classroom.

The value of ethnography is its ability to provide insights which other research methods cannot; its limitation is that it is not capable of testing the hypotheses generated by the data. Relationships have not been established merely because this traditional teacher was observed expressing herself as traditionally aware of her femininity while at the same time in complete control of it, and the androgynous one as un stereotypically personal, and the mixed one as seductive. Relationships between the two variables, however, are strongly suggested. However, the complexity of sexual expression only begins to be revealed by this type of observation.

**Interpretation**

Having established that teacher expression of sexuality in the classroom was observable and variable, and moreover that in the three cases observed it occurred in the context of differing sex role orientations and teacher styles, we turn now to analyzing some of the components in the normal expression of sexuality.

We start with a tentative definition of sexual expression: the behavior of the person who has learned the pattern of culturally comprehensive cues which are intended to provoke or elicit erotic interest or arousal, and who has learned under what conditions the cues are appropriate or inappropriate. The cues include dress and adornment, body use and display, voice and tone, as well as language and action.

Being "properly feminine" in the Western tradition has meant subduing, if not denying, one's sexuality. This mode of non-erotic feminine response conjures up the metaphor of "the Lady". Demonstrably female, she does not actively involve herself in consciously arousing sensuality or in satisfying her own desires should she have them. In fact, the lady, although capable of creating desire merely by being female, must never actually consciously provoke or gratify. She must remain on her pedestal and never involve herself emotionally with others. On the other hand, other images of feminine sexuality have abounded, depicted by the lustier "earth mother" or the powerful "witch" figures (O'Faolín and Martines, 1973). The sexual duplicity perceived in feminine nature has been variously dealt with in different cultural representations, but in recent times the solution has been to deny sex appeal to some kinds of women, such as wives, mothers, and
teachers, and to assign it to special categories of other women, such as pubertal girls, public entertainers, and prostitutes. The "proper" or respectable woman downplays, at least in the public sphere, her sexual nature, unless she is "manhunting" and then it must be used very subtly. The other exception is that even the "good woman" should present herself as a pleasing object for others to admire. The line between provocative and pleasing is indeed a fine one.

The sexual style of an androgyne is difficult to conceptualize. By definition androgynous expression originates in the situation rather than in femininity or masculinity. The model of androgyne is one of accepting sexuality in one's self and expressing it toward an appropriate actor or object on an appropriate "stage" in the drama which life represents, rather than denying or affirming one's sexual nature because of sex or gender.

Teacher T’s and Teacher A’s sexual styles make sense in view of the sex role ideologies to which they subscribed; that is, the traditional teacher exhibited the model of the "lady" and the androgynous one, the "androgyne". There was no basis, however, upon which the sexual expression of Teacher M could have been predicted. Why, after all, would a mixed sex role ideology be associated with such provocative behavior as was observed? The answer to this question lies in the way in which sexual expression appears as a function of personality integration, including such elements as teaching philosophy and style, and sex role ideology. The observable, behavioral link between the three components—classroom style, sex role orientation, and sexual expression—can be found in the teachers’ use of power. This is interesting in view of the prominent part played by the concept of power in the contemporary paradigms used to analyze relations between the sexes.

It is difficult to escape the conclusion of extreme control in the case of Teacher T. She played the role of teacher to the hilt. As the teacher, she demanded the complete attention and respect due her. To use Weber's term for her power it was legitimate, since it was derived from the authority of the position. (See Gerth and Mills, 1958:295-300). At first this aspect of Teacher T is confusing; exercise of authority seems more in keeping with the masculine stereotype. However, the parallels between her use of power as a control technique and the situation of the mother in the patriarchal family clarify the matter. In both roles the female acts as the agent of the male, carrying out his mandate in the secure knowledge that behind her actions lies his ultimate authority. Her authoritarian philosophy of education enabled her to play the traditional role of the teacher without conflict with the feminine role of the "lady".

Teacher A, whose philosophy of education was reflected in her awareness of the effects of her decisions on interaction with the students, was more flexible in her use of power than Teacher T. However, she was far from abandoning formal authority. Consequently, she sought and received the
respect of her students even if she did not require the degree of control over them that Teacher T did over hers.

Teacher M, torn between uncertainty as to the appropriate feminine role and the need to maintain discipline in the classroom, had neither respect nor authority to fall back upon in order to maintain the atmosphere required for the tasks at hand. Affected by the ideas on change in sex roles and education, she was unable to incorporate them effectively into a teaching or personal style of her own. Influence, the type of informal power in which the personal qualities of the teacher count rather than her position, might have been an alternative to an authoritarian or strict rule. However, Teacher M's sex stance interfered with tactics which could have resulted in her making friends with the students and winning their loyalty and admiration. In fact, she did just the opposite; she lost their esteem since her feminine search for the approval of the "other" prevented her from carrying through on her disciplinary threats. In all, her mixed sex-role ideology did not provide her with a definite sense of a unified, nor compartmentalized, self. The ambivalence of that ideology combined with her inability to exercise the legitimate power inherent in the teacher role. The potency of her sexuality, being totally inappropriate, also failed her.

In sum, what was discovered during the observation was that these three teachers' use of power was linked closely with the way sexuality was denied, accepted, or flaunted. Whether she controlled her class by stern admonition, coy rejoinders, or fair play was associated with sexuality as a part of the teacher's total personality, most notably in her sex-role orientation.

Conclusion

At the outset of this paper we observed that sexual expression may be an important component in student-teacher interaction, viewed as a special genre of the privileged relationship. Our findings suggest that erotic expression appears to be variable, at least for female teachers.

We recommend caution, however, in the application of results to social studies teaching or programs. Not only does the hypothesis derived from the ethnography require testing on larger numbers of teachers, both with further observation and survey methods, but it is especially important to know whether students, female and male, observe the differential sexual behavior of teachers with various sex role ideologies. We have examined only the role models, not their effects on students. Additionally, conceptualizing control, or use of power, as a factor in the normal expression of feminine sexuality and relating it to sex-role ideologies raises questions about how power as a social element might enter into the normal expression of masculine sexuality. Masculine sexuality appears to be much more invariably connected to gender identity than feminine sexuality, according to Person (1980:624). At the same time, the dualism found in the stereotypical notions of femininity is absent from those of masculinity. It is necessary,
therefore, to determine whether the general hypothesis of a relationship be-
tween sexual expression and sex-role orientation can be substantiated for
male teachers as well. Theoretically, the case can be made that "traditional",
"androgynous", and "mixed" male teachers would behave very differently
from their female counterparts since there is only one acceptable male role
and direct power is legitimate for it. This is an important issue since there is
evidence that sex of teacher has implications for learning and gender role
development (see Gordon, 1974, and Deem, 1978 for surveys of the lit-
terature).

Even so, social studies teachers in high schools, and all educators at all
levels of education, may want to consider some of the questions raised by
the research concerning the influences of social structures on our percep-
tions of how social control works. Why have we been so oblivious to the
potential for erotic display and arousal in the classroom? What do young
women (and men) learn when they perceive adult women in a position of
potential direct power subverting their own authority by following one or
both of the sexual scripts of the gender ideology? What can social studies
teachers do to undo the sexist lessons in the "hidden agenda"; i.e., either
that a woman's biology (sex) renders her incapable of directly exercising
power, or that her indirect exercise of it limits her authenticity or emotional
involvement with the young people in her charge?

It is to be hoped that through this study and further research on the topic,
including rigorous tests of the hypothesis derived from the ethnography, in-
sight can be gained into the relationship between a teacher's sex-role
ideology and his or her classroom interactions. The concept of the use of
power is one which may provide social studies teachers with a tool for better
understanding their own and others' classroom interactions and their rela-
tions with the mostly male administrators in secondary schools. Finally, if
social studies education is to be viewed in the broadest sense of social learn-
ing, then the subject of expression of sexuality in the classroom by teachers
is one of which educators must be made more aware.

References

Butler, Sharon and Seymour L. Zelen
1977 Sexual intimacies between therapists and patients. Psychotherapy: Theory,
Research and Practice 14:139–145.

Davidson, Virginia
1977 Psychiatry's problem with no name: therapist-patient sex. American Journal
of Psychoanalysis 37:43–50.

Deem, Rosemary

Dreeben, R.
1970 The Nature of Teaching: Schools and the Work of Teachers. Glenview, Il-
linois: Scott, Foresman.
Educational Horizons

ERIC
1966 ED 073383. Study at SUNY Buffalo giving information about students' backgrounds, values and experiences, etc., including family relationships, peer relationships, heterosexual relationships.

Farely, Lin

Gentry, W. C.

Gerth, H. H. and C. W. Mills (eds.)

Ginsberg, Gilbert and Jean Galloway Koreski

Gordon, Deborah E.
1974 The Psychological Implications of Teacher Sex on Learning and Sex Role Development at the Elementary School Level. M.A. Thesis, University of Toronto.

Greenberg, Joel
1979 Psychology students and the couch (sexual contact between psychologists and patients). Science News (Sept. 22) 116-201.

Holroyd, Jean C. and Annette M. Brodsky

Jackson, P. W.

Kaplan, Alexandra G.

McKinnon, Catherine A.

Martin, Wilfred B.

Momentum

O'Faolin, Julia and Lauro Martines (eds.)
Parsons, Talcott

Person, Ethel Spector

Peterson, Elizabeth L.

Pope, Kenneth S., Hanna Levenson, and Leslie R. Schover

Smith, L. M. and P. F. Kleine

Stone, Michael H.

Stroup, Herbert W., Jr., and Norma Schweitzer Wood

Stubbs, Michael and Sara Delamont (eds.)

Taylor, B. J. and N. M. Wagner
1976 Sex between therapists and clients: review and analysis. Professional Psychology 7:593–600.

Tikunoff, William J., David C. Berliner and Roy Crist

Times Education Supplement
Social Studies Project Evaluation: Case Study and Recommendations

John Napier
Mary Hepburn
University of Georgia

Introduction

Systematic evaluation of social studies development projects is of vital concern not only to educators and educational researchers but to public officials and taxpayers seeking evidence of the effectiveness of programs developed with public funds. In social studies education there have been a variety of projects since the sixties which have left us a legacy of curriculum materials and recommended teaching strategies. They have not, however, provided us with models for project evaluation.

In fact, we find very few project evaluation reports in the published social studies literature. As far as we can tell, few projects have systematically evaluated the development process or student learning outcomes in varied school settings over a period of time. Many projects have depended on impressionistic review by teachers or by the developers themselves. The net result is that there are very few cases and very little analysis in the limited evaluation literature on social studies projects. Consequently, when we began our work as members of an evaluation team for a large social studies improvement project in 1978, we became acutely aware of the absence of social studies project evaluation models. We had to look to the general evaluation literature for guidelines from which to synthesize an appropriate design.

The purpose of this paper is to report the evaluation design which we used
to determine project effectiveness and to examine the evaluation process as a case in point in an analysis of ideals and practicalities of project evaluation. Based on four years of experience we delineate what we consider to be fundamental or minimum requirements for project evaluation. And, beyond offering a usable model, we address the problem of reconciling the ideals of evaluation to the realities of the school system and its students, teachers, and administrators. Finally, we hope our reflections and recommendations will be useful to other social studies educators working with curriculum projects.

**The Project**

The Improving Citizenship Education Project is a curriculum improvement project of the Fulton County (Georgia) School System funded by the Georgia Department of Education under Title IV-C of the federal Elementary and Secondary Education Act. Its first objective was development and testing of an improvement model. Now the project is disseminating that model. The Fulton County school district is one of the largest and most populous in Georgia, spanning urban, suburban, and small town communities in its 420 square miles. Approximately 35,000 students attend its 60 elementary and middle schools and 19 high schools.

The purpose of the project in the first three years was to produce a workable model for systematically infusing political-citizenship education into the existing elementary and secondary social studies curriculum. The idea was not to throw out the existing curriculum or to develop a completely new set of instructional materials. Instead, the plan was to use “staff development” (teacher education), administrator cooperation, and teacher and student resource materials to bring about improved student learning. The project was funded to design, implement, field test, and evaluate the citizenship-social studies program over a three-year period. If the model proved effective, it would be disseminated statewide.

Nine student educational goals guided the development, implementation, and evaluation of the program (See Table 1). Utilizing results of national and state assessments of student political knowledge (National Assessment of Educational Progress, 1976; Atlanta Assessment Project, 1977), the staff and a teacher team had drawn up the goals in the first months of the project. From these goals eight cognitive content areas were defined by teachers and staff (See Table 2).

The project staff was composed of school district and university personnel with experience and/or interest in citizenship/political education. The evaluation team was composed entirely of university personnel. The social studies supervisor of the school system served as the project director. The development plan was initially a three-year and later a four-year plan as follows:

1) The first year, 1977–78, was for planning the improvement process.

42
<table>
<thead>
<tr>
<th>Table 1: Goals of the Improving Citizenship Education Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Understanding and acceptance of the responsibilities of U.S. citizenship;</td>
</tr>
<tr>
<td>2. Understanding of the structure and operations of local, state, and national governments;</td>
</tr>
<tr>
<td>3. Understanding of the roles of individuals in the political decision-making processes and developing the skills for participating;</td>
</tr>
<tr>
<td>4. Understanding of the principles of individual rights and learning to make choices in the context of concern for the society as a whole;</td>
</tr>
<tr>
<td>5. Understanding of the U.S. legal system and a commitment to rule of law;</td>
</tr>
<tr>
<td>6. Knowledge of current public issues and skills for evaluating alternative choices in regard to these issues;</td>
</tr>
<tr>
<td>7. Understanding of the interdependence of the global community and the political processes in operation internationally;</td>
</tr>
<tr>
<td>8. Understanding of the means to participate in school, local, state, and national political processes;</td>
</tr>
<tr>
<td>9. Understanding of the need for government and community services and the ways to secure, utilize and contribute to these services.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Content Areas of the Improving Citizenship Education Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. National Government—federalism, bicameral lawmaking, executive power and roles, judicial powers, separation of powers, checks and balances, policy-making processes, interest groups, lobbies, electoral college, two-party system, bureaucracy, national revenues and expenditures, U.S. citizenship.</td>
</tr>
<tr>
<td>2. State and Local Government—intergovernmental relationships, state and local lawmaking, state and local citizenship, services of state and local governments, financing services, executive leadership, bureaucracy, local interest groups.</td>
</tr>
<tr>
<td>3. Democratic Principles—popular sovereignty, limited powers of government, majority rule, minority rights, reserved powers, general welfare, constitutional republic.</td>
</tr>
<tr>
<td>4. Politics—political symbols, campaigns, elections, voting behavior, representation, public opinion, means and limits of political power, political socialization, political leadership.</td>
</tr>
<tr>
<td>5. Law and Individual Rights—rule of law, equality before the law, justice, the Bill of Rights, judicial principles, individual rights and responsibilities, limits of individual freedom.</td>
</tr>
<tr>
<td>7. Analytical Skills—making informed decisions on public issues, value analysis, evaluating information and information sources, interpreting and evaluating quantitative data, distinguishing fact from opinion, synthesizing data from charts and graphs.</td>
</tr>
<tr>
<td>8. Participation Skills—conflict resolution skills, news awareness skills, registration and voting skills, skills for jury duty and participation in judicial process, party and interest group participation, career awareness, securing basic governmental services.</td>
</tr>
</tbody>
</table>
Elementary and secondary teacher teams examined the possibilities for implementing improved citizenship education in the curriculum and designated content areas and specific learning activities for each grade. The staff designed a teacher education program and a procedure for increasing administrators' awareness and cooperation. A community advisory board was established. Specific project objectives were stated based on the nine goals. An evaluation plan and the needed instruments were developed near the end of the first year.

2) The second year, 1978–79, was designated for pilot testing the process. The staff development (teacher education) program was piloted, reviewed, and modified; instructional materials were reviewed, teacher-prepared lesson plans were edited and organized into teaching guides, and needed new materials were developed. A pilot study of the school implementation with matched project and control groups was conducted in the spring of the second year. Modifications were made based on pilot results.

3) The third year, 1979–80, was designated for the implementation and testing of the improvement model. Project schools and matched control schools were identified. Staff development sessions were held with teachers from the project schools in late summer and early fall. Specific objectives and recommended materials were made available to the teachers. Meetings with building administrators of project schools informed them of project goals and enlisted their cooperation. The improvement program was implemented for six weeks in the winter quarter by teachers who had participated in the staff development program. Both process and product (student outcomes) were evaluated.

4) A fourth year was added to allow a second, longer implementation of the elementary program.

**The Evaluation Plan**

In the absence of guidelines for or discussion of systematic project evaluation in the social studies literature, the evaluation team turned to the general literature on project evaluation to obtain ideas for a design. Three sources proved helpful in planning a series of steps for conducting the evaluation of the Improving Citizenship Education Project. From Rutman (1977) we obtained the idea of reviewing preconditions to the evaluation and making sure that processes and outcomes, and the linkage between the two were adequately stated. From Cook, Cook, and Mark (1977) we fastened on the idea of attempting to satisfy four types of validity in the process and project evaluation. Later, from Rossi, Freeman, and Wright (1979) we gained a perspective on determining the cost efficiency of the project, i.e., estimating the costs to users rather than long-term monetary benefits.
From these three sources we synthesized a three-step evaluation design for the ICE project (See Figure 1).

The resulting design or model is versatile and comprehensive. It can be applied in any social studies project evaluation, quantitative and/or qualitative, and it provides information on the success of all phases of an intervention. As we trace evaluation procedures based on the design, we will discuss the difficulties which arise in attempting to apply the ideals of the evaluation model to the realities of the needs and wishes of school personnel and state education officials.

**First Step — Reviewing Preconditions.** According to Rutman (1977) evaluation of a curriculum project must begin with meeting three preconditions: 1) a clearly articulated program, 2) a clearly specified set of goals and effects, and 3) a linking rationale between program and goals and effects. Without these preconditions evaluation is of little value. Ideally the evaluation plan should be developed at the earliest stages of a project to assure clarity of purposes, specificity of objectives, and linkage between project activities and goals. Consideration of these preconditions as early as the proposal writing stage would assure their realization and provide the evaluation team with sufficient time to design an appropriate evaluation plan.

Development of the ICE Project began approximately seven months before the evaluation was designed. The improvement model, the goals (Table 1) and content areas (Table 2) were developed in the early months of the project and presented to the evaluation team near the end of the first year. Because the goals as stated were too general for development of measurement instruments, the first task of the evaluation team was to develop more specific objectives. In consultation with the project staff the goals and content areas were restated as cognitive and affective objectives. (Table 3 presents the specific objectives developed from the nine project goals. Relationship of objectives to project goals is shown in Table 4.)

In addition, the evaluation team assisted the project staff in developing a linking rationale between program processes and the specific objectives. Based on the curriculum process model the linking rationale was outlined as follows:

Teachers will take part in a staff development program which will make them aware of the goals of the project, improve their skills in using various activities, teaching strategies, and materials to achieve the goals. Teachers will also be given administrative support by principals who are familiar with the purposes of the ICE project. With this background and support, teachers will spend more time on citizenship content, employ a greater variety of activities and strategies, and use a greater number of materials in their classrooms. As a result students will be exposed to an improved citizenship education program, and will show increased knowledge, skills and attitudes.

45
Comprehensive Evaluation Process for Social Studies Improvement Projects

Initial Project Evaluation Plan

Review Preconditions:
1. Clearly stated program
2. Specified goals and outcomes
3. Linking rationale between 1 and 2

If Met

If Not Met

Test Validity of Process and Product
1. Construct validity
2. Internal validity
3. Statistical validity
4. External validity

If Met

If Not Met

Cost and Efficiency Analysis
1. Effect size
2. Monetary outlay

If Not Met

Project Dissemination

Redesign

Figure 1
Table 3: Specific Objectives for the Improving Citizenship Education Project

1.0 Know and apply specific facts, basic concepts, and processes related to government and politics.
1.1 Identify specific facts, processes, and basic concepts of national government.
1.2 Identify specific facts, processes, and basic concepts of state and local government.
1.3 Identify specific facts, processes, and basic concepts of democracy.
1.4 Identify specific facts, processes, and basic concepts of politics.
1.5 Identify specific facts, processes, and basic concepts of law and individual rights.
1.6 Identify specific facts, processes, and basic concepts of global affairs.
1.7 Identify participation skills related to government and politics.
1.8 Utilize analytical skills with government and political data and issues.

2.0 Demonstrate commitment to democratic institutions, principles, and processes.
2.1 Express commitment to democratic institutions.
2.11 Identifies magnitude of agreement with statements about political institutions as others should view them.
2.12 Identifies magnitude of agreement with statements about political institutions as self views them.
2.2 Express commitment to community democratic processes.
2.21 Identifies magnitude of agreement with statements about community democratic participation as others should view them.
2.22 Identifies magnitude of agreement with statements about community participation as self views them.
2.3 Express commitment to school democratic processes.
2.31 Identifies magnitude of agreement with statements about school democratic participation as others should view them.
2.32 Identifies magnitude of agreement with statements about school democratic participation as self views them.

Table 4: Relationship between Goals and Specific Objectives of the Improving Citizenship Education Project

<table>
<thead>
<tr>
<th>Specific Objectives</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.2</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.3</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.4</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.5</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.6</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1.8</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>2.11</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>2.12</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>2.21</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>2.22</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>2.31</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>2.32</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

47
A pilot or formative evaluation was undertaken to assess the accuracy of the linking rationale. This formative evaluation was conducted in the spring of 1979 with teachers involved in the first staff development program. The results of the formative evaluation indicated that the project teachers did not produce significantly improved knowledge and attitudes in their students when compared to matched control students. (Instruments and matching procedures used to obtain control students are described in the Evaluation Design section below.)

In the pilot study, evaluation of the staff development program indicated that teachers did increase their knowledge of recommended teaching procedures. A questionnaire given at the end of the pilot treatment period to project and control teachers indicated that the project teachers 1) taught more of the content areas (Table 2) associated with the cognitive objectives than did control teachers, and 2) used more of the activities advocated in staff development sessions than did the control teachers. However, no data were collected on how the teachers used the objectives of the project. In fact, follow-up discussion with the teachers indicated that they needed specific objectives based on the goals. Apparently project teachers had interpreted the nine general goals differently from one another and the project staff.

Results of the pilot study led the project staff and evaluation team to reformulate the linking rationale as follows:

Teachers will take part in a staff development program which will make them aware of specific objectives, cognitive content areas, and goals of the project. In addition teachers will be instructed in the use of various activities, teaching strategies, and materials to achieve specific objectives. Teachers will be given administrative support by principals familiar with the specific objectives of the project. With this background and support, teachers will spend more class time teaching the content areas, employ a greater variety of activities and strategies, and use a greater number of materials in their classrooms to achieve the specific objectives. Therefore, students will be exposed to an improved citizenship education program, and as a result they will increase their achievement of the specific objectives of the project.

This case illustrates the effort which must be exerted in less than ideal situations to achieve preconditions necessary for effective evaluation. The pilot evaluation had been conducted jointly by project staff and the evaluation team. Obviously the linking rationale used in the pilot study was not adequate. Information collected from the teachers suggested that modifying the rationale by increasing its specificity would satisfy the necessary preconditions for formal summative evaluation recommended by Rutman (1977).

Second Step—Evaluation of Process and Product. Once preconditions are met, then the evaluation design should encompass two types of review: 1)
determining whether the program is carried out as planned (process evaluation), and 2) assessing the impact of the program on the students (product evaluation). According to Cook, Cook, and Mark (1977), four types of validity must be examined to assure dependability and veracity of the process and product evaluation. They are:

1) construct validity—the accuracy of manipulating the process and measuring the product;
2) internal validity—the accuracy of determining whether a selected process actually produced the observed products within the sample;
3) statistical validity—the accuracy of analyzing the data to determine association between process and product.
4) external validity—the accuracy of generalizing to a target population.

Cook et al. (1977) argue that the four validity considerations should be met in all types of evaluation of product and process related to educational improvement projects. We agreed with this notion. However, when working in the real world of the schools, we found that several compromises with these ideals had to be made.

In making compromises priorities must be set among the types of validity. We considered construct validity to be most important because an evaluation is meaningless if the accuracy of process manipulation and product measurement is not assured. We considered internal validity as second most important because failure to control rival processes would produce misinformation. Statistical validity was considered third in importance because evaluation data which were construct and internally valid could be reanalyzed if the first statistical analysis selected was invalid. Finally, we viewed external validity as relatively less important because replications could be done if this type of validity was faulty.

**Construct validity considerations.** One of the first considerations in shaping an evaluation design is how to accurately measure manipulated processes and their products. Ideally, the outcomes or products of the project should be measured using a variety of instruments and procedures. Cook et al. (1977) stress that the best means to assure construct validity when measuring product is to have multiple operationalizations of constructs (i.e., defining a concept like political knowledge in several ways). For example, to measure product related to the first goal (Table 1) we wanted to use observations of student behavior to assess actual acceptance of responsibility and to use written tests or interviews to determine if students understood these responsibilities. Likewise, the evaluation team sought to monitor processes by a variety of measures. For example, the effectiveness of the staff development program could be examined by paper and pencil tests and questionnaires plus an observation instrument to assess if teachers were implementing more citizenship content, using more citizenship activities and strategies, and using more citizenship materials. We recommended observation of project and control teachers in the classroom along with written
questionnaires. Such multiple measurements would provide information on whether teacher behavior varied as expected. For manipulated processes, Cook et al. (1977) stress that the construct validity of the processes can only be assumed if the processes vary as hypothesized. That is, did the teachers actually use more content, activities, and materials than they would without the treatment?

However, the ideal design was not possible. First, the school officials were hesitant to involve teachers or students, especially control teachers and students, in time-consuming tests and observation. A total of only four class periods was released for measurement purposes. Hence, time constraints inhibited the use of multiple measurements of processes and products. Also, the project budget and staff were not adequate for observation in 20 classrooms. Finally, the state education department required the results of the summative evaluation within a relatively short time period—by spring of 1980. As a result, multiple measurement of processes and products was impossible.

The actual measurement of the product was conducted using paper and pencil instruments—the Citizenship Knowledge Test and the Opinionnaire on Political Institutions and Participation which were developed and validated in 1978 and 1979 (Napier, Hepburn, and Strickland, 1981; Napier and Hepburn, 1981) and based on the specific objectives (Table 3) established the first year. Both instruments have elementary and secondary forms.

One of the major concerns of the evaluation team was the use of paper and pencil measures with early elementary students. Initially, the third grade was set to be the lowest grade tested. Therefore, the third grade was used to test reliability and validity of the elementary cognitive and affective instruments. Later, first and second graders were included in the study at the request of the state education agency, so a parallel cognitive measure was constructed especially for the first and second grades. The affective measure was found to work acceptably without revisions.

According to Cook et al. (1977) the multiple measures of the product constructs should covary or relate to each other, and although an ideal multiple measurement of the constructs political knowledge and political attitudes could not be used, the cognitive and affective tests were constructed with subtests. Factor analysis was used to examine the relationship among subtests, and the results indicated that there was one common factor among the subtests while each subtest retained a unique factor (Napier, Hepburn, and Strickland, 1981; Hepburn and Napier, 1980; Napier and Hepburn, 1981). Thus, construct validity for the product measures was stronger than if single, unidimensional measures were used.

The teacher education or "staff development" process was monitored using four instruments: 1) the secondary form of the Citizenship Knowledge Test, 2) the secondary form of the Opinionnaire on Political Institutions
and Participation, 3) a test of knowledge of teaching procedures, and 4) a test of attitudes toward use of the teaching procedures. In the pilot study only one instrument showed significant change from pretest to posttest. This was the test of the knowledge of teaching procedures. Project teachers in the pilot study were found to have high citizenship knowledge and attitudes as well as positive attitudes toward the use of recommended teaching procedures before their project involvement, so it was decided that it would be unnecessary to monitor those three areas in the formal evaluation study.

Administrative support and classroom procedures of both project and control teachers were monitored with questionnaires and a daily log. Three questionnaires were developed for the project teachers—one on principal support, one on objectives, and one on specially provided materials (see Margolis, 1981, for copies of these instruments.) In the pilot study the log was filled out at the end of the field test. This procedure was altered in the formal evaluation so that project and control teachers completed the logs daily during the implementation period. This change allowed them to record time spent on each content area. As it turned out, time accounted for an important variance between project and control classes. By showing differences between project and control teachers based on log summaries, the construct validity of the process was increased.

In summary, the actual procedures used to measure process and product of the project had to be a compromise between evaluation ideals and the requirements of school system and state education agency. Nevertheless, we consider the actual procedures to be as construct valid as the situation permitted. In the real world of evaluation, absolute construct validity is rare. However, it is worth the effort for evaluators and school personnel to be aware of and to attempt to increase construct validity of both process and product evaluation as much as possible within the constraints of the school situation.

**Internal validity considerations.** One purpose of any evaluation is to get an honest assessment of outcomes. To do so requires that all alternative causes of processes other than the curriculum change model or alternative causes of the products other than processes be controlled. Ideally, the most efficient means of controlling alternative cause is to employ true experimental design which uses random assignment of subjects to control and project groups. However, use of true experimental design is seldom practical in school. In this evaluation setting the school administrators would not assign teachers randomly to control and project groups because they felt that teachers who volunteered to participate in the project had to be rewarded with project participation rather than be part of the control group. It was also administratively impossible to randomly assign students to project and control classes.

Since the Improving Citizenship Education project was a total infusion
endeavor, a longitudinal study would have been the ideal way to judge its success. This would have involved continuous monitoring of both processes and products over the twelve-year schooling period. Obviously, limited resources and the funding agency deadlines prohibited even a short-run longitudinal study. We viewed the next best alternative to be trend analysis using cohorts (i.e., subjects paired by grade level) over a period of time. However, the education department required more immediate information which had to be submitted to a state review panel which would be using national Joint Dissemination Review guidelines. Hence, the trend analysis alternative was ruled out.

Since neither a true experimental design nor a trend analysis was possible, an alternative design had to be selected. Pre-experimental designs were selected for evaluation of the staff development and administrative support components of the process. Quasi-experimental design was used to evaluate classroom processes and products. (Elementary and secondary school students had to be separately evaluated because two forms of the cognitive test were not equivalent.)

For teachers in the staff development program there was no realistic control group. To give the very specific tests of the staff development process to teachers who had not participated in it would produce only guess scores. Likewise, to ask control teachers to review project objectives and principal support, when they had neither, would provide useless data. Therefore a pre-experimental design was the best option. Of course, pre-experimental designs fail to control a number of alternative causes. However, in regard to the staff development and principal support measures, the evaluators felt that uncontrolled alternative causes could be ruled out on logical grounds because they were highly unlikely to occur in this situation.

In regard to the use of quasi-experimental design, the major alternative cause which it fails to control outright is the interaction between selection and maturation or history. That is, if a difference is found between project and control subjects, it must be considered possible that it is attributable to differences in mental growth of subjects or to environmental factors rather than the treatment. To attempt to control for these alternative causes, the evaluators with the assistance of school system personnel matched project and control teachers and their students on the basis of socioeconomic information about the school population. Each project teacher was paired with a control teacher of the same course and grade level teaching students from the same socio-economic level. The use of pretests for the product evaluation allowed examination of any initial differences between project and control students.

The pilot and formal evaluations of the ICE Project employed the same design. As with any evaluation using human subjects, some problems arose in the implementation. One was the loss of subjects. Two project teachers dropped out of the study, and some students in project and control classes
failed to take pretests or posttests because of absence. Data from the matching control teacher and class were eliminated if a project teacher dropped out. The loss of a project teacher is a concern in assessing external validity which will be discussed below.

If either project or control students failed to take a pretest or posttest, they were dropped from the study. Differential rates of loss between project and control groups could lead to observed differences between project and control students on the project instruments, a matter of internal validity. However, in this study the rate of loss was the same for both project and control student groups, so this alternative cause seemed unlikely. Like the loss of teacher subjects, the loss of student subjects thus became an external validity concern.

In summary, the procedures used in this project to assure internal validity did not absolutely control alternative causes. Given the constraints of what was possible in the schools, however, it was decided that the best alternative to a true experimental design had been used by matching teachers based on socioeconomic characteristics, course, and grade of their students. Again, in the real world of project evaluation total control of alternative cause is highly unlikely.

Statistical validity considerations. When employing statistical analysis of data to ascertain an association between process and product, the power of the statistical analysis is extremely important. According to Cook et al. (1977) problems with statistical power can be minimized by:

1. using a large sample;
2. selecting subjects from homogeneous populations and using standardized measurement settings to decrease sources of errors;
3. using analysis of covariance or blocking procedures to account for extraneous sources of variation;
4. measuring the product with instruments having high reliability;
5. making the implementation of the processes standardized, and giving sufficient exposure to the processes.

Ideally, the evaluation team sought to have a large number of project teachers involved in the study. We wanted to use classes as the unit of analysis in the product evaluation which would require a large sample of teachers. As to exact size, initially a 30–35 teacher sample was considered adequate. Results of the pilot study could be utilized to secure a better estimate of the exact size needed for formal evaluation (Borg and Gall, 1979). Another initial commitment was to using multivariate analysis of covariance procedures in the product evaluation. Additionally, we recommended that the implementation period be at least 10 weeks and that the subjects include only students present for all administrations of the measurement instruments, thus eliminating students unlikely to have adequate project exposure. With these procedures, we felt that statistical validity of the evaluation design would be maximized.
The reality was that the size of the sample of teachers was dictated by the number who volunteered from the schools designated as project schools. The number of elementary teachers varied from as low as 7 in the formal evaluation to as high as 17 in the re-evaluation. The number of secondary teachers varied from a low of 19 in the formal evaluation to a high of 15 in the pilot (formative) evaluation. Thus, our ideal of a large sample was not met. An outside evaluation consultant brought in by the state education agency to review the formal evaluation plans recommended that the unit of analysis for the product evaluation be changed from classes to students because teachers could be considered part of the random nature of the treatment and the logical target unit was students rather than classes. Additionally, analyses of race and gender were of interest and using class as the unit of analyses would prohibit examination of the interaction of these two factors and the treatment. Therefore, for the formal evaluation of the product, sample size was not a problem, but in the process evaluation we were forced to use small samples.

Another problem, length of treatment, arose in the formal evaluation. Although a 10-week exposure was planned, the state education agency set a deadline for formal evaluation results by early spring 1980. Consequently, the implementation would have to take place in the preceding winter quarter which would have a scheduled Christmas break after four weeks. Staff and evaluators both felt that this break would interfere with the 10-week implementation. Therefore, a shorter 6-weeks implementation was used following the holiday. In elementary grades where daily treatment time is considerably less than in secondary grades, the shortened treatment period apparently influenced the finding that project students did not do better than their control pairs. Thus, a re-evaluation with a new elementary sample was slated for the next year using the full 10-week treatment. With that full treatment the elementary project group did score significantly higher than the control group.

Multivariate analysis of covariance had been proposed by the evaluators. We considered this statistical procedure most appropriate because the cognitive and affective measures had been found to have related subtests which were also unique. However, based on guidelines issued by the state education department, it was felt that anything other than simple analysis of covariance would confuse outside reviewers.

In summary, there was one major problem with the less than ideal procedure used in the formal evaluation. The small sample of elementary teachers and the short exposure of elementary students to the treatment led to a Type II error, the initial false conclusion that the project was not effective with elementary students. However, re-evaluation with a larger sample of elementary teachers and a full 10-week exposure in the subsequent school year indicated that the project did work with elementary students. Obviously, it is important that evaluation of an educational project attend to
statistical validity considerations to avoid such "false" conclusions. The five "minimizing" guides offered by Cook et al. are extremely helpful.

**External validity considerations.** Evaluators should be concerned with the population about which the results can be generalized (i.e., target population), and under what conditions the results will be generalizable. The ideal way to assure that the results will be generalizable to the intended population and conditions is to randomly sample subjects from the target population and design an evaluation to approximate the normal setting in which the project will be used. Under such ideal conditions the external validity of evaluation will be maximized.

In this evaluation the project teachers were volunteers, which raises some questions of external validity. However, the matching control teachers were also volunteers who agreed to administer control tests. Consequently, the target population was changed from "all teachers" to "volunteer teachers." Our concern over this was alleviated when we realized that if the project were disseminated, the teachers adopting the program would also be "volunteers." Therefore, we did not consider external validity to be threatened by this factor.

As for the evaluation of the processes, the setting did have limiting factors. Project teachers were aware of the questionnaires on principal support, use of objectives, use of materials, and the daily log. Therefore, the results may be generalized only to those situations where these instruments are used.

In summary, the external validity of the evaluation could not be fully assured by the actual procedures used. As with any "experimental" type evaluation, the external validity is not well protected. The only feasible means to overcome the inherent weaknesses when experimental procedures are used is to replicate. That is precisely what is now occurring with the ICE project. Other school systems adopting the project are being monitored using evaluation procedures very similar to the one outlined here. If the results are replicated, then the external validity of the original results will be increased. All project evaluators should encourage replication since control over external validity is nearly impossible.

**The Third Step—Project Efficiency.** If a project shows that it is effective in changing subject behavior, the final question in a comprehensive evaluation is whether the change is of practical significance. Ideally, because the ICE Project is designed as a complete infusion project, per pupil costs of applying the project throughout 12 years of schooling and the effect to be expected from 12 years of exposure to the project should be shown. In that way efficiency is demonstrated by monetary outlays and effect size. Such an analysis is termed cost-efficiency by Rossi et al. (1979).

It was impossible, for obvious reasons, to estimate the cost per pupil over a twelve year period. Instead, we estimated "start-up" costs per school
system as well as additional costs per teacher. The cost per school system involves training a local person to run a staff development workshop employing a standardized procedure (see Margolis, 1981). The cost per teacher involves purchases of the Elementary or Secondary Handbook which assists teachers implementing the project. Finally, the cost of supplemental materials is estimated per school system when one set is adequate, and per teacher when multiple sets are needed. A school system can use these monetary figures to judge whether the project is “worth” adopting.

The procedure used to estimate the effect of the project on student performance was to calculate the “effect size” of the statistically significant results. This estimate of effect is a statistical rule of thumb suggested by Tallmadge (1977, p. 34). An “effect size” is calculated by subtracting the posttest mean score of control subjects from the mean of project subjects. Then the remainder is divided by the standard deviation of the control group. The resulting value is known as the effect size. Effect size can range from 0% to 100% (or 0 to 1 if percentage are converted to fractions). For the cognitive measures, an effect size of at least 25% was sought and at least 10% for the affective measure. A lower effect size was expected for the affective measure because attitudes are difficult to change in any short-term treatment (Shaw & Wright, 1967).

In summary, cost-efficiency was estimated based on the available information from a one-time intervention. While this is not the ideal, considering the purpose of the ICE project, it was the best available information on project efficiency. Such information can be helpful to potential adopters making the decision of whether or not to adopt the project.

Reflections

Overall, the evaluation model served as a sound guide for conducting a comprehensive evaluation. The compromises we had to make because of the constraints of the school system or funding agency were based on the priorities related to construct, internal, statistical, and external validity we had established for decision-making. On reflection, however, there were two facets of the evaluation which could have been improved. One was the result of insufficient time, and the other was due to an error in judgment using the validity priority system.

First, the cognitive measure used with elementary students could have been better structured. The preparation of four instruments in the first project year resulted in less time given to the development of the elementary instrument. Given more development time the secondary and elementary cognitive instruments could have been constructed and validated as equivalent measures. We were able to do this with the affective instrument, but by the time the evaluation team realized the great value of having equivalent measures, it was too late to re-develop the elementary cognitive instrument. It is possible that the weaker instrument may have contributed
to the non-significant results in the first formal evaluation of the elementary portion of the project.

Second, as we look back, we are convinced that we should not have acquiesced to the state agency's request for early spring data. We feel that we should have stood firm in our requirement of a ten-week implementation period as the minimum approximation of true implementation. But, we had used the priority of internal validity over statistical validity in this situation. On reflection, we should have insisted that the priority of "giving sufficient exposure to the process" within statistical validity be of equal priority to internal validity. We later learned that the state funding agency had accepted delays from other projects and probably would have waited for results of a formal evaluation in the spring quarter from our project. It is likely, given the results of the re-evaluation, that the shorter implementation contributed to the non-significant results in the formal evaluation of the elementary portion of the model.

It is important to note that project development activities and evaluation progress benefited from continuing communication with the Georgia Department of Education. While the state's approach to project review was rather nebulous when the project was launched (the guidelines merely stressed monitoring processes), by the second year the Georgia Department of Education had developed a full set of formal review requirements according to which Title IV-C projects would undergo review by a panel of evaluation specialists from outside the state. The review guidelines were modeled after the Joint Dissemination Review Panel guidelines (Tallmadge, 1977) which are used for national project validation.

It was our good fortune in the state that more rigorous review requirements were established. It was also our good fortune in the project that we had designed a comprehensive evaluation including both process and product in the first year. As a result, it was not difficult for us to meet the new evaluation requirements of the funding agency. Others of the five projects in the state had designed only process evaluation plans or product estimates and found that they would be unable to provide all of the data needed for project evaluation. Some had to completely redesign their evaluations. Clearly, a comprehensive evaluation, which monitors both the process and product, not only provides all of the people involved in the project with useful feedback, but it also provides for the collection of several types of data which can satisfy the legitimate requirements of the funding agency.

The state agency also provided guidance by bringing in additional evaluation consultants to review and critique the evaluation plan. In addition, the school system provided assistance in the form of the personnel needed to make possible the monitoring of processes and products. Without the support of the social studies supervisor and department heads in the schools, we could not have gained the cooperation of project and control
teachers in the collection of extensive data. While university staff members and consultants could provide expertise in political science, social studies education, materials, and evaluation, the success of the project and its evaluation depended greatly on teacher involvement and teacher support. Teacher cooperation was crucial.

Recommendations

With so little information in the social studies literature regarding project evaluation, we hope that this discussion provides some insights into project evaluation planning and conduct. Based upon our experiences, we make the following five recommendations for all social studies projects.

1. We recommend using a comprehensive evaluation similar to the one employed in the ICE project. The comprehensive evaluation includes three steps: 1) analysis of preconditions, 2) process and project evaluation, and 3) cost-efficiency analysis. Utilizing this type of evaluation assures the project staff and evaluation team that a variety of data will be collected that will provide the practical information needed by school personnel and will meet the guidelines of nearly any funding agency.

2. Although there are numerous models which can be used to gain information about the three steps (for example, see Steele, 1973), we strongly recommend using the Cook, Cook, and Marsh (1977) model for process and product evaluation with our priorities for construct, internal, statistical, and external validity. The combination provides a helpful framework for making decisions when ideal evaluation procedures must be compromised because of the constraints of real world schools.

3. We strongly recommend involving the evaluation team at the beginning of the project. Although an ideal plan may be designed to measure processes and products of any curriculum project, the constraints of the real world may play havoc with design. Therefore, an evaluation team which includes an expert in evaluation should be involved in the project and its realities from its inception.

4. The inclusion of teachers is an important component in project development. With their involvement, the project becomes theirs, and they are more likely to give their cooperation during the implementation and evaluation of the project.

5. Finally, we strongly recommend that the staff and evaluation team of future school based projects (e.g., those funded under federal block grants to the states) go out of their way to establish mutually satisfying working relationships between school system, state education agency, and university.

Endnotes

1 An earlier draft of this paper was presented at the 1981 Annual Meeting of the National Council for the Social Studies in Detroit.
The authors wish to thank Helen W. Richardson, Project Director, and Sheila L. Margolis, Project Coordinator, for their cooperation. Without their concern for careful review of the Project, the Project evaluation and this report would not have been possible. We thank Gerald Klein of the Georgia State Department of Education for his guidance. Also, thanks to teachers, students, administrators, and parents who participated in the evaluation.

As a university member of the evaluation team, the first author acted as investigator. The second author was a member of both the evaluation team and the project staff and acted as the liaison between university and school personnel. The ICE Project is located in and was developed in the school system. Neither author had a vested interest in the success of the project. Rather, both had as their main objective lending their skills to the school system to assist in effective evaluation of the improvement project.

The major problem in using students rather than classes as the unit of analysis was making a Type I error (i.e., saying there is a difference between groups when there is no real difference). Recently Hopkins [K. D. Hopkins, The unit of analysis: Group Means versus individual observations. American Educational Research Journal, 1982, 19(1), 5-18] has suggested a means of using class as a random factor in the analysis of variance/covariance design so examination of treatment would be done by class unit while additional factors like race and gender would still be included in the design. At the time of this evaluation, however, this suggestion was not available. Instead, the use of effect size was considered a reasonable control on the Type I error possibility.

For additional information concerning the Elementary and Secondary handbooks, contact:

Sheila L. Margolis, Project Coordinator
Improving Citizenship Education Project
Liberty-Guinn Center
Fulton County School System
4820 Long Island Drive
Atlanta, GA 30342

Calculation of effect size followed the procedures suggested by Glass (Glass, G. V. Integrating findings: The meta-analysis of research. Review of research in education, Vol. 5. Itasca, IL: Peacock, 1978).

The Cook et al. model can be used with both quantitative and qualitative procedures since both should be concerned with construct, internal, statistical (in the sense of data from samples), and internal validity. The difference between the two approaches is evident in the steps used to demonstrate validity. For a recent discussion of the validity of qualitative type research, see: LeCompte, M. D. and Goetz, J. P. Problems of reliability and validity in ethnographic research. Review of Educational Research, 1982, 52 (1), 31–60.

References

Atlanta Assessment Project. How well prepared are young people in Georgia today to live in the society of the future? Atlanta: Atlanta Public Schools, 1977.


We are seeking critical reviews of scholarly works related to the concerns of social educators. This includes books on education, the social sciences, history, philosophy, research and any other works which might make a contribution to the field.

Normally, textbooks will not be reviewed with the exception of those which appear to advance theory and research. Essay reviews of two or more works on the same topic will be considered if they conform to manuscript guidelines for reviews. Reviews which exceed the guidelines for length must be handled on a case by case basis as space permits. Reviewers who have suggestions for reviews which might exceed the guidelines are urged to contact the book editor prior to submitting the review.

Reviewers should provide sufficient detail regarding the book's substance and approach, including positive and negative evaluations where relevant. Finally, the review should include the specific importance of the book for social educators.

**Manuscript Form**

The length may vary from 1,000 to 2,000 words; the manuscript must be typed, double-spaced (including quotes) on 8½" × 11" paper. The format is as follows for the top of the first page of the review, left side:

- Book Author's Name (Last Name first),
- Title, City of publication: Publisher, Date;
- Total pages; list price (if known).
- Reviewer's Name (Last Name Last)
- Institution

Submit review manuscripts to:

Professor William B. Stanley, Book Editor
Department of Curriculum and Instruction
64 Long Field House
Louisiana State University
Baton Rouge, LA 70803
Occasionally, I encounter a book whose substance and style are so appealing that I do not want it to end. Such was my experience with The Myth of Educational Reform. My expectations were high and, with very few exceptions, well-fulfilled.

The research underlying The Myth was part of a large-scale evaluation study of Individually Guided Education (IGE), a comprehensive curriculum management system developed at the University of Wisconsin R and D Center in the 1960s and implemented in numerous elementary schools in the 1970s. In addition to being a report of a four-year study of six exemplary IGE schools, The Myth is an inquiry into questions of schooling, educational research, and reform. In chapters one and two, Tom Popkewitz, Bob Tabachnick, and Gary Wehlage present an overview of the framework and methodology of their inquiry (methodological details are offered in an appendix) and provide a sociohistorical account of the genesis of IGE and its underlying assumptions about schooling, learning, and change. In chapters three-six, they present and interpret their findings regarding the different configurations of schooling they identified and, in the final chapter, consider implications for school reform and research.

IGE is a systems approach to school reform based on the assumption that increasing delivery system efficiency will increase students’ chances of success in school and society. Increasing efficiency, according to the IGE model, involves systematically sequencing and implementing behaviorally stated objectives and continuously evaluating student mastery of them. Individualization is accomplished by means of a nongraded unit organization, with each unit being composed of approximately 100 students and four to five teachers plus aides. Within each unit, students are grouped and regrouped on the basis of their progress in attaining the stated objectives. By linking a belief in individualism with faith in the rationality and effectiveness of scientific management, the IGE model is congruent with societal expectations for school efficiency, accountability, and meritocratic reward. IGE is assumed by its developers to provide a universal, nonideological model for elementary school reform.

Rather than treating IGE assumptions and procedures as given, the authors sought to describe and analyze how the IGE reform was interpreted by school personnel and students as it was incorporated into the daily ac-
tivities of schooling. Of particular interest was the extent to which the IGE reform served to change or sustain the schools' social patterns and values, particularly with respect to principles of authority, legitimacy, and control. The following questions guided the investigation:

1. How are school programs actually used?
2. What social values and meanings to reforms and their technologies generate as they are used in schools?
3. What are the assumptions behind and the implications and consequences of using reform technologies in day-to-day schooling? (p. 7)

Three interrelated aspects of schooling provide the framework for their analysis: patterns of work, conceptions of knowledge, and professional ideologies. Patterns of work refer to the activities, social relations, and sentiments that school work entails or creates. Conceptions of knowledge encompass definitions of what constitutes knowledge and how it is selected, organized, distributed, and used. Professional ideologies consist of teachers' beliefs about the appropriate relationship with students, other educators, and parents, and their beliefs about how and why students differ and learn. In the course of the book, community context is related to these three aspects of schooling. Patterns of work and conceptions of knowledge are seen as constituting a pedagogical context, which is influenced by professional ideologies that are, in turn, influenced by perceived community needs and demands.

In contrast to conventional studies of school change, the authors' purpose was not to assess extent or fidelity of IGE implementation. The six schools had already been identified as successful implementers of the IGE reform. Instead, they sought to uncover the meaning of the IGE reform in the everyday school experiences of students, teachers, and administrators by means of extensive observation, formal and informal interviews, and examination of school materials and documents. The Myth might well be viewed as a study of the social education of the children in these schools.

What was found to be the nature of the social and other education available to students in these schools? Contrary to the expectations of its developers, IGE did not provide a common school experience across the varying school settings. The schools differentially adapted the IGE model and maintained different patterns of work, conceptions of knowledge, and professional ideologies, which "reflected (or were reactions to) particular social values and interests found in the larger social/cultural community" (p. 4). On the basis of these observed differences, the authors distinguished three institutional configurations or conditions of schooling—technical, constructive, and illusory. Their observations regarding patterns of work and conceptions of knowledge in technical, constructive, and illusory schools are illustrative.

In the three schools characterized as technical, they found that students were typically offered a preplanned series of activities intended to produce
measurable competencies that could be recorded as students demonstrated mastery of them, often by completing worksheets. Content was limited, for the most part, to presumably discrete skills and bits of information. Official learning (what counted for grades) tended to be mechanical and unrelated to students' experiences, the apparent assumption being that knowledge is standardized and that right answers can be predetermined. Students were expected to strive to attain the objectives set for them—within the bounds of explicit discipline and management procedures intended to promote efficiency.

In the constructive school, students were offered opportunities to pursue their own interests, engage in a variety of activities, and examine a broad range of content and possibilities. Learning activities were designed to foster students' interpersonal and communication ability, and students were expected to demonstrate independence and initiative as well as responsibility. Comprehension rather than memorization was valued, with content often being integrated across subject boundaries and related to students' experiences. It was assumed that knowledge is tentative, that there are multiple ways of learning and knowing, and that different perspectives ought to be considered. For example, along with government structures and processes, students in a constructive school might examine various meanings of democracy and possibilities of economic and social as well as political democracy.

In contrast to the technical and constructive schools, the two illusory schools offered form and appearance but little or no academic substance. The range of content and learning activities was limited; rote memorization of information unrelated to students' experiences was stressed, the implicit assumption apparently being tended that knowledge is static and that conscientious students will accept predetermined right answers. It seemed to be further assumed that the students needed to be properly socialized before they could benefit from substantive learning opportunities. Consequently, discipline and ritual were emphasized. For example, students in an illusory school might regularly recite the Pledge of Allegiance but rarely if ever engage in examination of the concepts of republic, liberty, or justice.

These differences in patterns of work and conceptions of knowledge, which gave different interpretations to authority, legitimacy, and control, were found to be related to teachers' beliefs about particular groups of students and their needs. Professional ideologies, in turn, were found to vary with the community context of the school. While differences among the schools were related to sociocultural variations among the six communities, the schools did not simply mirror the economic situation of their communities; there was not a one-to-one correspondence between community economic status and school practices. Technical schools, for example, were found in a poor rural community, a working-middle class suburb, and an affluent, business-oriented community.
After examining IGE in use, including the social values and meanings it generated, the authors consider the functions of reform and critique conventional approaches to school reform and its study. They show how reforms such as IGE entail ritual, ceremony, and language that encourage the belief “that things are getting better,” reinforcing confidence in institutional processes and professional competence by presenting institutions and their personnel as “progressive, responsive, and, above all, consumer-oriented” (p. 169). For example, while procedures such as unit organization and computerized record keeping might encompass and justify a range of practices, they serve to “produce an image of rational enterprise,” which “creates a sense of control and projects an image of competence” (p. 170).

By providing legitimation, reform can serve to stabilize and conserve institutional patterns. It was found that the IGE reform gave credence to and thereby justified school practices and values that it was intended to change. The schools differed prior to the introduction of IGE, and the IGE reform served to maintain these differences—thus, the “myth” of school reform.

For all its strengths, I wanted even more from The Myth. I would have liked a deeper historical dimension to the study of these schools. How have they come to be the ways they are? Although there is some hint of the history of Belair and Kennedy schools, the others seem to have no past. Particularly in view of the seeming inadequacy of current economic determinist and quasi-determinist accounts of schooling, I would also like to see further exploration of contemporary school-community linkages, i.e., the socio-cultural contexts of schooling. The complexity and import of these relationships are evident; they merit our continued attention.

In sum, The Myth is powerful and sensitive. It is powerful conceptually in the range, depth, and interconnectedness of the ideas explored—beyond surface features of schooling to underlying assumptions, social meanings, and their implications—thus seriously questioning our ordinary ways of thinking and acting with respect to schooling, change, and research. Its power also derives from the interweaving of ideas and illustrations, both narrative and numerical. The Myth is sensitive to people in and associated with schools, to their situations and beliefs, and to the complexity of contemporary schooling and society. Largely unencumbered by educational or ideological language that tends to obscure more than to inform, The Myth is a compelling account of schooling and efforts toward its improvement. It challenges us to expand and perhaps reorient our visions of schooling and its study. This is an important book not only for graduate students but for all students of educational research, schooling, and reform.
Abstracts

Cognition Research: Some Implications for the Design of Social Studies Instructional Materials

A basic purpose of the paper was to consider several major areas in which cognition research would seem to have implications for the conscious design of social studies texts, and to sketch the ways in which the structure of texts might be shaped if research were applied. Some of the more salient findings from cognition research were reviewed under four headings: information processing styles, developmental differences, retention of key material, and instructional principles. Illustrative implications of these findings for the design of social studies textbooks were presented, along with suggestions for how existing texts could be modified.

School Uses of Materials Traditionally Associated with Children's Play

The structure of daily classroom interaction helps to socialize young children to appropriate pupil behavior. The incorporation of play experiences and play materials into the kindergarten curriculum provides the teacher with particularly significant opportunities to emphasize expected classroom attitudes and behavior while providing children opportunities to play. Familiar play objects in the classroom serve to ease children's transition to kindergarten while highlighting the novelty of the context in which the children now find these things. In addition, the pattern of play in the classroom introduces children to the relationship between work and play in school.

Sexuality in the Classrooms of Teachers with Various Sex Role Orientations

An ethnographic study of teacher-student interaction in the classrooms of high school teachers with sex role ideologies classified as Traditional, Androgynous, and Mixed resulted in unanticipated findings of sexuality as an apparently unconscious element in teacher behavior. The expressions of sexuality appear to be related to the sex role ideology subscribed to by the teachers. An attempt is made to analyze components of sexual expression and to point out the lack of attention in the literature to this aspect of teacher personality. Finally, the implications for social studies education are stressed.

Social Studies Project Evaluation Case Study and Recommendations

The paper presents a comprehensive evaluation model which was designed for a social studies improvement project in a large school system. The model involved three components: reviewing preconditions, evaluating process and product, and judging project efficiency. Within each component, the ideal procedures selected for obtaining information is presented followed by the actual procedures used in a
specific case. The case study illustrates the real world constraints on an ideal project evaluation plan, and the decision-making process used to formulate the actual evaluation plan. The authors include reflections on their experience in making decisions with the project evaluation and offer recommendations for evaluation of social studies projects.