Any attempt to derive conclusions from the literature regarding leadership in education is limited because much of the literature consists of untested opinion and most empirical studies are based on information from one school or one school district and are thus very limited. Most studies utilize variables that are unique to the particular study and/or are nontheoretical, thus providing little in the way of a framework for integrating findings from different studies. Generally, studies fail in specifying the conditions under which a relationship between variables is tested, thereby running the risk that an actual relationship is undetected; and investigators tend to use perceptions of involved persons rather than observations by a third party as sources of information regarding variables, thus making findings subject to the attitudes or the memory of participants in the organization. In many cases, information regarding both variables is obtained from the same respondents, thus inflating the obtained measure of relationship because of a response set. Finally, the models used for differentiating among behavior patterns provide for limited alternatives and frequently involve one specified behavior (with the only other option being the absence of the one specified), thus running the risk of not differentiating what are actually meaningfully different behaviors. (Author)
CHARACTERISTICS OF RESEARCH ON LEADERSHIP IN EDUCATION

Paul C. Buchanan, Ph.D.

November 1971

Social Learning Curriculum
F.G.S. Yeshiva University, New York, N.Y.
Herbert Goldstein, Director

November 1971
Curriculum Research and Development Center Staff

Director: Dr. Herbert Goldstein

Development Staff: George Mischio, Coordinator-Academic Skills Program
Sarah Olberg, Coordinator-Social Learning Curriculum
Jerry Bernstein, Development Project Specialist
Pat Buxton, Development Project Specialist
Nancy Chaikin, Development Project Specialist
Sam Frankel, Development Project Specialist
Karen Kass, Development Project Specialist
Judy Walden, Development Project Specialist
Mary Korb, Development Project Assistant

Research: Dr. I. Leon Smith, Research Coordinator

Evaluation Staff: Dr. Barry Lehrer, Evaluation Coordinator
Gerald Melnick, Senior Evaluation Scientist
Sandra Greenberg, Evaluation Specialist

Field Operations Staff: Marjorie Fratkin, Field Operations Coordinator

Media Staff: David Bocher, Media Coordinator
Carol Richter, Media Specialist

Administration: Caroline Cordts, Administrative Assistant and Editor
Dale Zevin, Administrative Secretary

Curriculum Research and Development Center in Mental Retardation
Department of Special Education
Ferkauf Graduate School of Humanities and Social Sciences
Yeshiva University
55 Fifth Avenue
New York, New York 10003

The Curriculum Research and Development Center is supported by the U.S. Office
of Education, Bureau for the Education of the Handicapped (HEW) Project
Number 6-1368

(c) Copyright - 1969 - Curriculum Research and Development Center in
Mental Retardation
Yeshiva University, New York
The purpose of this study is to examine the quality of research regarding leadership in public education.

This issue arose in the context of an attempt to determine what is known about (a) prevailing patterns of interaction between leaders and other adults in public schools and about (b) conditions under which specific patterns of behavior are connected with (i.e., cause) differences in the performance of others. It soon became apparent that limitations in the design of studies critically restricted any generalizations which could be made regarding either of these questions. These limitations and their implications are discussed in the remainder of this paper.

Opinion vs. research.

Much of the material which fills the innumerable volumes which have "supervision" or "leadership in education" in their titles consists of statements regarding what a person or a committee thinks should be done by a leader, and in many cases such statements are presented as if they were descriptions: "the effective supervisor does..."! (See ASCD, 1960, as an example.) I have checked the footnotes which are frequent in such publications, and a high proportion turn to be not research but opinions of another person or committee. Several of these conclusions - "principles" - have been subject to empirical research and found to be unverified (Gross and Herriott, 1965; Hemphill, et al., 1962).
It is my impression that the literature on education contains a much higher proportion of "practitioner opinions" which are neither supported by empirical research nor by sound theory than in comparable literature regarding behavior in other types of organizations (industry, hospitals, etc.), and that practices considered desirable tend to be more faddish than in the other fields. (Harris [1963] states that the purpose of his book is to "order and relate many of the diverse and sometimes conflicting concepts of supervision that are in vogue from time to time" p. viii.) Whether this is so, and if so, what leads to this condition, would appear to be a worthwhile study in its own right. Lack of money for research is undoubtedly one contributor; while 4.6% of the expenditures on the nation's health is devoted to research, development and innovation, only 0.3% of similar funds in education is for these purposes (according to a study by Rand Corporation, quoted in Hot Line [1971]).

Sampling.

Of the empirical studies, most are based on information from one school system which typically is selected not as a sample but as a situation which is available. While recognizing the difficulty of obtaining meaningful samples, and that one-school studies may be very useful when applied to the school in which they were conducted (Fox, 1969: 98), one must note this limitation in attempting to generalize from them.

Theoretical base.

Another characteristic of educational research on leadership is that with few exceptions, such as the work of Getzels (1969) and Halpin (1966), empirical work is either atheoretical or utilizes constructs which are unique to a particular study. As a consequence, as Getzels (1969: 461) points out, there is little in the way of "an articulated framework...within which to see the
relationships among topics, among the settings, and between the topics and
the settings."

Related to this is an observation by Halpin (1966: 121) that no
satisfactory "ultimate" criteria of leader behavior (i.e., measures of change
in organizational output attributable to the behavior of the leader) have been
devised. Thus various types of "intermediate" ones are used, such as personal
characteristics (he is "viewed positively," [Bidwell, 1957]; he "generates
confidence" [Hills, 1963]). Other types consist of the effectiveness of performance
as judged by one or more role groups, and in some such cases the aspect of
performance being judged is not even specified. An observation made by Halpin
in 1957 still appears to be the case:

The most critical research task ... in educational ad-
ministration is to prove that the various intermediate
criteria of administrator effectiveness now used ... are, in fact, significantly correlated with changes in
the organization's achievement. (Halpin, 1966: 53.)

Contextual considerations.

Recently, an increasing number of studies have attempted to link contextual
and organizational variables to differences in the output of schools (Coleman,
et al. 1966; Dyer, 1968; McDill, et al. 1967). However, these typically have
not included leader behavior as one of the organizational variables. Studies
of leadership tend to be one-dimensional rather than multi-dimensional; they
seldom specify the conditions under which relationships between variables are
tested even though there is rather strong indication that conditions make a
difference (Gibb, 1969: 268).

Sources of information.

Most studies of leadership in education are based on perceptions of people
involved in the situation. While this yields important information regarding
sentiments, beliefs, etc., exclusive use of such data in a given study, as White (1969: 48) points out, leads us to "study (people's) reactions without learning what it is they are reacting to." Furthermore, the nature of reactions obtained depends upon the attitude of the respondents to the researcher and to the situation (Argyris, 1968), and as Gross and Herriott (1965: 217-235) have shown, responses vary significantly with the age, sex, and career satisfaction of respondents. As was determined in the only study of the question which I am aware of (Dean, 1958), even when the task of the respondent is to provide factual information there is great divergence from that noted by a trained observer.

Studies of leadership behavior are also contaminated by the fact that, in many, information regarding both variables whose relation is to be examined is the perceptions of the same sample of respondents. For instance, in a study (Hills, 1963) of the relationship between some aspects of the principal's performance and the morale of his staff, information on both variables was obtained from the same teachers. Such procedure runs the risk of spuriously increasing the magnitude of the obtained relationship due to the "set" of the observer. The magnitude of this increase is suggested by a study by Gross and Herriott (1965: 52-3). Using the same teachers as sources of information on both variables, they found a correlation of .71 between teacher morale and teacher performance. They then split the teachers from a given school into two halves and used the responses from one sub-group to assess morale and the other half to assess performance. The two correlations which they obtained (i.e., one using sub-group A as the source of information regarding morale and sub-group B as the source regarding performance, and the other using sub-group B as the source regarding morale and sub-group A as the source regarding performance) were .47 and .50, even though the correlation was corrected for reduced size of sample. While
these correlations were still statistically significant, the proportions of variance apparently accounted for, dropped by more than half (from .56 to .22 and .25).

**Limitations of models.**

Research on leadership is further limited by the fact that most studies utilize models or typologies which provide for limited differentiations in leader behavior (Buchanan, 1970). For example, the Lippitt-White typology which consists of three "styles" of leadership is typically treated as a continuum between democratic-autocratic (Gibb, 1969: 258). When this is done, behavior which cannot meaningfully be encompassed into one of the two categories must either be ignored or included with Procrustian effects. This means that failures to find differences in consequences of behaviors categorized by the model can be due to the fact that behavior patterns which do have predictable consequences were not "captured" by the model, or to the Procrustian effects, rather than to there actually being no consistent relation between leader behavior patterns and other variables.

Because of the importance of this question three studies which illustrate this aspect of research on education will be presented here in some detail. It will help clarify the point if, in presenting the summaries, we use a model which makes it possible to differentiate among behaviors. The "Managerial Grid" of Blake and Mouton (1964) involves a typology of seven "styles": integrative, autocratic, compromise, human relations, laissez-faire, paternalistic, and "wide arc pendulum" (the latter meaning a swing, over time, from autocratic to human relations or from autocratic to laissez-faire). For purposes of this discussion we will assume that these seven styles are conceptually distinct.1

---

1 Blake and Mouton avoid use of such titles as I have used, since such terms tend to be loaded with value connotations. However, these terms are adequate approximations here and are used to avoid the lengthy explanation required to meaningfully present Blake-Mouton terms. The reader is strongly encouraged to read the original work by Blake and Mouton.
In a well-known study, Moeller (1964) undertook to determine the relationship between the teachers' sense of power and the extent of bureaucratization of the school system. While he expected to find an inverse relation—i.e., the greater the extent of bureaucratization the less the feeling of power—he found the opposite. His hypothesis was based on the expectation that "Teachers in (bureaucratic) systems are confronted with increased regulations, structuring of the curriculum, and other bureaucratic devices of coordination and control ... (which) induce in teachers a sense of powerlessness to affect school policy." He explains his unexpected findings by noting that predictability of policies, decisions, and the like is a major characteristics of high but not of low bureaucratic systems:

We may postulate that teachers' knowledge of ... policy is, in itself, a form of power. When policy is applicable to all, then any individual who knows the rules ... is able to predict how any particular situation will be handled. This factor enables the teacher ... to anticipate how the administration will act in most problems ... knowledge of policy enables teachers to know the most effective course of action to take in order to influence the policy-maker.

His study leaves the implication that bureaucracy is desirable since it reduces the sense of powerlessness of teachers: "Certainly bureaucracy ... does not induce in teachers feelings of powerlessness or alienation from the system."

Aside from drawing an unwarranted conclusion from perceptual data this study suffers from the fact that unclear and restricted options underlay its construction: school systems were characterized only in terms of being high or low in bureaucracy. This means that those systems described as low on bureaucracy were actually not characterized, although Moeller proceeded as though they were.

---

2One can conclude from Moeller's study that bureaucracy is associated with less powerlessness than is non-bureaucracy, but not, as Moeller did, that it does not induce powerlessness. The responses he analyzed were teachers' perceptions, a kind of response which provides only ordinal data: but to assert that bureaucracy does not induce powerlessness would require responses yielding ratio measurements. (See Gross and Herriott [1965: 24-25] for discussion of limitations of perceptual ratings.)
For instance, he concluded that communication was unpredictable in schools which were ranked low on "communication through established authority" - one of the dimensions of bureaucracy used in the study. Clearly, unpredictability is one possibility (the laissez-faire option in Grid terms) when respondents report that communication does not take place via established authority. But equally clear is the possibility that communications can be through close personal relations with the supervisor (human relations option), or through organic arrangements (integration option), etc., and still be predictable. Thus one does not know from Moeller's study if those schools in which teachers felt a lower sense of power were characterized by a laissez-faire, human relations, autocratic, paternalistic, or a pendulum pattern of leader behavior. On the basis of other studies and related theory (Lippitt and White, 1952; Coch and French, 1948; Likert, 1967) one would expect that the teachers' sense of power would vary greatly under these different types of leader behavior; but in studies where one category includes an unspecified mixture of types, differential effects of each would be likely to "wash out." Thus Moeller's conclusion that bureaucratic systems are desirable is clearly not justified on the basis of his study.

Another research study which illustrates the limiting consequences of using undifferentiated options is Gross' and Herriott's (1965) work on "staff leadership in public schools." In this study, leadership was conceptualized as "the effort of an executive...to conform to a definition of his role that stresses his obligation to improve the quality of staff performance" (p.8). Operationally, the principals who were the focus of the study were ranked on the basis of their scoring higher or lower on Likert-type scales which constituted the measure of Executive Professional Leadership (EPL). This is a unique way of conceptualizing leadership differences, making it difficult to compare this study with others. However, examination of the questionnaire items which make up the EPL scale indicates that the scale includes several behaviors typical of the integrative
Grid style; but unlike the integrative Grid style the focus of EPL is clearly on the teacher in the classroom, not at all on involving the teacher in policy or other school-wide issues; and there is no clear indication of how conflict is viewed and resolved by principals ranking high in EPL. As in the Moeller study, Gross and Herriott leave unspecified the behavior of principals who were described as lower on EPL - one only knows they are viewed as not doing what EPL calls for, and apparently this category includes unspecified proportions of principals whose behavior is characterized as laissez-faire, autocratic, etc. Gross and Herriott's comparisons, then, are between the impact of principals who are high on EPL vs. principals having an unknown mixture of other styles. This may account for the relatively low relationship found between the principal's EPL and teacher performance (Gross and Herriott, 1965:57; Hilfiker, 1969).

It will be useful to consider one other study in the field of educational leadership, both because it conceptualizes leadership in terms of a restricted number of options, yet gets at some of the fundamental issues involved in leadership. Seeman (1960) began his study with the belief that "analysis of cultural background, even if incomplete and tentative, can provide a rich source of hypotheses regarding leadership in specific institutions" (p.1); and he set out with the question: "In what way is leader and member behavior a function...of conceptions or social structures which characterize the culture of which both the group and its leader are a part?" (p.4). One variable in the study - "leadership ideology" - was the view of school administrators and teachers (from 26 communities in the midwest) regarding "an ideal school leader." To obtain information on this, he constructed a 10-item forced-choice instrument such that "in each pair of choices there is a relatively directive (leader-centered) and a relatively non-directive (member-centered) alternative" (p.21). Examination of the specific scales in the instrument Seeman used indicates that
in Grid terms the directive responses portrayed a mixture of autocratic and compromise behaviors and the non-directive ones a mixture of human relations and laissez-faire. Thus an integrative option was not provided. Since the choice in each item was between two specified alternatives, Seeman's study avoids the shortcoming of "unspecified opposites" encountered in Moeller and in Gross and Herriott; but since the respondent was forced to choose between only two ideologies, and each of these included a mixture of behavior styles, one does not from this study know how much distortion was generated by not providing other and more differentiated options.

Because of the quality of research which it represents and because of its extensive use in education, it is relevant to note the Leader Behavior Description Questionnaire (LBDQ) and the model underlying it (Halpin, 1966; Gibb, 1969). This model characterizes behavior in four conceptually meaningful categories (high-low on "consideration" and "initiating structure"), and thus goes a long way in avoiding the limitations of such models as those summarized above. Yet in the LBDQ, conflict resolution is not included as a behavior (only one of the 30 items in the questionnaire deals with conflict - "He speaks in a manner not to be questioned") and the categories do not provide for integrative actions. In his more recent work, Halpin (1966), one of the developers of the concept and of the instrument, indicates awareness of shortcomings of it. He sees ambiguity in the concept of "consideration" (Halpin, 1966: 197-8), notes that the instrument deals with phenotypical rather than genotypical issues (202), and that it does not involve authenticity (203-224) - all of which appear to be related to what I am referring to as an integrative behavior pattern.

One further observation regarding the concepts and variables used in research: for the practitioner one of the important things, as Halpin (1966: 202) has pointed out, is that concepts be ones along which change can be
induced - they must be modifiable. They also need to be of sufficient importance that if change is induced in the behavior underlying the variable it will be more than trivial. One aspect of modifiability is the extent to which the behavior described can be observed directly or whether it is inferred - a distinction which is especially important if data is to be used as feedback in a training project. This is a condition difficult to meet when constructs are generated by factor analysis or similar procedure. Since many of the constructs used in research on leadership in education are a theoretical and many of the more rigorous studies involve factor analysis, it appears likely that such research will have restricted practical value. One should also note, however, that it may be difficult to get hold of variables which are at the same time directly observable and at a genotypical level.

Summary.

Any attempt to derive conclusions from the literature regarding leadership in education is limited by the fact that (a) much of the literature consists of untested opinion; (b) most empirical studies are based on information from one school or one school district (and are in effect single case studies) and thus are of very limited generalizability; (c) most studies utilize variables which are unique to the particular study and/or are non-theoretical, thus providing little in the way of a framework for integrating findings from different studies; (d) most do not specify the conditions under which a relationship between variables is tested, thereby running the risk that an actual relationship is undetected; (e) most use perceptions of involved persons rather than observations by a third party as sources of information regarding variables, thus making findings subject to the attitudes and in many cases to the memory of participants in the organization; (f) in many cases information regarding both variables is obtained from the same respondents, thus inflating the obtained measure of
relationship due to a response set; and (g) the models used for differentiating among behavior patterns provide for limited alternatives and frequently involve one specified behavior with the only other option being the absence of the one specified, thus running the risk of not differentiating what are actually meaningfully different behaviors.
REFERENCES


Buchanan, Paul C. Some thoughts on leadership and teaching in the public schools, 1970 (mimeograph).


Dean, Lois. Interaction reported and observed: the case of one local union. Human Organization, 1958, 17 (3), 36-44.


Hilfiker, Leo R. The relationship of school system innovativeness to selected dimensions of interpersonal behavior in eight school systems. Technical Report No. 70, Madison, Wis.: Center for Cognitive Learning, 1969.


Hot Line, 1971, 4 (1), Published by American Association of School Administrators, Washington, D.C.


