

1969

Organizational Climate, Teacher Age, and Student Achievement.

Dennis Jerome Deloria
Northern Michigan University

Follow this and additional works at: <https://commons.nmu.edu/theses>

Recommended Citation

Deloria, Dennis Jerome, "Organizational Climate, Teacher Age, and Student Achievement." (1969). *All NMU Master's Theses*. 218.
<https://commons.nmu.edu/theses/218>

This Open Access is brought to you for free and open access by the Student Works at NMU Commons. It has been accepted for inclusion in All NMU Master's Theses by an authorized administrator of NMU Commons. For more information, please contact kmcdonou@nmu.edu, bsarjean@nmu.edu.

TITLE OF THESIS

ORGANIZATIONAL CLIMATE, TEACHER AGE,
AND STUDENT ACHIEVEMENT

by

DENNIS JEROME DELORIA

(name)

This thesis is recommended for approval by the student's thesis committee.

Richard I. Egelston

Chairman

A. C. Ferguson

Approved by *Roland S. Strover* _____, Dean of Graduate Studies.

January 19, 1970

(date)

Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Arts.

Northern Michigan University
Marquette, Michigan

(date)

ORGANIZATIONAL CLIMATE, TEACHER AGE,
AND STUDENT ACHIEVEMENT

A Thesis
Presented to
the Faculty of the Graduate School
Northern Michigan University

In Partial Fulfillment
of the Requirements for the Degree
Master of Education

by
Dennis Jerome Deloria
August 1969

ProQuest Number: 10804842

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10804842

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Statement of the Problem	1
Methods	2
Hypotheses	4
Hypothesis One	4
Hypothesis Two	5
Hypothesis Three	5
II. BACKGROUND	7
The Problem of Criteria	9
Armchair Approach	10
Rational-analytic Approach	14
Empirical Approach	19
Three Classes of Criteria	22
Conclusion	27
Potential Correlates	28
Central Variables	28
Noncentral Variables	29
Two Relevant Potential Correlates	31
Organizational Climate	32
Relevant Research	36
Teacher Age	36
Organizational Climate	41

CHAPTER	PAGE
III. EXPERIMENTAL DESIGN	45
Instruments	45
Stanford Achievement Battery	45
Lorge-Thorndike Intelligence Test	46
Organizational Climate Description	
Questionnaire	49
Development of the OCDQ	50
Data Collection	54
Scoring	55
Statistical Analysis	57
IV. RESULTS	60
Achievement Analysis of Covariance	60
Hypothesis One: Open vs. Closed	68
Hypothesis Two: Young vs. Old	74
Hypothesis Three: Young/Open vs. Old/Closed	78
Correlational Analysis	80
V. DISCUSSION	86
BIBLIOGRAPHY	91
APPENDIX A: THE OCDQ INSTRUMENT	95
APPENDIX B: THE EIGHT DIMENSIONS OF ORGANIZATIONAL	
CLIMATE	97
APPENDIX C: OCDQ ITEMS ARRANGED BY SUBTEST	99
APPENDIX D: OCDQ CLIMATE DESCRIPTION	102

Chapter I

INTRODUCTION

Statement of the Problem

In this study two variables were examined for their possible influence on the process of education. The first variable investigated was "organizational climate," as measured by an instrument recently devised by Halpin and Croft (1963). Organizational climate can best be described as the prevailing mood, or degree of harmony, that exists among teachers and between teachers and their principal. The sharp observed differences between Open and Closed climate schools in the dimension of school climate (Halpin, 1966) lead one to expect parallel differences in other dimensions of school functioning, such as, perhaps, student achievement.

The second variable under examination is teacher age. Recent research has suggested that teachers become less effective as they approach old age. "Generally speaking, scores of teachers fifty-five and above showed this group to be at a disadvantage when compared with young teachers. . . (Ryans, 1960)." Peterson says, "There are some indications that young adulthood is the most effective age for a teacher (1964)." Indeed, Halpin suggests that perhaps the Closed

climate might characterize schools made up of older teachers (1966), implying that systematic relationships might exist between teacher age and school climate.

The purpose of this investigation was to examine schools that differed with respect to teacher age, school climate, or both, to see whether they also differed with respect to pupil achievement levels.

Methods

The facilities and financial resources of the Upper Peninsula Multi-District Planning Project (UPMDPP) were used for this study, in both the data-gathering and the data-analyzing processes. The UPMDPP was an Elementary and Secondary Education Act Title III planning project (OE 66-1397), funded by the United States Office of Education, and sponsored by the Ishpeming, Michigan, school district. Students and teaching staffs throughout the Upper Peninsula were the object of extensive investigation for the UPMDPP, and a portion of the resulting data were used for this study. Approximately 18,000 students drawn from the fourth, sixth, ninth and eleventh grades were tested. They represented 104 public and private school systems occupying almost 250 school buildings, with a combined total student enrollment of about 70,000. Almost

1,300 teachers and principals, out of approximately 3,500 possible, completed questionnaires describing their personal background and their schools' climates. Further information about the UPMDPP can be obtained from the final project report (Egelston et al., 1967).

The Organizational Climate of each school was determined by means of the recently-devised Halpin-Croft Occupational Climate Description Questionnaire (OCDQ). This questionnaire, composed of sixty-four items with Likert-type responses, contains eight subtests: disengagement, hindrance, esprit, and intimacy (descriptive of teacher behaviors); and aloofness, production emphasis, thrust, and consideration (descriptive of principal behaviors). The OCDQ separates schools into six climate categories: open, autonomous, controlled, familiar, paternal, and closed, depending upon how closely each school's subtest profile resembles the normative prototypic profile for each category.

The ages of teachers in the sample were obtained from the birthdate item in the identification grid of the OCDQ answer sheet. Teachers were divided into three age groups: young, up to age thirty-five; middle, ages thirty-six to fifty-five; and old, fifty-six and over. All comparisons were made between the young group and the old group; the middle group was dropped from the sample.

The student achievement levels for each school were determined by administering the appropriate Stanford Achievement Battery to all students in each of the four grades tested. The Stanford Achievement Test is described as ". . . a series of comprehensive achievement tests developed to measure the important knowledges, skills and understandings commonly accepted as desirable outcomes of the major branches of the elementary curriculum (Kelly, Madden, Gardner & Rudman, 1964)." Battery subtests include word meaning; paragraph meaning; spelling; language; arithmetic computation, concepts, and applications; social studies; and science.

To insure that school differences were due to differences in achievement, rather than differences in intelligence, Loevinger-Thorndike Intelligence Test scores were used as a covariate to control the effects of intelligence upon the achievement scores.

Hypotheses

Hypothesis one. This hypothesis tests the effects of school climate upon student achievement. Since Halpin (1966) states that Closed climates appear to be less desirable than Open climates, differences in student achievement might be expected. Halpin describes an Open climate as one ". . . in which the members enjoy extremely high Esprit . . . The behavior of the principal represents an appropriate integra-

tion between his own personality and the role he is expected to play as principal." On the other hand, "the closed climate marks a situation in which group members obtain little satisfaction . . . group achievement is minimal. . . the principal is ineffective in directing the activities of the teachers." Evolving from the stated differences between these climates, the first hypothesis can be formulated as follows:

H₁: students from schools described by the OCDQ as having an Open climate will score higher on achievement measures than students from schools having closed climates.

Hypothesis two. This hypothesis tests the effect of teacher age on student achievement. Following the remarks of Ryans (1960) and Peterson (1964) cited above, the second hypothesis can be stated as follows:

H₁: students from schools employing more young (under 35) than old (over 56) teachers will score higher on achievement measures than students from schools employing more old than young teachers.

Hypothesis three. This hypothesis tests the combined effects of school climate and teacher age upon student achievement. Although neither condition alone tested in the hypotheses above might be sufficient to influence student achievement appreciably, both together might have an important effect. Presumably, the subset

of schools having both an Open climate and mostly young teachers would be more conducive to student achievement than the opposite group of schools.

H₁: students from schools jointly in the Open and young categories above will score higher on achievement measures than students from schools in the Closed and old categories above.

Chapter II

BACKGROUND

Much past research on teacher effectiveness has fallen short of intended objectives because insufficient thought was given to preliminary decisions about the dependent and independent variables investigated. This chapter, first, examines theoretical considerations influencing the researcher's choice of a dependent variable; second, presents a theoretical framework to categorize pertinent independent variables; and third, describes several previous studies relating to the present investigation.

The two basic aspects of any research on teacher effectiveness, as outlined in the Handbook of Research on Teaching (Gage, 1963), are the potential correlates studied and the criteria of teacher effectiveness used. According to the Handbook, whenever the idea of effectiveness enters research, the problem of a criterion of effectiveness appears. The structure of the research takes this form:

Identify or select a criterion (or set of criteria) of teacher effectiveness. This criterion then becomes the dependent variable. The research task is then (1) to measure this criterion, (2) to measure potential

correlates of this criterion, (3) to determine the actual correlations between this criterion and its potential correlates (Gage, 1963).

Early in his research report, Characteristics of Teachers (1960), Ryans gives this description of what constitutes a "criterion" as the word is used in a research context:

A criterion is a standard description, or definition, which is accepted in undertaking research and is used to provide a frame of reference for judging whether or not some phenomenon occurs (and often the degree to which it occurs). It is a base, often of a rather arbitrary nature and ultimately involving value judgments, against which comparisons may be made (Ryans, 1960).

A potential correlate can be any one of many variables in the school situation, but in order for an investigation to be research on teaching, according to the Handbook, at least one of the potential correlates must be a central variable:

A central variable in research on teaching is one referring to a behavior or characteristic of teachers. Such variables are central by definition, not as a matter of empirical determination, because they represent the center of our concern (Gage, 1963).

A survey of past research on teacher effectiveness reveals much disagreement both about the criteria to be used in judging teacher effectiveness, and also about the potential correlates of these criteria. Since both criteria and correlates are of fundamental importance in any research dealing with teacher effectiveness, yet are quite

distinct from one another, they are discussed in two separate sections of this chapter. The last section of the chapter reviews relevant research.

The Problem of Criteria

True, it may be said that teaching is effective to the extent that the teacher acts in ways that are favorable to the development of basic skills, understanding, work habits, desirable attitudes, value judgments, and adequate personal adjustment of the pupil. But even such an operational-appearing definition really is very general and abstract and is not easily translatable into terms relating to specific teacher behaviors.

Embarrassing as it may be for professional educators to recognize, relatively little progress has been made in supplementing this definition with the details that are necessary for describing competent teaching or the characteristics of effective teachers for a specific situation or cultural setting (Ryans, 1960).

Here Ryans provides the distinction between effective teaching, which is a global concept, and effective teaching for a specific situation, a more limited concept which can be concretely applied to particular geographical areas, cultural settings, subject areas, or grade levels. Since subcultures judge different teachers and methods to be effective according to their individual objectives, teacher effectiveness is not an absolute concept but one that changes from setting to setting. Herein lies one major difficulty in selecting criteria around which to build meaningful educational research:

to choose any criterion requires subjective value judgments, which will be different for different people with different backgrounds. Therefore, the act of choosing a criterion lies beyond the domain of empirical research and must ultimately be a decision made by some individual or group.

Armchair approach. According to Ryans, the actual decisions about criteria appear to have been arrived at by one of three approaches, which he named the "armchair approach," the "rational-analytic approach," and the "empirical approach" (Ryans, 1960). The armchair approach to criterion selection seems to best describe most early efforts in research on teacher effectiveness:

In spite of the importance of the criterion problem, it frequently has been neglected or at best treated in a commonsense manner with little apparent attention to possible biasing conditions. Thus, hundreds of "criteria" for teacher load, salary schedules, vocational success, and the like have been arbitrarily set up and employed by professional educators. Typically, such criteria are based on an armchair approach to the criterion. This approach is highly subject to intentional and unintentional selection, or bias, in that it utilizes unanalyzed retrospective impressions, based upon non-systematic observation and often characterized by free association. As a result, it is likely to result in incomplete and unsatisfactory descriptions of criteria (Ryans, 1960).

One early research effort that attempted to describe successful teachers, apparently using the armchair approach to criteria selection, is the Commonwealth Teacher Training Study completed by Charters and Waples in 1929. In this

study forty-one administrators, twenty-seven teachers, fourteen parents, ten pupils, three representatives of teachers' agencies, and two professors of education were interviewed to discover what they believed to be the traits most essential for success in teaching. Trained interviewers then asked selected questions of the interviewees, reproduced in part below:

1. What are the important qualities that a high-school teacher should possess? How are they shown?
2. Think of the best teacher among all the high-school teachers you have ever known. Why was he or she the best teacher?
3. Think of the poorest teacher whom you have ever known among your co-workers. Why was he or she so poor? (Charters & Waples, 1929)

From these interviews were secured approximately 2,800 actions indicative of particular traits. These actions were then "translated" by twenty-one judges, consisting of graduate assistants and hourly employees, who stated what traits were considered to be exemplified by each action. Then the staff defined the traits by consulting four dictionaries and selecting the definition agreed upon by at least three of the four. This produced a list of eighty-three action-traits, which were then "telescoped" and ranked by a new group of fifteen judges to twenty-five traits, a number chosen beforehand because it was considered practical for use in teacher-training institutions. The

final list of twenty-five traits, synthesized from the original 2,800 trait-actions, is reproduced in Table II-1.

TABLE II-1
FINAL TEACHER TRAITS

1. Adaptability	14. Industry
2. Attractiveness	15. Leadership
3. Breadth of interest	16. Magnetism
4. Carefulness	17. Neatness
5. Considerateness	18. Open mindedness
6. Co-operation	19. Originality
7. Dependability	20. Progressiveness
8. Enthusiasm	21. Promptness
9. Fluency	22. Refinement
10. Forcefulness	23. Scholarship
11. Good judgment	24. Self-control
12. Health	25. Thrift
13. Honesty	

(Charters & Waples, 1929)

These researchers apparently felt that no individual could give a complete description of the ideal teacher, but then felt the problem of fragmentary knowledge could be overcome by polling enough people, "assuming that the persons to be interviewed know successful and unsuccessful teachers as well as good and poor teaching . . . (Charters & Waples, 1929)." The researchers evidently assumed that the interviewees would select some criteria for judging teachers to be good or poor, since the questions did not specify any criteria against which the teachers' actions

might be compared. In the last analysis, each trait-action can represent only an opinion, formed perhaps without any deliberate consideration either of the broad educational objectives of school systems, or of the measured effects of those traits upon actual pupil changes. The final traits are also lacking in precision; the same process that compressed the 2,800 specific actions into twenty-five general traits also deprived them of much of their meaning. What, for example, constitutes "good judgment" in a teacher? It seems likely that if a teacher's actions in a particular situation were reviewed by a group of judges, there would be disagreement as to whether or not good judgment was actually used. The decisions would probably be affected greatly by the judges' backgrounds and mores, but probably would not be affected as much by the actual changes produced in students by the act of "good judgment."

Charters' and Waples' study is deficient because it concentrates its efforts on just one of the two aspects of research described by the Handbook, that of potential correlates, ignoring the relation of the potential correlates to any but armchair criteria. The list compiled in this study would still be extremely useful to researchers if the traits were operationally defined and used as theoretical predictors of criteria to guide research efforts.

After examination of many early studies, the Committee on the Criteria of Teacher Effectiveness summarized their shortcomings in these words:

In the earlier decades of this century, research seems to have been aimed at discovering the characteristics of the effective teacher. Teacher effectiveness was generally assumed to depend on one or more pattern or dimension which held true regardless of the kind of person the teacher was, the kinds of pupils he affected, the nature of the educational program, the kind of school administration and supervision, or the kind of environment of which the pupils were a part. . . . Further, if the assumption is sound that teacher effectiveness is a function of the extent to which it represents changes in pupils, then the ratings of teachers so commonly used do not correlate highly with such a criterion or other multidimensional criteria (Remmers, et al., 1953).

Rational-analytic approach. Although studies using the armchair approach toward criteria are still being published, most modern educational research uses either the rational-analytic approach, the empirical approach, or a combination of both. Ryans continues his description of the three approaches:

The gap between armchair approaches and a rational, or logical, approach (which may appear to be superficially similar to the armchair technique) is a large one. The rational-analytic approach plays an important part in every serious investigation involving a criterion and is centered in systematic observation and the logical analysis of the criterion behavior and its products, leading to an inclusive and exclusive designation of the components of the standard to be employed in making comparisons. Rational analysis is systematic and comprehensive. It aims to result in a description based on the relevancy of possible criterion components, judged from the standpoint of belongingness and representative sampling (1960).

The Committee on the Criteria of Teacher Effectiveness apparently used an approach similar to the rational-analytic in arriving at their hierarchy of criteria according to "ultimacy." First they selected what in their judgment was the ultimate criterion of teacher effectiveness: "It is assumed that the measure of teacher effectiveness is the change produced by the teacher in pupils. . . . Determining the effectiveness of teachers consists of finding the effects of teachers on pupils (Remmers et al., 1953)." Then they examined other possible criteria of teaching effectiveness, evaluating their relative importances, and ranked them on a continuum descending from the ultimate to the proximate. According to the committee, each lower criterion depends upon its functional relationship with the top criterion for its validity:

The ultimacy of criteria may be considered in terms of the "distance" of the behaviors or characteristics--of how far they are removed--from the ultimate criteria. Yet, all behaviors and characteristics of teachers can hardly be assumed to occupy merely different positions on the same dimension. They are not the same kinds of things (1953).

Thus, the criteria are in a logical sequence which comprehensively takes in many levels of complexity. Table II-2 presents the hierarchy, ranked in order of decreasing "ultimacy-proximacy," as it appeared in the first report of the Committee of the Criteria of Teacher Effectiveness.

TABLE II-2
 CRITERIA OF TEACHER EFFECTIVENESS

ULTIMATE CRITERION

Teachers' effect on:
 pupils' achievement and success in life
 pupils' achievement in subsequent schooling
 pupils' achievement of current educational objectives
 pupils' satisfaction with the teacher
 parents' satisfaction with the teacher
 superintendents' satisfaction with the teacher
 Teachers' "values" or evaluative attitudes
 Teachers' knowledge of educational psychology and mental hygiene
 Teachers' emotional and social adjustment
 Teachers' knowledge of methods of curriculum construction
 Teachers' knowledge of the subject matter
 Teachers' interest in the subject matter
 Teachers' grades in practice teaching courses
 Teachers' grades in education courses
 Teachers' intelligence

(Remmers et al., 1952)

Although the committee apparently used systematic and comprehensive rational analysis, as Ryans described the approach, its list of criteria leaves much to be desired from a researcher's viewpoint. The list is concise, but lacking in precision, so that before any of the criteria could be paired with potential correlates it would be necessary to define operationally each criterion. Because of this flaw, Cox presented a minority report along with the full report:

I am compelled to dissent from the report of this committee primarily because I do not believe this report brings us any nearer than we were to the solution of the practical problems which research workers face in this area.
 . . I think that the committee made an unfortunate choice early in its deliberations when it decided to define teaching effectiveness in conceptual terms rather than in operational terms (Remmers et al., 1953).

At least one other deficiency mars their ranking of criteria. The lower criteria on the list more nearly resemble potential correlates, whose "distances" from the ultimate criterion have yet to be empirically determined, than they resemble criteria whose relationships with the ultimate criterion can be determined by the judgments of a committee. Does "teachers' intelligence," for example, really lead to greater "pupil achievement and success in life"? If so, what are the bounds; that is, how much intelligence in a teacher leads to greater pupil achievement and success in life? These are questions that can best be answered by empirical research.

A later report of a similar nature also appears to have been arrived at by means of a rational-analytic approach, but it eliminated many of the shortcomings of the ultimacy report. The idea of preparing this list first appeared at the 1948 American Psychological Association Convention in Boston, when a group of college examiners expressed interest in a theoretical framework

which could be used to facilitate communication among examiners. After a number of informal meetings they agreed that

such a theoretical framework might best be obtained through a system of classifying the goals of the educational process, since educational objectives provide the basis for building curricula and tests and represent the starting point for much of our educational research (Bloom, 1956).

The culmination of this idea is the Taxonomy of Educational Objectives, Handbook I: Cognitive Domain (Bloom, 1956).

This taxonomy represents a very thorough and precise description of cognitive educational objectives, together with examples of suitable criteria to determine whether the objectives described have been achieved in any given situation. All cognitive skills are reduced to a few distinct groups:

As the taxonomy is now organized, it contains six major classes:

- 1:00 Knowledge
- 2:00 Comprehension
- 3:00 Application
- 4:00 Analysis
- 5:00 Synthesis
- 6:00 Evaluation

Although it is possible to conceive of these major classes in several different arrangements, the present one appears to us to represent something of the hierarchical order of the different classes of objectives. As we have defined them, the objectives in one class are likely to make use of and be built on the behaviors found in the preceding classes in this list (Bloom, 1956).

The committee states that pupil change is the central criterion around which they build the hierarchy of goals:

By educational objectives, we mean explicit formulations of the way in which students are expected to be changed by the educative process. That is, the ways in which they will change their thinking, their feelings, and their actions (Bloom, 1956).

Their guiding principles closely parallel Ryans' description of the rational-analytic approach. This report seems to represent the approach at its finest.

Empirical approach. The last of Ryans' three approaches to choosing a criterion is the empirical approach:

The empirical approach to the criterion is a pragmatic one which both follows and proceeds beyond rational-analytic efforts. It consists essentially of trying out hypothesized descriptions of the criterion, or dimensions composing the criterion, and accepting, modifying, or rejecting the criterion framework in the light of experience (Ryans, 1960).

This approach is exemplified by Ryans' own study, discussed in this chapter under Relevant Research.

Measures of criteria. Immediately upon encountering the empirical approach to criteria selection, a whole new group of difficulties appears. "Accepting, modifying, or rejecting the criterion framework in the light of experience" implies the necessity of methods of measuring the criteria in empirical situations, and it requires judgments about

which methods of criteria measurement are most appropriate to a given study. Ryans describes the nature of this new problem in the following words:

It is apparent that if decisions are to have the advantage of empirical evidence as well as rational support, data must be obtained through the use of selected criterion measures. Thus a circularity is introduced: criterion measures cannot be chosen until certain decisions have been made regarding the nature of the criterion; and some of the judgments relative to the criterion must remain tentative until reliable estimates of the hypothesized dimensions have been made. This is characteristic of all research, however, in that hypotheses are rationally derived, empirically tested, and then altered as a result of empirical evidence and further rational analysis.

So, criterion definition and criterion measurement interact and, in a sense, all conclusions about a criterion are relative to the type of measurement or observation employed. Therefore, in discussing the composition of the criterion, particularly from the standpoints of dimensionality, weighting, and generalizability, we must recognize that the terms "criterion" and "criterion measure" cannot be entirely independent (Ryans, 1960).

Mitzel feels that the relationship between the criterion and criterion measure is so close that he says, ". . . teacher effectiveness as a concept has no meaning apart from the criterion measures or operational definitions of success as a teacher (Mitzel, 1960)."

Any criteria chosen for use in research must be operationally defined to prevent ambiguity. Co-researchers cannot be sure they are talking about the same concept, even though they use identical words, unless each concept

being used in an investigation is given a working definition. However, one might wonder if some essential relationships are lost in the translation of broad educational objectives into the criterion measures (operational definitions) used in research. Mitzel does not seem to think this presents a problem:

The ultimate goals are really the goals of society at large; for practical purposes, the unique contribution that can be made by schools and teachers is impossible to isolate. It is therefore suggested that product criteria should consist primarily of measures of the goal-oriented effects of teachers (and schools) on children, obtained immediately after the periods in which children attend the schools (Mitzel, 1960).

The Committee on the Criteria of Teacher Effectiveness arrived at a similar conclusion a few years earlier:

Realistically, it seems necessary to assume that changes at the time the pupil is under the teacher's influence are sufficient to serve at least as a first approximation in evaluating teacher effectiveness (Remmers et al., 1953).

Somewhere in the criterion-selection process the final choice of a criterion measure passes beyond the limits of purely empirical processes, so that recourse to subjective judgment must once again be made, even as the initial selection of a criterion was ultimately a subjective judgment. To aid researchers in making necessary judgments, Mitzel describes four basic attributes that all criterion measures should possess: relevance, reliability, freedom from bias;

and practicality. Selected excerpts from his discussion are reproduced below:

Relevance as a criterion attribute is the product of a rational analysis of the job functions and the job objectives. In so far as a criterion measure reflects the behaviors required in the achievement of job objectives, it is relevant.

Relevance, whether direct or indirect, is the paramount attribute of any criterion measure. It overshadows consideration of the other desirable characteristics: reliability, freedom from bias, and practicality.

Reliability as an attribute of a criterion measure--like reliability as a characteristic of a test--is necessary but not sufficient.

Freedom from bias in a criterion measure is always a desirable attribute, even though elimination of bias may lower reliability. . . Sources of potential criterion bias include: school differences, grade differences; and student differences in initial ability, interest, and motivation.

Practicality is related to convenience and economy, but it is probably the least essential of the four criterion attributes (Mitzel, 1960).

Although the final choice of criterion measures must be left to the researcher's judgment, the above guidelines can help him prevent poor judgment either because of oversight, or through overemphasis of one of the attributes.

Three classes of criteria. The emphasis to this point has been upon pupil change as the ultimate criterion, because of its acceptance among authorities:

Determining the predictors of the ultimate criteria (changes in pupil behavior) is the essence of the problem of determining teacher effectiveness (Remmers et al, 1953).

However, in his article in the Encyclopedia of Educational Research quoted earlier, Mitzel lists three different classes of criteria grouped according to goal-proximity. These classes are named product criteria, process criteria, and presage criteria. Criteria of pupil change belong to the class of product criteria, because pupil changes are products of teacher behaviors.

Whether a particular operationally defined measure is or is not a product criterion depends on the answer to the question, "Is student change in this trait or characteristic an appropriate goal for our schools?" (Mitzel, 1960).

Ryans, who uses process criteria in his study, Characteristics of Teachers, feels that product criteria have many disadvantages outweighing their advantages for use in research:

The chief disadvantage in the use of products as criterion measures is the difficulty of adequately controlling external factors in order to provide reasonable assurance that the hypothesized product is truly a product of the criterion behavior rather than that of a wide range of uncontrolled conditions occurring before and during the criterion behavior (Ryans, 1960).

Another difficulty with the use of product criteria is the effect of time separation:

An additional consideration in the use of products for the purpose of criterion measurement involves the proximity in time of the product evaluation to the occurrence of the behavior which is hypothesized to be the producer. Frequently the passage of time and

the consequent introduction of additional influences have considerable effect upon a product . . . What really is being measured in each case is a product once removed from its hypothesized producer (a product of pupil behavior is measured and taken as a measure of a product of teacher behavior) (Ryans, 1960).

Nevertheless, Ryans gives qualified approval to the basic idea of product criteria:

However, if the rationale of the product (pupil performance) criterion is accepted, and if the complex control problem presented by a multiplicity of producers and the multidimensionality of the criterion can be satisfactorily handled, pupil change becomes an intriguing approach to the measurement of teacher behavior criteria (Ryans, 1960).

The second class, following product criteria, is called process criteria. "Process criteria consist of those aspects of teacher and student behavior which are believed to be worthwhile in their own right (Mitzel, 1960)." Typically, according to Mitzel, these will be conditions, climates, or typical situations involving the social interactions of students and teacher. Ryans gives his reasons for preferring process criteria to product criteria in the following paragraph:

One serious impediment to the understanding of teacher behavior has been the frequent failure of researchers to recognize that before progress can be made in any field of science there must be a long period of investigation devoted to (a) the identification of the elements or components that make up the phenomena under consideration and (b) the study of their interrelationship--the descriptive

phase of research. For many years, there appeared to be little recognition that study and assessment of teacher effectiveness involved the adoption of a value system and that value judgments necessarily were relative in nature. The estimation of teacher effectiveness must be preceded by (a) description of teacher behaviors, (b) subjective judgment and choice of values and associated teacher behaviors that are acceptable to a particular group as manifestation of teacher effectiveness, and (c) well-designed study of the relationships between observable teacher characteristics and teacher behaviors on the one hand and the accepted evaluative criteria on the other (1960).

Ryans indicates that he does not disagree with the idea that pupil gain ("the accepted evaluative criteria") is the ultimate criterion; rather he feels that the present state of educational research is not advanced enough so that researchers can move directly to the determination of the effectiveness of teacher behaviors from pupil behaviors:

It seems probable that, without losing sight of the importance of developing means of recognizing "good" teachers, attention of the researcher might first more properly and profitably be directed at the identification and estimation of some of the major patterns of personal and social characteristics of teachers (1960).

Ryans acts according to his views, since in his own study he directs his efforts toward "identification and estimation" of characteristics of teachers:

It is with problems relating to the description of teacher characteristics that the research reported in this volume was primarily concerned. In much of

the research of the Teacher Characteristics Study, consideration of the effectiveness or ineffectiveness of particular behaviors was intentionally set aside (1960).

When Ryans does include estimations of teacher effectiveness, as he does in the section entitled, "A Comparison of Teachers Assessed as Generally High and Generally Low," he never proceeds to a direct comparison of teacher behavior and pupil behavior, but instead uses as criteria "teacher behaviors that are acceptable to a particular group as manifestations of teacher effectiveness (1960)."

Mitzel, citing Ebel, suggests that claims for the superiority of process over product criteria may be merely defense mechanisms on the part of the claimants:

It is a hedge against the possible finding that a new learning technique does not yield a better educational product than some old technique employed for the same purpose. If certain defineable and observable educational means are clearly better than others, then their effects should be discoverable in measured educational ends (Mitzel, 1960).

Following product and process criteria is the class named presage criteria. From a logical standpoint, presage criteria are completely removed from the goals of education. "In a sense they are pseudo criteria, for their relevance depends upon an assumed or conjectured relationship to other criteria, either process or product (Mitzel, 1960)." Mitzel mentions four types of presage variables in common

use as criteria of teacher effectiveness: teacher personality attributes, characteristics of teachers in training, teacher knowledge and achievement, and inservice teacher status characteristics. He suggests that some presage variables might best be presented as potential predictors of process or product criteria, rather than as criteria themselves. To all appearances, the lower criteria listed in the "Ultimacy" report (Table II-2) are presage variables.

Conclusion. The discussion above may suggest the complexity of the criterion problem. It cannot be ignored, since the value of an entire research undertaking rests upon the fundamental criteria employed. "Calling a particular measure a criterion lends to it connotations of worth and value. Criteria cannot be trivial; otherwise evaluations are made against trivial standards (Mitzel, 1960)." Even if the problem of criteria is given careful consideration, using all available facts, there are many compromises to be made due to practical necessity. No available criterion will solve all of the researcher's difficulties. Mitzel makes a succinct generalization about proposed criteria:

Perhaps the most significant statement that can be made about any proposed criterion variable is that it is a partial criterion. Relevance for a particular criterion should be defended logically without undue

apology for its lack of comprehensiveness, since it seems unlikely that we shall soon find a generally accepted formula for weighting school goals according to their importance for over-all effectiveness (1960).

The present study has adopted a criteria of pupil change (product criteria), using achievement tests constructed to sample all of the broadly agreed-upon areas of academic content in which teachers are expected to produce pupil change.

Potential Correlates

Central variables. After the criterion is chosen, research is almost ready to begin in earnest. However, one more decision must be made before actually beginning: from a broad field of potential correlates the researcher must select one or more central variables that might possess relationships with his chosen criterion. The central variable, defined in the introduction to this chapter, can belong to any of three possible categories of variables according to the Handbook of Research on Teaching: first, teaching methods; second, instruments and media of teaching; and third, the teacher's personality and characteristics. "For an investigation to be 'research on teaching' by our definition, it is necessary and sufficient that it deal with such a central variable (Gage, 1963)."

Teaching methods are often considered to play a central part in educational processes. When studying teaching methods as central variables the individual differences among teachers are temporarily overlooked; the teacher is viewed as acting according to a method, or role, such as "lecturing" or the "project method." Instruments and media of teaching can be considered central variables because "writers of textbooks, producers of films, programmers of teaching machines, or makers of reinforcement schedules all exhibit teaching behavior, no less so because it is 'fixed' on a page, film or tape. . . Accordingly, research on teaching must deal with such instruments and media (Gage, 1963)." The third central variable includes both teacher behavior and teacher characteristics. "When the teacher's behavior is considered to be a reflection of his personality, research attention centers on individual differences among teachers in ability, knowledge, attitude, temperament, and the like (Gage, 1963)." Teacher characteristics, also included as part of this central variable by the Handbook, are nonbehavioral properties such as teacher's age, sex, social class, or years of experience.

Noncentral variables. "Although a piece of research on teaching must by definition be concerned with one such

central variable, it will usually also be concerned with noncentral variables, i.e., with variables we shall term 'relevant' or 'site,' to be defined below (Gage, 1963)."

Relevant variables can be grouped into two classes, according to the Handbook, either as social interaction in the classroom or as social background of teaching. "Interactive variables are those characterizing not the individuals related to each other in teaching but the 'field,' or relationship, between them (Gage, 1963)." "Other relevant variables take account of the social background provided by the teaching and administrative staff off the school and school system, by the homes from which learners come, and by the community and society at large (Gage, 1963)."

Site variables are another group of noncentral variables considered by educational researchers. They are usually held constant so as to characterize the site in which other variables are studied. The two main site variables are grade level and subject matter.

The three classes of potential correlates described by the Handbook provide a convenient frame of reference against which to view research variables pertinent to the present study.

Two relevant potential correlates. Two types of studies relate to the present investigation. The first examines the role of teacher age in influencing teacher behavior; in these studies age is a central variable in the sense defined by Gage. Describing a school, Peterson says:

It is . . . a social system in which individuals develop, mature, age, and perform with various degrees of effectiveness at various career points and at various points of change in their social worlds (1964).

Despite its importance to educational researchers, age has frequently been ignored as a potential correlate of teaching effectiveness.

The second type of study contains potential correlates that are noncentral, namely, the social background of teaching. In this context, the word "background" has a meaning similar to "ground" in the figure-ground concept from psychology, rather than its customary meaning of "historical" or "chronological" background. The social background might be described as an atmosphere or climate against which the teaching-learning processes are viewed. "Social environment" could be another name for this group of potential correlates.

Charters describes the importance of social background to the education process in his chapter in the Handbook:

The most immediate environment for the teacher and the teaching-learning process is, of course, the school and the system of social relationships of which it consists. . . . The behavior of a particular teacher, caught up in the system of interlocking roles, is bounded by the fact that he is dependent upon the role performance of other school personnel and they are dependent upon his. . . . The teacher's involvement in this system of relationships has important consequences for him and for his conduct of the teaching-learning process (Gage, 1963).

Organizational climate. In order to work with social background as a research variable it is necessary to define it operationally. One research project has accomplished that objective. In September of 1959 Halpin and Croft, under a contract from the United States Office of Education, Department of Health, Education and Welfare, began an investigation intended to

map the domain of organizational climate, to identify and describe its dimensions, and to measure them in a dependable way which would minimize those limitations that necessarily inhere in every instrument which must, in the final instance, rely upon some form of subjective judgment (Halpin, 1966).

The organizational climate that these researchers were trying to map out was different from the traditional concept of morale:

One obvious approach to the domain of Organizational Climate is the attempt to encapsulate everything important to be said about the climate within the single global concept of morale. With this approach the best we can hope to do is to estimate how high or how low the morale of a given organization is. . . . The difficulty, of course, is that this approach rests upon the

a priori assumption that a single dimension--that is, morale, can usefully summarize the essence of the variations that occur in organizational climates . . . But the assumption of this research is untenable, for research on morale has yielded, above all, one unequivocal finding: morale, whatever it may or may not be, is not unidimensional in its structure (Halpin, 1966).

Another traditional approach discarded by Halpin and Croft used many names to describe a school's social background:

The investigator who uses this tactic proceeds to generate elaborate lists of adjectives which presumably can be used to describe a host of possible aspects of organizational climates. . . Obviously, these adjectives do describe the climate, but since they have been generated ad hoc, without an organizing principle, the investigator can group the behaviors to which the adjectives refer on only an arbitrary basis (Halpin, 1966).

The approach they used tried to accomplish goals similar to those of the two approaches above, but in a more rigorous, objective way:

We were seeking to map the same domain of inquiry that other investigators have described as morale, but we were seeking to conceptualize--or, if you will, to map--this domain in a different way. . . We wanted first to observe the behavior that defined different Organizational Climates and were willing to name these climates only after we had analyzed the specific behavior. In short, we were committed to an inductive, empirical approach (Halpin, 1966).

Specifically, they constructed an Organizational Climate Description Questionnaire (OCDQ) that permitted them to portray the Organizational Climate of an elementary school. The OCDQ is made up of sixty-four Likert-type items through

which teachers and principals could describe the climate of their school. The questionnaire, requiring less than thirty minutes to administer, was given in group situations. Halpin and Croft used this instrument to analyze the climate of seventy-one elementary schools chosen from six different regions of the United States, involving 1151 respondents. The sixty-four items in the OCDQ were assigned to eight subtests which were delineated by factor-analytic methods; four of the subtests applied to characteristics of the faculty as a group, the other four to characteristics of the principal as a leader. A profile depicting each school's Organizational Climate was then constructed from the eight subtest scores. By comparing the profiles of different schools with each other, the distinguishing features of their respective Organizational Climates were found; by means of this comparison, six Organizational Climates were identified and arrayed along a continuum defined at one end by an Open climate and at the other by a Closed climate.

At the completion of the research described above, Halpin and Croft felt they had succeeded in achieving their original objective:

In social science research, findings are seldom as tidy as the investigators would like them to be. Our situation provides no exception, yet our objective has been attained. We have found that it is feasible to dimensionalize the behaviors which define

the Organizational Climate of elementary schools, and we have identified empirically six distinct Organizational Climates which make good sense, both practically and psychologically (Halpin, 1966).

The original objective of Halpin and Croft was to "identify and describe" occupational climates, not to evaluate them. The names chosen to describe the climates, however, are evaluative: "Open" has good connotations, whereas "Closed" carries with it undesirable connotations. Halpin admits that their decision to choose evaluative names was inescapable in light of their findings:

At the outset, Croft and I thought that we might be able to stick to describing climates without evaluating them. That was an innocent thought. The more we worked with the findings, the more did judgments about the climates force themselves upon our attention. The difference in the quality of the different climates became too vivid and too compelling to be ignored. In short, we were forced to admit that the Open Climates were "the good guys" while the Closed Climates were "the bad guys" (Halpin, 1966).

"Good" and "bad" imply that some standard or criterion was used to judge the climates, but such was not the case, for no formal criteria were built into the research structure: ". . . at this point we have not been concerned about the relationship between the profile scores on the OCDQ and the external criteria of a school's effectiveness (Halpin, 1966)." Thus, to this point the OCDQ was primarily a tool for describing the social background of schools, corresponding to the descriptive phase of research in the quotation by

Ryans on page 25 of this chapter. The next step, after identifying and describing the dimensions being studied, consisted of selecting criteria and determining empirically if those newfound dimensions are predictors of the criteria. This was the task of the present study.

Relevant Research

Teacher age. Two types of research studies relate to the present investigation. The first type deals with teacher age as a central variable. At least two studies have reported significant differences in teaching behavior between age groups.

The first study, the Characteristics of Teachers (Ryans, 1960), is considered to be the most extensive investigation of teachers to date. The study took six years and was made up of over 100 separate research projects, involving more than 6,000 teachers in 1,700 schools. The purposes of the study were threefold: to identify and analyze some patterns of classroom behavior, attitudes, viewpoints, and intellectual and emotional qualities which may characterize teachers; to develop instruments to predict certain patterns of classroom behavior and personal qualities of teachers; and to compare various groups of teachers. Through (1) a review of the literature, (2) an assembly of reports of

classroom "critical incidents," (3) the assessment of teacher classroom behavior, and (4) statistical analysis of the teacher assessments, three principal criteria were chosen for determining correlates:

- TCS Pattern X_o : understanding, friendly vs. aloof, egocentric, restricted teacher behavior
- TCS Pattern Y_o : responsible, businesslike, systematic vs. evading, unplanned, slipshod teacher behavior
- TCS Pattern Z_o : stimulating, imaginative, surgent or enthusiastic vs. dull, routine teacher behavior

(Ryans, 1960)

There were two substudies within Ryans' overall effort for which age differences in teachers were reported. The first was named the Basic Analysis Study and involved 3,883 elementary and secondary teachers from 377 schools in 33 school systems. The second was called the Survey Study, and included 1,640 teachers from 962 schools in 271 school systems. The former sample was directly observed, but for practical reasons could not be a representative national sample; the latter sample was questioned by mail, and responses from all states were received for each of the major teaching fields.

For each teacher included in the Basic Analysis Study, trained observers completed an assessment blank (The Classroom Observation Record) constructed for this

study. The observer ranked the teacher on a seven-point scale for eighteen pairs of teacher behaviors:

Partial	Fair
Autocratic	Democratic
Aloof	Responsive
Restricted	Understanding
Harsh	Kindly
Dull	Stimulating
Stereotyped	Original
Apathetic	Alert
Unimpressive	Attractive
Evading	Responsible
Erratic	Steady
Excitable	Poised
Uncertain	Confident
Disorganized	Systematic
Inflexible	Adaptable
Pessimistic	Optimistic
Immature	Integrated
Narrow	Broad

(Ryans, 1960)

One of the findings of the study was that teachers over 55 years old tended to receive the lowest assessments, with F-ratios significant at the .05 level for the secondary teacher sample, although not for the elementary teacher sample.

For the Survey Study a special paper-and-pencil instrument (Teacher Characteristics Schedule) was devised. Some of these Schedules were sent to each principal of a stratified sample of 2,800 principals with instructions to pass them on to randomly selected teachers in their systems. A total of 1,640 usable schedules were returned, 670 from elementary teachers and 970 from secondary teachers, closely

approximating the geographic distribution of the United States as a whole. The Teacher Characteristics Schedule, completed by the teachers themselves, is intended to be an indirect estimation of "classroom behaviors, attitudes, educational viewpoints, verbal intelligence, and emotional adjustment of teachers from their responses to multiple-choice and check-list items relating to preferences, judgments, activities, family and home backgrounds, and the like (Ryans, 1960)." In its final form the Schedule consisted of 350 items. Means and standard deviations were computed for ten schedule variables according to age groups, and 45 out of 60 sets of differences between the means of age groups were found significant at the .05 level.

There appears to be little doubt about the existence of significant differences between teachers comprising different age groups. . . Generally speaking, scores of older teachers (55 years and above) showed this group to be at a disadvantage compared with younger teachers . . . (Ryans, 1960).

As a result of the findings from the Basic Analysis and Survey Studies, Ryans draws this conclusion:

The implication for research and practice in the area of teacher personnel probably is that age must be taken into account as a relevant independent variable when teacher characteristics are considered-- that personality-wise, teacher variation with age likely interacts with the main effects sometimes described as contributors to teaching performance, and may tend either to make differences in such main effects appear important when they do not exist, or, on the other hand, to obscure main effects (1960).

Peterson, using a different technique from that used by Ryans, completed another study using age as a central variable. Peterson describes his methods in these words:

Techniques were qualitative rather than quantitative and concerned social patterns and processes rather than prevalence or reliability. The validity of the data rested on the honesty and competence of teachers as observers of their own role in a social system. The objectivity of the results depended upon the responsibility and degree of insight possessed by the interpreter (1964).

His sample was confined to white, female high school teachers thirty to seventy years of age. Seventy-two teachers were drawn at random from an annual directory of personnel in an urban school system, of which fifty-six completed the interview. The data-gathering technique used was an informal, conversational interview which was partially structured according to topic but largely unstructured in the wording and order of questions. The interviews were conducted in the respondents' homes, usually lasted two hours, and were recorded on tape unless the respondent objected.

After carefully analyzing the contents of the recorded interviews, Peterson described some general findings:

Teachers seem to recognize a short orientation period, a 'best teaching period,' while still young and vigorous, and a decline beginning at the age of thirty-five or forty. Almost without exception, teachers who were in their thirties at the time of interviewing expressed concern about losing some kind of intimate, informal contact with students. In comparison, middle-aged teachers seemed secure, relaxed,

and appeared to have accepted increased distance from students--having established for themselves a parentlike authority. Old teachers seemed frustrated and tended to complain about students--'getting worse every year.' They vacillated between two themes: 'Teaching is for the young and vigorous,' and 'The younger generation is undisciplined, uninterested, morally decadent'(1964).

He stresses that exceptions were found to this picture, but the pattern generally holds true as teachers progress in their career. The findings of the study prompted Peterson to make this conclusion about age as a variable in the measurement of teacher effectiveness:

One practical implication should be stressed, however. Previous views of teacher role have seen the teacher in ageless abstraction. Exhaustive discussions of teacher effectiveness stress methods, sensitivity to students, and control of the classroom. Studies of teacher-pupil interaction usually differentiate among teachers in terms of subject matter, grade, or other criteria functional to school tasks. It is the contention here that such views, recommendations, and studies ignore the important facts of aging, commitment, and job satisfaction (1964).

Organizational Climate. In addition to research using teacher age as a potential correlate of the criterion, studies using social background as potential correlates are pertinent to the present investigation. At least three researchers have investigated social background using the OCDQ as the measure. In the development of the OCDQ, Halpin and Croft concluded their efforts with the identification and description of school climates; the next became to

determine whether the findings correlated with accepted criteria. The studies reported here attempted to accomplish this, using pupil achievement as their criterion.

The first of these researchers, Andrews (1965), performed correlations between an index called Pupil Achievement and each of the OCDQ subtest and climate scores. The Pupil Achievement index was calculated by finding the means for ninety-five Alberta, Canada, schools on the Grade IX Departmental Examinations and removing the effect of pupil academic ability through statistical means. Using the Pearson Product Moment Correlation technique only one of the correlations was significant at the .01 level, that of the Pupil Achievement index with the Intimacy subtest score. This led him to conclude that "schools that produce high examination results, then, tend to be characterized by strong social relationships among teachers (Andrews, 1965)." This conclusion is questionable, however, since in forming it Andrews capitalizes on chance. There were eight subtest scores and six climate similarity scores, yielding fourteen correlations. Each correlation was tested at the .01 level, so the level of significance for the entire group of correlations is fourteen times .01, or .14; thus, fourteen out of one-hundred, or approximately one out of seven, can be expected to be significant at the .01 level due to chance alone. It

is improbable that the difference in achievement was due to actual differences in the sample of schools tested.

Another researcher, Millar (1965), performed correlations between the Grade IX Departmental Examinations of eight Alberta, Canada, urban schools attended by pupils of similar socioeconomic status, and found that both Intimacy and Aloofness subtests correlated with pupil achievement. The Intimacy subtest correlated .804 with pupil achievement, supporting Andrews' conclusions; the Aloofness subtest correlated $-.821$ with pupil achievement. Millar, like Andrews, found little relationship between the global concept of Organizational Climate and pupil achievement. Millar's findings are weak because he was working with a sample of only eight schools, so he provides little support for Andrews even though duplicating the statistical significance for the Intimacy subtest.

The third researcher, Feldvebel (1964), used the Grade V Stanford Achievement Battery as his criterion measure. In a sample of 30 schools in the suburban Chicago area, he obtained correlations of $-.399$ and $.391$ respectively, between achievement and the Production Emphasis and Consideration subtest scores of the OCDQ. As with the other two studies, Feldvebel found no statistically significant relations using the climate scores.

Much research needs to be done to determine whether the dimensions measured by the OCDQ can be used as valid predictors of criteria. None of the studies above conclusively demonstrates or refutes relationships, so the need for further research is clear. The present investigation attempts to overcome the inadequacies of these studies by using a larger sample, by examining correlations by grade level and subject area, and by statistically subtracting the results of student intelligence differences.

CHAPTER III

EXPERIMENTAL DESIGN

Instruments

Stanford Achievement Battery. The most direct means of evaluating the relationship of teacher age, Organizational Climate, and school effectiveness in this study appeared to be through a criterion of pupil achievement. Once this criterion was selected, the problem became that of choosing an acceptable measure of student achievement. This measure had to (1) reflect the subject matter taught in schools; (2) provide comparable scores from school to school; (3) provide comparable scores from grade to grade; (4) yield reliable scores; (5) be as short as possible commensurate with complete subject area sampling; (6) contain up-to-date norms and items; and (7) be machine scorable. The Stanford Achievement Battery met all these requirements and was chosen as the criterion measure.

Table III-1 below shows the level of test battery used in each grade for the achievement test.

TABLE III-1
LEVEL OF TEST BATTERY USED IN EACH GRADE

Grade	Stanford Battery
4	Intermediate I
6	Intermediate II
9	Advanced
11	High School

Achievement battery subtests administered to each grade level are presented in Table III-2. More detailed information about the Stanford Achievement Batteries can be obtained from Kelly, Madden, Gardner & Rudman (1966); Gardner, Merwin, Callis & Madden (1965); Buros (1965); and, with regard to the present project, Egelston et al. (1967).

Lorge-Thorndike Intelligence Test. In order to insure that school differences in achievement level were due to the independent variables being examined, not to the extraneous variable of student intelligence, some measure of school differences in student intelligence level was necessary. For reasons similar to those guiding the selection of the Stanford Achievement Battery, the 1964 Multi-Level Edition of the Lorge-Thorndike Intelligence Test

TABLE III-2
SAT SUBTESTS USED IN EACH BATTERY

Intermediate I	Intermediate II	Advanced	High School
Word Meaning	Word Meaning		
Paragraph Meaning	Paragraph Meaning	Paragraph Meaning	Reading
Spelling	Spelling	Spelling	Spelling
Word Study Skills [#]			
Language*	Language*	Language*	English
Arithmetic Computation	Arithmetic Computation	Arithmetic Computation	
Arithmetic Concepts	Arithmetic Concepts	Arithmetic Concepts	Numerical Competence
Arithmetic Applications	Arithmetic Applications	Arithmetic Applications	
Social Studies**	Social Studies**	Social Studies**	Social Studies*
Science	Science	Science	Science A
			Mathematics A

* These subtests were administered, but because of technical data-processing problems they were not available for this particular analysis.

** The Social Studies subtest consisted of two parts, Content and Study Skills, which were treated as individual subtests in this study.

[#] The Word Study Skills subtest consisted of two parts, Phonics and Syllabication, which were treated as individual subtests in this study.

was selected as the intelligence measure. The test manual provides this description of the test:

The Lorge-Thorndike Intelligence Tests are a series of tests of abstract intelligence. Abstract intelligence is defined as the ability to work with ideas and the relationships among ideas. The tests are based on the premise that most abstract ideas with which the school child or the working adult deals are expressed in verbal symbols, so much so that verbal symbols are the appropriate medium for the testing of abstract intelligence. Nevertheless, they take account of the fact that for some--the young, the poorly educated, or the poor reader--printed words may constitute an inadequate basis for appraising an individual's abilities. Consequently, a parallel set of nonverbal tests is provided to accompany the basic verbal series (Lorge et al., 1964).

The Multi-Level Edition provides both Verbal and Nonverbal Batteries for grades three through thirteen in a single booklet. The items in each battery are divided into eight different but overlapping scales for use in each separate grade. The Verbal Battery is made up of five subtests which use only verbal items: vocabulary, verbal classification, sentence completion, arithmetic reasoning, and verbal analogy. The Nonverbal Battery uses items which are either pictorial or numerical. It contains three subtests involving pictorial classification, pictorial analogy, and numerical relationships.

Table III-3 shows the level of the Lorge-Thorndike Test used in each grade.

TABLE III-3
LEVEL OF TEST BATTERY USED IN EACH GRADE

Grade Level	Lorge-Thorndike Battery
4	B
6	D
9	F
11	G

More detailed information about the Lorge-Thorndike can be obtained from Lorge, Thorndike & Hagen (1966); Buros (1965); and, with regard to the present project, Egelston et al. (1967).

Organizational Climate Description Questionnaire.

The OCDQ, described briefly in Chapter II and presented in Appendix A, questions teachers and principals about the frequency of occurrence of certain personnel behaviors in their schools. Through analysis, Halpin and Croft determined that eight different dimensions of school climate were being measured by their instrument. On the basis of these dimensions, they found it was possible to group schools into six categories according to school climate: Open, Autonomous, Controlled, Familiar, Paternal, and Closed. Descriptions of the eight dimensions and six climates are

given in Appendices B and D, respectively. Briefly, the OCDQ is a quick measure of eight different dimensions including, yet surpassing, the traditional concept of morale.

Because of the uniqueness of the OCDQ and the limited availability of technical information about the instrument, a short description of its development will be included to supplement the description above.

Development of the OCDQ. After first choosing a Likert-type questionnaire, Halpin and Croft selected this four-response frequency of occurrence scale:

1. Rarely occurs
2. Sometimes occurs
3. Often occurs
4. Very frequently occurs

Then they constructed about 1,000 items intended to gather information about teachers' and principal's activities, screening the items by administering three preliminary forms of the OCDQ. In successive parings the number of items was trimmed from 1,000 down to 80, and a later analysis showed that the number could be further lowered to 64 items, its present size, without significant loss of information. Through factor analysis they identified eight dimensions of Organizational Climate, represented by eight subtests; by inspection, four were found to relate to the behavior of

teachers, the other four to the behavior of principals. The eight dimensions are named in Table III-4.

TABLE III-4
OCDQ SUBTESTS

Relating to Teacher's Behavior	Relating to Principal's Behavior
1. Disengagement	5. Aloofness
2. Hindrance	6. Production Emphasis
3. Esprit	7. Thrust
4. Intimacy	8. Consideration

(Halpin, 1966)

Halpin's description of each of the eight subtests can be found in Appendix B, and the items comprising each subtest are grouped in Appendix C.

After initial analysis was completed at the item level, analysis was begun at the subtest level. Each of the eight subtest scores for every respondent (from a sample of 1,151 respondents in 71 schools) was calculated by averaging the subtest item scores; then the means were converted to standard scores. A factor analysis was performed on these scores, yielding three predominant factors which Halpin named (1) Social Needs, an individual factor; (2) Esprit, a group factor; and (3) Social Control, a leader factor.

To this point the analysis had been conducted at the item and subtest level for individuals, but since the major purpose of the study was to describe the Organizational Climate of schools, further analysis was performed at the school level. Four more steps were required to complete the basic analysis. The first step was to construct profiles made up of the eight subtest scores for each school. The second step was to factor-analyze the 71 school profiles, which yielded three profile factors and six patterns of factor loading. Each school profile was then categorized with respect to one of the six patterns of loadings. In the third step, Halpin and Croft constructed a representative profile for each of the six patterns, using mean scores. These "average-profiles" were then designated as prototypic profiles, from which the six Organizational Climates were defined. The fourth and last step in the analysis was to rank the six Organizational Climates in respect to Openness versus Closedness. This was done by examining the content of subtest items in order to discover the behavior characterizing the principals and teachers for each climate. The six climates, in order, were named Open, Autonomous, Controlled, Familiar, Paternal, and Closed. Climate descriptions are given in Appendix D.

When the basic development of the OCDQ was completed, each school taking the questionnaire was categorized according to the prototypic profile it most resembled. This was accomplished by computing the absolute difference between each subtest score in a school's profile and the corresponding score in the first prototypic profile, then in the second prototypic profile, the third, and so on. For each comparison with a prototypic profile the absolute differences were added together to form a climate similarity score. Thus, for each school the lowest climate similarity score (the least sum of differences) determined which prototypic profile it most resembled, hence its climate. For example, if the lowest of a school's six climate similarity scores were the Openness score, it was classified as having an Open climate. Halpin and Croft set the cutoff score at forty-five (Croft, undated); that is, if the lowest similarity score obtained by a school was larger than the cutoff score, then the school did not clearly resemble any climate. There is an apparent contradiction between this recommendation and the method used to categorize schools in Halpin's report (1966), where no cutoff score was used at all.

Data Collection

The instruments chosen were designed and standardized to be administered by classroom teachers, and each school was asked to administer its own tests. Special training sessions and instruction sheets were used to inform teachers about the school and student identification required by the UPMDPP, and to insure that the machine scorable answer sheets would be coded properly. Lorge-Thorndike tests were delivered to the schools first, and administered by most of the schools in early November of 1966. The Stanford tests were delivered next, and administered by most of the schools in December of 1966. The OCDQ booklets were delivered in March, 1967, and completed answer sheets continued arriving until the end of the school year.

With each succeeding instrument fewer schools were willing to continue participation in the investigation. Of the original 241 schools administering the Lorge-Thorndike, only 113 administered the OCDQ, representing 1253 usable respondents out of an estimated total population of 3,500 teachers. This sample is nevertheless larger than that used by Halpin and Croft to develop their instrument.

Scoring

Because of the large number of students and teachers participating in this investigation, machine scoring of answer sheets was essential. All answer sheets used for this study were designed to be scored with the Digitek Optical Scanner, Model 100, leased for the duration of the project. Answer sheets for the Stanford Achievement Batteries and the Lorge-Thorndike Intelligence Test were commercially available from the test publishers, but the answer sheet used with the OCDQ had to be designed by the author and custom printed. Many of the items on the OCDQ might have been answered differently if the respondent thought his identity could become known to his principal, so no identification was required on the OCDQ answer sheets other than school code, birth year, and sex. Each Digitek answer sheet punched one output card at scoring time. The tests administered used from two to five sides per test, and following scoring the results from each test were combined on a single card for ease of handling.

Scoring the OCDQ was not a matter of simply adding up correct responses, as was the case for the achievement and intelligence tests. For such a sizable number of schools, the required scoring procedure (described briefly in the Instrument section, above) would be too difficult to compute

by hand, so Halpin and Croft provided a computer program which automatically calculated both the subtest scores and the climate similarity scores. The storage demands of this program exceeded that available in the IBM 1620 computer used by the project, and most of the program's optional features had to be written out before the computer could accommodate it. Even after the investigator finished modifying the program, the small computer system created another difficulty: initial calculations made after a trial run showed that the computer time for processing the 1253 respondents from 113 schools would be forty-three hours, and the actual run took somewhat longer because of excessive missing data in several schools. To correct the problem of missing data (that is, respondents omitting an excessive number of items), about fifteen OCDQ answer cards had to be removed from the rest of the data.

Using the birth year data from the OCDQ answer sheets, a specially written computer program calculated each teacher's age, grouped them in five-year categories from below 20 years to over 85 years, then printed and punched a distribution of teachers' ages for each school. The age array output cards were then used as data for another custom program which computed for individual schools the total personnel, the number

of teachers under 35 and over 55, and the percentage of teachers under 35 and over 55.

At this point all data necessary to test the three hypotheses were on punched cards, ready for use with the appropriate statistical programs.

Statistical Analyses

Lorge-Thorndike intelligence scores were used as the covariate in a one-way analysis of covariance on the Stanford Achievement Test raw scores. The hypotheses tested differences between schools, so each individual school comprised a level of the one-way design. Computations were carried out using the CANOV 1 computer program written by Paul Lohnes, State University of New York at Buffalo, and run on the project's IBM 1620 computer. Output from this program included an analysis of variance table for the criterion variable, together with group means, standard deviations and variances; analysis of variance table for the control variable, including group means, standard deviations, and variances; pooled within regression weights for all groups; and an analysis of covariance table for the criterion variable, together with adjusted school means.

Significant differences from the covariance analysis could only show that two (or more) of the school adjusted means differed from one another significantly, but could not

reveal which two (or more). In order to test the hypotheses, select groups of schools still had to be compared statistically after it was determined there was an overall significant F-ratio. One common method for doing this is to perform t-tests between groups of scores. Since the adjusted means were to be used for this analysis, and adjusted individual student scores were difficult to obtain with the available resources, a non-parametric counterpart of the t-test was used to test the specific hypotheses. This test, the Mann-Whitney U test (Siegal, 1956), was selected because of its high power-efficiency of 95.5% when substituted for a t-test with large samples. Computations for this test were carried out by computer program Mann-Whit written by the investigator and run on an IBM 360/67 computer at the University of Michigan. Output included the value of U; its Z score and exact probability for large samples; the sum of ranks for each group (to indicate the direction of significance); and the number of tied scores and the clusters they formed. A 0.05 level of significance was used for all tests, and because the hypotheses were directional, all significances reported are one-tailed.

Since the hypotheses formulated about the OCDQ were very selective, ignoring much of the information available

from the eight subtest scores and the remaining four climate scores, a correlational-type analysis was also carried out to uncover promising directions for future research. Pearson product-moment correlation coefficients were used for this analysis, computed between the OCDQ climate and subtest scores for a school and all of its Stanford subtest raw scores at each grade level. Computations were carried out using subprogram MDRS adapted from Fortran Programming for the Behavioral Sciences (Veldman, 1967) and run on an IBM 360/67 computer. A t-test was performed for each correlation to determine whether it was significantly different from zero. The method given in Hays (1963, p. 529) was used in all tests.

CHAPTER IV

RESULTS

Achievement Analysis of Covariance

Before the hypotheses could be tested statistically, it was necessary to perform the analysis of covariance on school achievement scores, covarying for intelligence scores. This analysis answered the global question, "are there any achievement differences at all between schools when scores are adjusted for intelligence?" Only when it was determined that at least some schools did differ significantly on achievement scores was it permissible to begin grouping schools to test the three hypotheses presented in Chapter I. Tables IV-1, IV-2, IV-3, and IV-4 present the subtest, degrees of freedom, F-ratio, and significance of adjusted school means for grades four through eleven, respectively.* As the tables show, there are significant differences between schools at every grade level on every subtest, since all F-ratios exceeded the .01 level of significance. This outcome might have been predicted merely because of the large N, however, with the resultant high power to detect minute differences between schools (Bakan, 1966).

*Adjusted means for individual schools can be found in Egelston et al. (1967, Tables 5-19, 5-20, 5-21, 5-22).

TABLE IV-1
 ACHIEVEMENT ANALYSIS OF COVARIANCE
 GRADE 4

Subtest	Df1	Df2	F	p
Arithmetic Computation	137	4205	8.3023	< .01
Arithmetic Concepts	137	4212	5.2285	< .01
Arithmetic Applications	136	4184	3.0110	< .01
Word Meaning	137	4195	3.0833	< .01
Paragraph Meaning	137	4202	2.7467	< .01
Spelling	137	4215	2.6783	< .01
Phonics	137	4221	4.3649	< .01
Syllabication	137	4218	4.2047	< .01
Social Studies Content	137	4208	2.1762	< .01
Social Studies Skills	137	4176	3.7527	< .01
Science	137	4210	4.2077	< .01

TABLE IV-2
 ACHIEVEMENT ANALYSIS OF COVARIANCE
 GRADE 6

Subtest	Df1	Df2	F	p
Arithmetic Computation	132	4431	9.5064	< .01
Arithmetic Concepts	132	4432	4.8162	< .01
Arithmetic Applications	132	4434	3.6885	< .01
Word Meaning	132	4370	4.0282	< .01
Paragraph Meaning	132	4362	3.4884	< .01
Spelling	133	4460	2.3136	< .01
Social Studies Content	133	4467	3.1319	< .01
Social Studies Skills	133	4461	3.5641	< .01
Science	133	4463	4.0706	< .01

TABLE IV-3
 ACHIEVEMENT ANALYSIS OF COVARIANCE
 GRADE 9

Subtest	Df1	Df2	F	p
Paragraph Meaning	45	4041	5.0258	<.01
Spelling	45	4041	6.3709	<.01
Arithmetic Computation	45	4050	9.5519	<.01
Arithmetic Concepts	45	4033	6.8356	<.01
Arithmetic Applications	45	4032	8.7038	<.01
Social Studies Content	45	4044	5.3866	<.01
Social Studies Skills	45	4039	4.9733	<.01
Science	45	4037	8.6572	<.01

TABLE IV-4
 ACHIEVEMENT ANALYSIS OF COVARIANCE
 GRADE 11

Subtest	Df1	Df2	F	p
English	47	3503	4.4113	<.01
Numerical Competence	47	3519	4.7509	<.01
Mathematics	47	3518	6.7641	<.01
Reading	47	3521	5.2056	<.01
Science	47	3538	5.0789	<.01
Social Studies	47	3526	5.0949	<.01

Only 113 schools out of 241 agreed to administer the OCDQ, as pointed out under Data Collection in Chapter III. This fact dictated the necessity of an additional analysis to determine whether the 128 schools which refused to participate differed with respect to adjusted achievement scores from the 113 schools which did administer the OCDQ. Since almost all of the schools from both categories were represented in the analysis of covariance on achievement scores, this post hoc test was relatively simple to execute. A Mann-Whitney U test on the adjusted means from the CANOV 1 program was used to test for differences. Tables IV-5, IV-6, IV-7 and IV-8 present the subtest, group N, mean of adjusted scores, direction of the difference, the value of U, the Z-score, and the significance level for grades four through eleven, respectively.

Only two out of thirty-four U's reached the .05 level of significance, which is 6.25% or approximately the amount to be expected by chance alone. This result tends to suggest that there are no important achievement differences between schools administering the OCDQ and those refusing to administer the OCDQ. However small the magnitude of differences, though, the consistent direction of differences in favor of schools who refused to administer

TABLE IV-5
SCHOOLS ADMINISTERING OCDQ VS. SCHOOLS REFUSING
GRADE 4

Subtest	Schools Administering OCDQ		Schools Refusing OCDQ		Direction of Difference*	U	Z	p**
	N	Mean	N	Mean				
Arithmetic Computation	66	13.010	72	13.725	-	1722	2.788	<.05
Arithmetic Concepts	66	13.364	72	13.907	-	1956	1.790	N.S
Arithmetic Applications	66	13.284	71	13.642	-	1995	1.499	N.S
Word Meaning	66	17.568	72	17.516	-	2296	.341	N.S
Paragraph Meaning	66	26.947	72	27.076	-	2342	.145	N.S
Spelling	67	23.884	71	24.395	-	2190	.803	N.S
Phonics	67	23.571	71	23.807	+	2241	1.586	N.S
Syllabication	66	12.106	71	12.439	-	2061	1.215	N.S
Social Studies Concepts	66	10.757	71	11.074	-	2278	.280	N.S
Social Studies Skills	66	12.642	70	12.453	+	2099	.919	N.S
Science	66	28.175	69	27.940	+	2127	.660	N.S.

* A difference in which the schools administering the OCDQ score higher is noted with a plus (+).

**A two-tailed test was used for this analysis.

TABLE IV-6
SCHOOLS ADMINISTERING OCDQ VS. SCHOOLS REFUSING
GRADE 6

Subtest	Schools Administering OCDQ		Schools Refusing OCDQ		Direction of Difference*	U	Z	p**
	N	Mean	N	Mean				
Arithmetic Computation	62	17.028	71	17.461	-	1982	.988	N.S
Arithmetic Concepts	62	15.037	71	15.336	-	1975	1.019	N.S
Arithmetic Applications	62	19.396	71	20.024	-	1686	2.323	.05
Word Meaning	62	24.498	71	24.996	-	1904	1.340	N.S
Paragraph Meaning	62	35.464	71	35.927	-	1940	1.177	N.S
Spelling	62	30.317	71	30.760	-	1997	.920	N.S
Social Studies Concepts	62	23.758	71	23.956	-	2035	.749	N.S
Social Studies Skills	62	16.536	71	16.841	-	1954	1.114	N.S
Science	62	31.943	71	32.287	-	1995	.923	N.S

* A difference in which the schools administering the OCDQ score higher is noted with a plus (+).

**A two-tailed test was used for this analysis.

TABLE IV-7
SCHOOLS ADMINISTERING OCDQ VS. SCHOOLS REFUSING
GRADE 9

Subtest	Schools Administering OCDQ		Schools Refusing OCDQ		Direction of Difference*	U	Z	p**
	N	Mean	N	Mean				
Paragraph Meaning	25	38.354	21	38.854	-	207	1.224	N.S
Spelling	25	35.748	21	35.183	-	259	.077	N.S
Arithmetic Computation	25	25.536	21	25.804	-	229	.739	N.S
Arithmetic Concepts	25	24.226	21	24.645	-	232	.673	N.S
Arithmetic Applications	25	17.656	21	17.831	-	258	.099	N.S
Social Studies Content	25	29.966	21	30.412	-	247	.342	N.S
Social Studies Skill	25	24.472	21	25.116	-	203	1.312	N.S
Science	25	36.521	21	36.967	-	253	.210	N.S

* A difference in which the schools administering the OCDQ score higher is noted with a plus (+).

**A two-tailed test was used for this analysis.

TABLE IV-8
 SCHOOLS ADMINISTERING OCDQ VS. SCHOOLS REFUSING
 GRADE 11

Subtest	Schools Administering OCDQ		Schools Refusing OCDQ		Direction of Difference*	U	Z	p**
	N	Mean	N	Mean				
English	24	56.354	23	57.671	-	204	1.532	N.S
Numerical Competence	24	28.442	23	29.192	-	218	1.234	N.S
Mathematics A	24	23.033	23	23.820	-	219	1.213	N.S
Reading	24	35.637	23	36.643	-	192	1.788	N.S
Science A	24	37.831	23	37.795	+	276	.000	N.S

* A difference in which the schools administering the OCDQ score higher is noted with a plus (+).

** A two-tailed test was used for this analysis.

the OCDQ suggested that there may indeed be a selective bias operating to limit the population to which this study can generalize its results.

Hypothesis One: Open vs. Closed

As soon as scoring of the OCDQ was completed, schools could be dichotomized into Open and Closed categories for use in testing the first hypothesis. Many schools were lost to the sample in this process, as shown in Table IV-9. This summary table is the tabulation of individual school climate similarity scores* from all participating schools. According to Halpin and Croft, only similarity scores below forty-five indicate a resemblance to a given profile (Croft, undated); however, since many of the lowest scores fell in the forty-six through forty-nine range, and since the cutoff point of forty-five was apparently selected arbitrarily, all scores below fifty were included with their appropriate climate-types in the summary table.

As Table IV-9 shows, out of 113 schools there were forty (35.4%) that did not clearly resemble any of the six climates. Nine of the 113 schools had characteristics resembling two or more of the climates, although frequently

* OCDQ subtest scores and climate similarity scores for individual schools can be found in Egelston et al. (1967, Table 7-3).

TABLE IV-9
SCHOOL CLIMATE DISTRIBUTION

	Number of Schools	
Single Climates	Open	15
	Autonomous	6
	Controlled	5
	Familiar	9
	Paternal	11
	Closed	37
Two or More Climates	9	
None of the Above Climates	40	

the resemblances were borderline (nearly fifty) toward one or both of the climates. Of the seventy-three schools that have at least one score less than fifty, fifteen (20.6%) belong to the Open climate category and thirty-seven (50.7%) belong to the Closed category. Upon classifying schools according to grade level, it was found that no eleventh grade schools had an Open climate as measured by the OCDQ, and only two ninth grade schools had an Open climate, so grades nine and eleven were dropped from further analysis relating to the OCDQ.

The fourth and sixth grade schools used to test the first hypothesis are presented in Table IV-10. These are the schools who had administered all three instruments (SAT, L-T, OCDQ) and were classified as either Open or Closed

TABLE IV-10
SCHOOLS USED TO TEST HYPOTHESIS ONE
OPEN VS. CLOSED

	School Code	Grade Levels	Open Climate Similarity Score	Closed Climate Similarity Score
OPEN CLIMATE	1365	6	47	97
	1596	4 6	41	104
	1676	4	38	111
	1696	4 6	23	113
	1975	4 6	46	104
	2015	4 6	31	119
	2036	4 6	47	102
	2065	4 6	35	103
	2195	4 6	33	112
	2196	4 6	44	108
	2215	4 6	38	109
	CLOSED CLIMATE	1305	4 6	111
1346		4 6	109	33
1366		4 6	101	46
1370		4 6	108	39
1416		4 6	111	42
1457		4 6	107	22
1549		4 6	108	31
1626		4 6	91	41
1647		4 6	94	43
1659		4 6	102	37
1660		4 6	111	23
1716		4 6	112	41
1756		4 6	113	41
2085		4 6	98	40
2136		4 6	101	49
2139		4 6	107	42

climates. Because there were far fewer Open than Closed climate schools, all of the Open climate schools were included in the sample, while only enough Closed climate schools were included to approximately equal the Open

climate schools. Beginning with the school which had the lowest Closed climate similarity score, all following schools were included until a natural gap was found in the scores at approximately the right number of schools. In both grades this left more Closed than Open climate schools, but this was judged desirable since it increased both the representativeness of the sample and the power-efficiency of the statistical test. Table IV-10 presents the climate of each school, the school code, the Open climate similarity score, and the Closed climate similarity score.

The last step in testing hypothesis one, after the Open and Closed samples were isolated, was to carry out calculations for the Mann-Whitney U test on the two samples. The results of these calculations are presented in Tables IV-11 and IV-12. Only two fourth grade U 's reached the .05 level of significance, and none from the sixth grades, about the pattern to be expected by chance alone. Moreover, since the direction of the difference was opposite to the direction predicted by the alternative hypothesis, the data failed to reject the null hypothesis for any subject area or either grade. Although most of the differences were small, contrary to the prediction of the alternative hypothesis they were consistently in favor of Closed climate schools.

TABLE IV-11
 OPEN CLIMATE SCHOOLS VS. CLOSED CLIMATE SCHOOLS
 GRADE 4

Subtest	Open Schools		Closed Schools		Direction of Difference*	U	P
	N	Mean	N	Mean			
Arithmetic Computation	10	13.087	14	12.916	+	69.0	N.S
Arithmetic Concepts	10	13.114	14	13.294	-	59.0	N.S
Arithmetic Applications	10	13.181	14	13.889	-	50.0	N.S
Word Meaning	10	17.714	14	18.078	-	66.0	N.S
Paragraph Meaning	10	26.720	14	28.132	-	42.0	N.S
Spelling	10	23.587	14	24.411	-	58.0	N.S
Phonics	10	23.476	14	24.267	-	51.0	N.S
Syllabication	10	12.523	14	12.024	+	54.0	N.S
Social Studies Content	10	10.565	14	10.878	-	49.0	N.S
Social Studies Skills	10	12.048	14	13.197	-	30.0	<.05
Science	10	27.325	14	29.327	-	40.0	<.05

*A difference in which Open schools score higher is noted with a plus (+).

TABLE IV-12
 OPEN CLIMATE SCHOOLS VS. CLOSED CLIMATE SCHOOLS
 GRADE 6

Subtest	Open Schools		Closed Schools		Direction of Difference*	U	p
	N	Mean	N	Mean			
Arithmetic Computation	10	16.553	16	16.370	-	71.0	N.S
Arithmetic Concepts	10	14.713	16	14.974	-	71.0	N.S
Arithmetic Applications	10	19.688	16	19.276	+	69.0	N.S
Word Meaning	10	24.834	16	24.698	+	71.0	N.S
Paragraph Meaning	10	35.814	16	35.371	+	62.0	N.S
Spelling	10	30.426	16	30.125	+	68.0	N.S
Social Studies Content	10	23.866	16	23.714	+	74.0	N.S
Social Studies Skills	10	16.661	16	16.867	-	78.0	N.S
Science	10	31.782	16	32.320	-	66.0	N.S

*A difference in which Open schools score higher is noted with a plus (+).

Hypothesis Two: Young vs. Old

Once the teachers' ages were tallied for each school as described under Scoring in Chapter III, it became possible to assign schools to a "young" or "old" category. The criteria for being assigned to the young category was that the school must employ more teachers younger than thirty-five years of age than it employed teachers older than fifty-five. To be assigned to the old category, the reverse relationship had to hold true. Table IV-13 presents the schools falling into either the old or young categories. Only one eleventh grade school and two ninth grade schools could be categorized old, so the ninth and eleventh grade schools were eliminated completely from further examination. Information contained in the table includes the school code, the grades taught, the total number of professional personnel, the percent under thirty-five years of age, and the percent over fifty-five.

The last step in testing hypothesis two, after the schools were assigned to the young or old samples, was to carry out calculations for the Mann-Whitney U test. The results of these calculations are presented in Tables IV-14 and IV-15. None of the U's reached the .05 level of significance, thus failing to reject the hypothesis of no difference between old and young categorized school; however,

TABLE IV-13
SCHOOLS USED TO TEST HYPOTHESIS TWO
YOUNG VS. OLD

Young or Old	School Code	Grade Levels	Number of Teachers	Percent Under 35	Percent Over 55
y	1305	4 6	18	44.4	11.1
y	1317	4 6	5	60.0	0.0
y	1346	4 6	24	37.5	20.8
y	1365	6	30	37.5	20.8
y	1367	4 6	6	33.3	16.6
y	1416	4 6	30	53.3	10.0
y	1417	4	33	87.8	0.0
y	1418	4 6	29	48.2	20.6
y	1457	4 6	16	43.7	37.5
y	1526	6	17	29.4	11.7
y	1527	4	8	50.0	25.0
y	1529	4	3	66.6	0.0
y	1549	4 6	8	62.5	37.5
y	1596	4 6	10	50.0	30.0
y	1601	4 6	3	66.6	33.3
y	1638	6	18	33.3	16.6
y	1659	6	2	100.0	0.0
y	1660	4 6	12	58.3	0.0
y	1665	6	5	80.0	0.0
y	1676	4	3	66.6	33.3
y	1696	4 6	9	33.3	22.2
y	1736	4 6	9	33.3	22.2
y	1738	4 6	4	50.0	25.0
y	1778	4 6	8	25.0	12.5
y	1975	4 6	17	41.1	29.4
y	2015	4 6	6	33.3	16.6
y	2036	4 6	8	25.0	12.5
y	2065	4 6	8	50.0	12.5
y	2105	4 6	6	33.3	0.0
y	2136	4 6	8	37.5	25.0
y	2138	4 6	10	40.0	30.0
y	2195	4 6	6	43.3	0.0
y	2196	4 6	9	55.5	22.2
y	2215	4 6	4	75.0	25.0

TABLE IV-13 (continued)

Young or Old	School Code	Grade Levels	Number of Teachers	Percent Under 35	Percent Over 55
o	1369	4 6	28	14.2	42.8
o	1370	4 6	2	0.0	100.0
o	1371	4 6	5	20.0	40.0
o	1372	4 6	6	16.6	50.0
o	1373	4 6	6	33.3	50.0
o	1437	4 6	12	8.3	50.0
o	1465	4 6	6	16.6	66.6
o	1547	4 6	11	27.2	45.4
o	1548	4 6	7	28.5	57.1
o	1550	4 6	10	30.0	50.0
o	1598	4 6	9	33.3	55.5
o	1605	6	3	0.0	100.0
o	1616	4	11	9.0	27.2
o	1617	4	7	0.0	57.1
o	1636	4 6	12	0.0	50.0
o	1637	4 6	6	0.0	16.6
o	1647	4 6	7	0.0	57.1
o	1648	4 6	7	14.2	42.8
o	1658	4	4	0.0	100.0
o	1666	4	12	41.6	50.0
o	1756	4 6	21	19.0	38.0
o	1776	4 6	10	30.0	60.0
o	1777	4 6	4	0.0	75.0
o	1786	4 6	4	0.0	50.0
o	1835	4	15	33.3	40.0
o	1965	4	4	0.0	50.0
o	2025	4 6	8	25.0	75.0
o	2085	4 6	8	25.0	50.0
o	2139	4 6	6	0.0	16.6
o	2165	4 6	10	20.0	70.0

there were a large number of differences in favor of schools categorized old, directly contrary to the direction predicted by the alternative hypothesis.

TABLE IV-14
YOUNG SCHOOLS VS. OLD SCHOOLS
GRADE 4

Subtest	Young Schools		Old Schools		Direction of Difference*	U	p
	N	Mean	N	Mean			
Arithmetic Computation	29	12.877	29	13.330	-	345.0	N.S
Arithmetic Concepts	29	13.261	29	13.369	-	389.0	N.S
Arithmetic Applications	29	13.124	29	13.521	-	337.0	N.S
Word Meaning	29	17.422	29	17.588	-	389.0	N.S
Paragraph Meaning	29	26.891	29	26.990	-	384.0	N.S
Spelling	29	23.419	29	23.797	-	406.0	N.S
Phonics	29	23.874	29	22.924	+	326.0	N.S
Syllabication	29	12.117	29	11.974	+	386.0	N.S
Social Studies Content	29	10.690	29	10.860	-	360.0	N.S
Social Studies Skills	29	12.466	29	12.780	-	352.0	N.S
Science	29	28.071	29	28.219	-	397.0	N.S

*A difference in which Young schools score higher is noted with a plus (+).

TABLE IV-15
YOUNG SCHOOLS VS. OLD SCHOOLS
GRADE 6

Subtest	Young Schools		Old Schools		Direction of Difference*	U	p
	N	Mean	N	Mean			
Arithmetic Computation	30	16.324	24	17.489	-	272.0	N.S
Arithmetic Concepts	30	14.706	24	15.284	-	274.0	N.S
Arithmetic Applications	30	19.368	24	19.461	-	349.0	N.S
Word Meaning	30	24.532	24	24.590	-	352.0	N.S
Paragraph Meaning	30	35.515	24	35.388	+	354.0	N.S
Spelling	30	30.673	24	29.822	+	283.0	N.S
Social Studies Content	30	23.660	24	23.662	-	359.0	N.S
Social Studies Skills	30	16.553	24	16.471	+	327.0	N.S
Science	30	31.870	24	31.784	+	357.0	N.S

*A difference in which Young schools score higher is noted with a plus (+).

Hypothesis Three: Young/Open vs. Old/Closed

The sample chosen to test this hypothesis consisted of schools who belonged jointly to the young and Open groups above, or to the old and Closed groups. The schools meeting these criteria are listed in Table IV-16.

TABLE IV-16

SCHOOLS USED TO TEST HYPOTHESIS THREE
OPEN/YOUNG VS. OLD/CLOSED

	School Code	Grade Levels
Young/Open	1365	6
	1596	4 6
	1676	4
	1696	4 6
	1965	4 6
	2015	4 6
	2036	4 6
	2065	4 6
	2195	4 6
	2196	4 6
2215	4 6	
Old/ Closed	1370	4
	1647	4 6
	1756	4 6
	2085	4 6
	2139	4 6

The results of the Mann-Whitney U tests on these samples are presented in Tables IV-17 and IV-18. Only one U reached the .05 level of significance, failing to reject the null hypothesis. The direction of differences was balanced.

TABLE IV-17
 OPEN/YOUNG SCHOOLS VS. CLOSED/OLD SCHOOLS
 GRADE 4

Subtest	Open/Young Schools		Old/Closed Schools		Direction of Difference*	U	p
	N	Mean	N	Mean			
Arithmetic Computation	10	13.087	5	11.648	+	14.0	N.S
Arithmetic Concepts	10	13.114	5	12.733	+	23.0	N.S
Arithmetic Applications	10	13.181	5	13.640	-	23.0	N.S
Word Meaning	10	17.714	5	18.432	-	19.0	N.S
Paragraph Meaning	10	26.720	5	28.546	-	9.0	< .05
Spelling	10	23.587	5	24.367	-	23.0	N.S
Phonics	10	23.476	5	24.130	-	17.0	N.S
Syllabication	10	12.523	5	11.572	+	15.0	N.S
Social Studies Content	10	10.565	5	10.857	-	17.0	N.S
Social Studies Skills	10	12.048	5	13.408	-	16.0	N.S
Science	10	27.325	5	28.742	-	19.0	N.S

* A difference in which Open/Young schools score higher is noted with a plus (+).

TABLE IV-18
 OPEN/YOUNG SCHOOLS VS. CLOSED/OLD SCHOOLS
 GRADE 6

Subtest	Open/Young Schools		Closed/Old Schools		Direction of Difference*	U	p
	N	Mean	N	Mean			
Arithmetic Computation	10	16.353	5	17.488	-	15.0	N.S
Arithmetic Concepts	10	14.713	5	14.761	-	23.0	N.S
Arithmetic Applications	10	19.688	5	19.023	+	21.0	N.S
Word Meaning	10	24.834	5	24.956	-	23.0	N.S
Paragraph Meaning	10	35.814	5	35.014	+	18.0	N.S
Spelling	10	30.426	5	29.844	+	17.0	N.S
Social Studies Content	10	23.866	5	23.290	+	19.0	N.S
Social Studies Skills	10	16.661	5	16.192	+	16.0	N.S
Science	10	31.782	5	31.517	+	25.0	N.S

* A difference in which Open/Young schools score higher is noted with a plus (+).

Correlational Analysis

For this analysis, all schools which administered the OCDQ, SAT, and L-T were included. Correlations were performed at each grade level between all adjusted SAT subtest means and each of the eight OCDQ subtests and six climate similarity scores. Tables IV-19, IV-20, IV-21 and IV-22 present correlation matrices for each grade showing all correlations significantly different from zero at the .05 level. Tables for the sixth, ninth, and eleventh grades contain approximately the number of significant correlations to be expected by chance alone, but a substantial portion (54.5%) of the fourth-grade matrix is filled with significant correlations. Five differentiable patterns seem to be apparent among the correlations in the grade four matrix. First, taking the three Arithmetic subtests as a group, they tend to correlate significantly and in a negative direction with almost all OCDQ subtests, and with the middle and closed climate similarity scores. Second, the Word Meaning, Paragraph Meaning, and Spelling subtests seem to behave as a group, and correlate significantly and positively with almost all subtests and climate scores. Third, neither Word Study Skill subpart seems to correlate highly with anything. Fourth, the two social studies subparts present a conflicting situation, where the pattern of significant

TABLE IV-19

SIGNIFICANT* CORRELATIONS BETWEEN THE OCDQ AND SAT

GRADE 4

N = 66 Schools

OCDQ Subtests	SAT Subtests	Arithmetic	Computation	Arithmetic Concepts	Arithmetic Applications	Word Meaning	Paragraph Meaning	Spelling	Phonics	Syllabication	Social Studies Content	Social Studies Skills	Science
DIS						.224	.328	.216			-.509	.237	.302
HIN	-.253	-.270		-.286		.432	.432	.309			-.570	.429	.312
ESP	-.470	-.388		-.443		.278	.278	.208				.283	-.210
INT	.252			.430		.349	.349	.207				.284	.281
ALO	-.461	-.298		-.330		.290	.395	.398			-.581	.404	
PRD	-.430	-.468		-.481		.318	.453	.409			-.552	.230	
THR		-.313		-.346					-.217		-.259		-.437
CON	-.438	-.385		-.498		.279	.279		-.412		-.548		-.310
OCDQ Climate													
Similarity													
Scores													
OPN						.295	.453	.401			-.349	.395	.403
AUT						.363	.218	.317			-.487	-.408	
CON	-.204	-.218					.565	.266			-.420		
FAM	-.393	-.372		-.393			.269	.249			-.452	.349	
PAT	-.290	-.209					.270	-.265					-.417
CLO	-.214			-.263									

* All correlations presented in this table are significantly different from zero (p .05).

TABLE IV-22
SIGNIFICANT* CORRELATIONS BETWEEN THE OCDQ AND SAT
GRADE 11
N = 24 Schools

OCDQ Subtests	SAT Subtests	English	Numerical Computations	Mathematics ^A	Reading	Science ^A
DIS						
HIN						
ESP						
INT						
ALO						
PRD						
THR						
CON						
OCDQ CLIMATE Similarity Scores						
OPN						
AUT						
CON						
FAM			.419			
PAT					.466	
CLO						

* All correlations presented in this table are significantly different from zero ($p < .05$).

correlations for each approximates a mirror image of the other, Content correlating in a negative direction and Study Skills in a positive direction. These subparts, one or the other, possess significant correlations with almost all of the

subtest and climate scores. Fifth, the Science subtest correlates positively with the Open climate similarity score and negatively with the two Closed scores, and alternates selectively among the subtests.

CHAPTER V

DISCUSSION

The data clearly fail to support any of the three hypotheses, and in the case of the Open vs. Closed comparison, the consistent direction of differences supports the reverse hypothesis, i.e., that Closed climate schools score higher on the Stanford Achievement Test than Open climate schools. However, it seems safe to conclude that on the whole neither school climate, teacher age, nor a combination of the two, is of vital importance in determining levels of student achievement among Upper Peninsula schools.

This conclusion demands curtailment of any thoughts of using the quick, easy OCDQ as a shortcut assessment of a school's achievement performance. In general the OCDQ results agree with other studies relating school climate and achievement (Andrews, 1965; Feldvebel, 1964; Millar, 1965), where the global concept of Organizational Climate had no demonstrable relationship to student achievement. The three studies above examined the ninth, fifth, and ninth grades respectively and their correlational analyses produced significant correlations barely above a chance level, foreshadowing the correlational results in this

study at the upper grade levels. The results of the fourth grade correlations in the present study suggest that something interesting might be happening in the lower grades, however, and point to this direction for future research. In all fairness, the OCDQ was designed for and with elementary schools, and even though examination of the instrument leads one to believe that it is just as well suited for upper elementary and high school grades, the results of this study argue against that belief.

The lack of relationship between Open and Closed climate schools comes as somewhat of a surprise after reading the harsh description of Closed climate schools by Halpin (1966), together with his glowing description of Open climate schools. Apparently the frustration versus cooperation that Halpin senses in these opposing climates is not manifest in the daily round of student activities or, if present, is minimal compared to other influences that the students are exposed to. It must be pointed out again that if grades below the fourth were examined, achievement differences might become evident as presently hinted at by the complex pattern of correlations found in the fourth grade. Perhaps students have closer emotional ties to teachers in these grades, and are more susceptible to the effects of inter-teacher dissonance or harmony.

Even if clear relationships between Open or Closed climate schools and student achievement had been delineated, the OCDQ would still have only very limited usefulness. The reason for this is the high percentage of schools that failed to be classified into any of the six climates (see Table IV-9). There are at least two possible explanations for this occurrence. The first is that Halpin's original factor analysis to isolate the three factors, or six opposites, might have contained a high amount of error variance, causing the factors to break down when replicated on a new sample. In support of this explanation, one researcher (Novotney, 1967) found that two of the subtest factors upon which the six climate scores are based did not replicate well on his sample. The second possible explanation is that for some as yet undetermined reason the Upper Peninsula schools used in this study might be atypical. A likely difference might be school size, for example. Because of the wide geographic distances separating people in the Upper Peninsula, there are many small schools sprinkled through a large area, and schools of this size may be immune to many of the personnel communication difficulties measured by the OCDQ in Halpin's sample of larger schools. School size is one suggestion, but any number of atypical characteristics might occur in a large

and fairly homogenous collection of schools such as those found in the Upper Peninsula. The only way of categorizing the presently unclassified schools is by redoing the steps in the developing of the OCDQ, this time using the Upper Peninsula sample in all item and factor analyses.

The complex pattern of significant correlations found between the SAT and the OCDQ in the fourth grade (Table IV-21) defies simple explanation, but possibly some well-thought-out hypotheses arising from these correlations could serve as the focal point for future studies. Three cautionary points must be made concerning the magnitude of the correlations, however: first, none of the correlations exceed .60, which means that less than 36% of the variance is accounted for by any individual r ; second, the correlations are not independent of one another, since the climate similarity scores are directly based on the subtest scores; third, although the level of rejection for each t-test performed on an individual r is .05, the experiment-wide error will be far higher than this because t-tests were indiscriminately performed on all possible r 's without regard to any theoretical rationale. Nevertheless, the dramatic difference between the fourth grade correlations and the other three grades highlights the fact that something is happening that requires further investigation.

In retrospect, with regard to teachers' ages, it seems naive to expect the usual small differences between percentages of young and old teachers within a school to have a pronounced effect, when frequently both categories combined account for less than half of the school's total professional personnel. In other words, the teachers from thirty-six to fifty-four outnumber either the young or old group in most schools, and outnumber both combined in many schools. Thus even if certain individual older teachers failed to inspire their students to high levels of achievement as successfully as younger teachers did, as implied by the findings of Ryans (1960) and Peterson (1964), in most schools staff ages were heterogenous enough to prevent this from becoming a significant factor in determining the overall achievement levels of students. This does not appear to be a fruitful direction for further investigation.

BIBLIOGRAPHY

Organizational Climate Description Questionnaire

The following 64 items describe situations or conditions that exist within your school. Please consider the items in terms of your immediate faculty and building principal, and not in terms of the whole school district.

All 64 items have the same response categories. Please use the following code for each item.

Mark response "a" if the item rarely occurs.

Mark response "b" if the item sometimes occurs.

Mark response "c" if the item often occurs.

Mark response "d" if the item frequently occurs.

Respond only once to each item.

Mark all responses on the special answer sheet provided.

-
1. Teachers' closest friends are other faculty members at this school.
 2. The mannerisms of teachers at this school are annoying.
 3. Teachers spend time after school with students who have individual problems.
 4. Instructions for the operation of teaching aids are available.
 5. Teachers invite other faculty members to visit them at home.

 6. There is a minority group of teachers who always oppose the majority.
 7. Extra books are available for classroom use.
 8. Sufficient time is given to prepare administrative reports.
 9. Teachers know the family background of other faculty members.
 10. Teachers exert group pressure on nonconforming faculty members.

 11. In faculty meetings, there is the feeling of "let's get things done."
 12. Administrative paper work is burdensome at this school.
 13. Teachers talk about their personal life to other faculty members.
 14. Teachers seek special favors from the principal.
 15. School supplies are readily available for use in classwork.

 16. Student progress reports require too much work.
 17. Teachers have fun socializing together during school time.
 18. Teachers interrupt other faculty members who are talking in staff meetings.
 19. Most of the teachers here accept the faults of their colleagues.
 20. Teachers have too many committee requirements.

 21. There is considerable laughter when teachers gather informally.
 22. Teachers ask nonsensical questions in faculty meetings.

23. Custodial service is available when needed.
 24. Routine duties interfere with the job of teaching.
 25. Teachers prepare administrative reports by themselves.
 26. Teachers ramble when they talk in faculty meetings.
 27. Teachers at this school show much school spirit.

 28. The principal goes out of his way to help teachers.
 29. The principal helps teachers solve personal problems.
 30. Teachers at this school stay by themselves.
 31. The teachers accomplish their work with great vim, vigor, and pleasure.
 32. The principal sets an example by working hard himself.

 33. The principal does personal favors for teachers.
 34. Teachers eat lunch by themselves in their own classrooms.
 35. The morale of the teachers is high.
 36. The principal uses constructive criticism.
 37. The principal stays after school to help teachers finish their work.

 38. Teachers socialize together in small select groups.
 39. The principal makes all class-scheduling decisions.
 40. Teachers are contacted by the principal each day.
 41. The principal is well prepared when he speaks at school functions.
 42. The principal helps staff members settle minor differences.

 43. The principal schedules the work for the teachers.
 44. Teachers leave the grounds during the school day.

 45. Teachers help select which courses will be taught.
 46. The principal corrects teachers' mistakes.
 47. The principal talks a great deal.
 48. The principal explains his reasons for criticism to teachers.
 49. The principal tries to get better salaries for teachers.

 50. Extra duty for teachers is posted conspicuously.
 51. The rules set by the principal are never questioned.
 52. The principal looks out for the personal welfare of teachers.
 53. School secretarial service is available for teachers' use.
 54. The principal runs the faculty meeting like a business conference.

 55. The principal is in the building before teachers arrive.
 56. Teachers work together preparing administrative reports.
 57. Faculty meetings are organized according to a tight agenda.
 58. Faculty meetings are mainly principal-report meetings.
 59. The principal tells teachers of new ideas he has run across.

 60. Teachers talk about leaving the school system.
 61. The principal checks the subject-matter ability of teachers.
 62. The principal is easy to understand.
 63. Teachers are informed of the results of a supervisor's visit.
 64. The principal insures that teachers work to their full capacity.
-
-

BIBLIOGRAPHY

- Andrews, J. H. M. What school climate conditions are desirable? The CSA Bulletin, 1965, IV (5), 4-20
- Bakan, D. The test of significance in psychological research. Psychological Bulletin, 1966, 6, 423-437.
- Bloom, B. S. (Ed.) Taxonomy of educational objectives. New York: David McKay, 1956.
- Buros, O. K. (Ed.) The sixth mental measurements year-book. Highland Park, N. J.: Gryphon Press, 1965.
- Charters, W. W., and Waples, D. The Commonwealth teacher-training study. Chicago: University of Chicago Press, 1929.
- Croft, D. B. Organizational climate scoring program. Salt Lake City: University of Utah, undated (mimeo).
- Egelston, R., Egelston, J., Deloria, D., LeSage, N., Lupton, P., and Pai, B.P. The Upper Peninsula Multi-district Planning Project final report. Ishpeming, Mich.: Ishpeming Public Schools, 1967.
- Feldvebel, A. M. Organizational climate, social class, educational output. Administrator's Notebook, 1964, XII (8).
- Gage, N. L. (Ed.) Handbook of research on teaching. Chicago: Rand McNally, 1963.
- Gardner, E. F., Merwin, J. C., Callis, R., & Madden, R. Stanford Achievement Test: high school battery manual. New York: Harcourt, Brace & World, 1965.
- Halpin, A. W. Theory and research in administration. New York: MacMillan, 1966.
- Halpin, A. W., & Croft, D. B. The organizational climate of schools. Chicago: Midwest Administration Center, University of Chicago, 1963.
- Hays, W. L. Statistics for psychologists. New York: Holt, Rinehart & Winston, 1963.

- Kelly, T. L., Madden, R., Gardner, E. F., & Rudman, H. C. Stanford Achievement Battery: directions for administering advanced battery. New York: Harcourt, Brace & World, 1964.
- Lorge, I., Thorndike, R. L., & Hagen, E. Technical manual: Lorge-Thorndike Intelligence Tests. Boston: Houghton Mifflin, 1966.
- Millar, D. E. Organizational climate and leader behavior. The CSA Bulletin, 1965, IV (5), 36-39.
- Mitzel, H. E. Teacher effectiveness. In C. W. Harris (Ed.), Encyclopedia of educational research. (3rd ed.) New York: MacMillan, 1960.
- Novotney, J. M. Organizational questionnaire in parochial schools. Catholic Educational Review, 1967, LXV (2), 92-113.
- Peterson, W. A. Age, teacher's role, and the institutional setting. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart & Winston, 1964.
- Remmers, H. H., et al. Report of the committee on the criteria of teacher effectiveness. Review of Educational Research, 1952, XX, 238-263.
- Remmers, H. H., et al. Second report of the committee on the criteria of teacher effectiveness. Journal of Educational Research, 1953, XLVI, 641-658.
- Ryans, D. G. Characteristics of teachers: a research study. Washington, D. C.: American Council on Education, 1960.
- Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.
- Veldman, D. J. Fortran programming for the behavioral sciences. New York: Holt, Rinehart & Winston, 1967.

APPENDIX

APPENDIX B

THE EIGHT DIMENSIONS OF ORGANIZATIONAL CLIMATE

Teachers' Behavior

1. Disengagement refers to the teachers' tendency to be "not with it." This dimension describes a group which is "going through the motions," a group that is "not in gear" with respect to the task at hand. It corresponds to the more general concept of anomie as first described by Durkheim.* In short, this subtest focuses upon the teachers' behavior in a task-oriented situation.
2. Hindrance refers to the teachers' feeling that the principal burdens them with routine duties, committee demands, and other requirements which the teachers construe as unnecessary "busywork." The teachers perceive that the principal is hindering rather than facilitating their work.
3. Esprit refers to morale. The teachers feel that their social needs are being satisfied, and that they are, at the same time, enjoying a sense of accomplishment in their job.
4. Intimacy refers to the teachers' enjoyment of friendly social relations with each other. This dimension describes a social-needs satisfaction which is not necessarily associated with task-accomplishment.

Principal's Behavior

5. Aloofness refers to behavior by the principal which is characterized as formal and impersonal. He "goes by the book" and prefers to be guided by rules and policies

*Emile Durkheim, Le Suicide (Paris: Librairie Felix Alcan, 1930), p. 277. Anomie describes a planlessness in living, a method of living which defeats itself because achievement has no longer any criterion of value; happiness always lies beyond any present achievement. Defeat takes the form of ultimate disillusion--a disgust with the futility of endless pursuit.

rather than to deal with the teachers in an informal, face-to-face situation. His behavior, in brief, is universalistic rather than particularistic; nomothetic rather than idiosyncratic. To maintain this style, he keeps himself--at least, "emotionally"--at a distance from his staff.

6. Production Emphasis refers to behavior by the principal which is characterized by close supervision of the staff. He is highly directive and plays the role of a "straw boss." His communication tends to go in only one direction, and he is not sensitive to feedback from the staff.
7. Thrust refers to behavior by the principal which is characterized by his evident effort in trying to "move the organization." Thrust behavior is marked not by close supervision, but by the principal's attempt to motivate the teachers through the example which he personally sets. Apparently, because he does not ask the teachers to give of themselves any more than he willingly gives of himself, his behavior, though starkly task-oriented, is nonetheless viewed favorably by the teachers.
8. Consideration refers to behavior by the principal which is characterized by an inclination to treat the teachers "humanly," to try to do a little something extra for them in human terms.

(Halpin, 1966, pp. 150-151)

APPENDIX C

OCDQ ITEMS ARRANGED BY SUBTEST

1. Disengagement

The mannerisms of teachers at this school are annoying.
There is a minority group of teachers who always oppose the majority.
Teachers exert group pressure on nonconforming faculty members.
Teachers seek special favors from the principal.
Teachers interrupt other faculty members who are talking in staff meetings.
Teachers ask nonsensical questions in faculty meetings.
Teachers ramble when they talk in faculty meetings.
Teachers at this school stay by themselves.
Teachers talk about leaving the school system.
Teachers socialize together in small select groups.

2. Hindrance

Routine duties interfere with the job of teaching.
Teachers have too many committee requirements.
Student progress reports require too much work.
Administrative paper work is burdensome at this school.
Sufficient time is given to prepare administrative reports.
Instructions for the operation of teaching aids are available.

3. Esprit

The morale of the teachers is high.
The teachers accomplish their work with great vim, vigor, and pleasure.
Teachers at this school show much school spirit.
Custodial service is available when needed.
Most of the teachers here accept the faults of their colleagues.
School supplies are readily available for use in classwork.
There is considerable laughter when teachers gather informally.
In faculty meetings, there is the feeling of "let's get things done."
Extra books are available for classroom use.
Teachers spend time after school with students who have individual problems.

4. Intimacy

Teachers' closest friends are other faculty members at this school.
Teachers invite other faculty members to visit them at home.
Teachers know the family background of other faculty members.
Teachers talk about their personal life to other faculty members.
Teachers have fun socializing together during school time.
Teachers work together preparing administrative reports.
Teachers prepare administrative reports by themselves.

5. Aloofness

Faculty meetings are organized according to a tight agenda.
Faculty meetings are mainly principal-report meetings.
The principal runs the faculty meeting like a business conference.
Teachers leave the grounds during the school day.
Teachers eat lunch by themselves in their own classrooms.
The rules set by the principal are questioned.
Teachers are contacted by the principal each day.
School secretarial service is available for teachers' use.
Teachers are informed of the results of a supervisor's visit.

6. Production Emphasis

The principal makes all class scheduling decisions.
The principal schedules the work for the teachers.
The principal checks the subject-matter ability of teachers.
The principal corrects teachers' mistakes.
The principal insures that teachers work to their full capacity.
Extra duty for teachers is posted conspicuously.
The principal talks a great deal.

7. Thrust

The principal goes out of his way to help teachers.
The principal sets an example by working hard himself.
The principal uses constructive criticism.
The principal is well prepared when he speaks at school functions.

The principal explains his reasons for criticism to teachers.
The principal looks out for the personal welfare of teachers.
The principal is in the building before the teachers arrive.
The principal tells teachers of new ideas he has run across.
The principal is easy to understand.

8. Consideration

The principal helps teachers solve personal problems.
The principal does personal favors for teachers.
The principal stays after school to help teachers finish their work.
The principal helps staff members settle minor differences.
Teachers help select which courses will be taught.
The principal tries to get better salaries for teachers.

APPENDIX D
OCDQ CLIMATE DESCRIPTIONS

The Open Climate

The Open Climate depicts a situation in which the members enjoy extremely high Esprit. The teachers work well together without bickering and griping (low Disengagement). The principal's policies facilitate the teachers' accomplishment of their tasks (low Hindrance). On the whole, the group members enjoy friendly relations with each other, but they apparently feel no need for an extremely high degree of intimacy.

The principal sets an example by working hard himself (high Thrust), and goes out of his way to help a teacher (high Consideration). He is not aloof, nor are the rules and procedures which he sets up inflexible and impersonal. He does not do all the work himself because he has the ability to let appropriate leadership emerge from the teachers (low Production Emphasis). Withal, he is in full control of the situation, and he clearly provides leadership for the staff.

The Autonomous Climate

The distinguishing feature of this Organizational Climate is the almost complete freedom that the principal gives to teachers to provide their own structures-for-inter-

action so that they can find ways within the group for satisfying their social needs. As one might surmise, the scores lean slightly more toward social-needs satisfaction than towards task-achievement (relatively high scores on Esprit and Intimacy). When teachers are together in a task-oriented situation they are engaged in their work; they achieve their goals easily and quickly (low Disengagement). A teacher does not have to run to the principal every time he needs supplies, books, projectors, and so on; adequate controls have been established to relieve the principal as well as the teachers of these details (low Hindrance). The morale of the teachers is high, but not as high as in the Open Climate.

The principal remains aloof from the teachers, for he runs the organization in a businesslike and a rather impersonal manner (high Aloofness). The principal appears satisfied to let the teachers work at their own speed; he monitors their activities very little (low Production Emphasis). On the whole, he is considerate, and he attempts to satisfy the social needs of the teachers as well as do most principals (average Consideration). The principal provides Thrust for the organization by setting an example and by working hard himself.

The Controlled Climate

The Controlled Climate is marked, above everything else, by a press for achievement at the expense of social-needs satisfaction. Nonetheless, since morale is high (Esprit), this climate can be classified as more Open than Closed. The teachers are completely engaged in the task. They are there to get the job done, and they expect to be told personally just how to do it (low Disengagement). Few procedures have been set up to facilitate their work; in fact, paper work seems to be used to keep them busy (high Hindrance). Accordingly, teachers have little time to establish very friendly social relations with each other, and there is little feeling of camaraderie (low Intimacy). Esprit, however, is slightly above average. We infer that the job satisfaction found in this climate results primarily from task-accomplishment, not from social-needs satisfaction.

The principal is described as dominating and directive; he allows little flexibility within the organization, and he insists that everything be done "his" way (High Production Emphasis). He is somewhat aloof; he prefers to publish directives to indicate how each procedure is to be followed. He cares little about how people feel; the important thing is to get the job done, and in his way. Accordingly, he does

not seek to satisfy the group's social needs (low Consideration). Nevertheless, he is trying to move the organization by working hard (average Thrust), and he personally sees to it that everything runs properly.

The Familiar Climate

The main feature of this climate is the conspicuously friendly manner of both the principal and the teachers. Social-needs satisfaction is extremely high while, contrariwise, little is done to control or direct the group's activities toward goal achievement. The teachers are disengaged and accomplish little in a task-oriented situation, primarily because the principal exerts little control in directing their activities. The principal does not burden the teachers with routine reports; in fact, he makes it as easy as possible for them to work. Procedural helps are available (low Hindrance). The teachers have established personal friendships among themselves, and socially at least, everyone is part of a big, happy family (high Intimacy). The Esprit that is found in this climate is one-sided in that it stems almost entirely from social-needs satisfaction.

The principal evidently is reluctant to be anything other than considerate, lest he, in his estimation, injure the "happy family" feeling (high Consideration.) Few rules

and regulations are published as guides to suggest to the teachers how things "should be done" (low Aloofness). The actions of members--at least in respect to task accomplishment--are not criticized (low Production Emphasis). Teachers do attribute Thrust to the principal but, in this context, this probably means that they regard him as a "good guy" who is interested in their welfare and who "looks out for them."

The Paternal Climate

The Paternal Climate is characterized by the ineffective attempts of the principal to control the teachers as well as to satisfy their social needs. This climate is, of course, a partly Closed one. The teachers do not work well together; they are split into factions. Group maintenance has not been established because of the principal's inability to control the activities of the teachers (high Disengagement). Few hindrances burden the teachers in the form of routine reports, administrative duties, and committee requirements, mainly because the principal does a great deal of the busywork himself (low Hindrance). The teachers do not enjoy friendly relationships with each other (low Intimacy). Obviously, low Esprit results when the teachers obtain inadequate satisfaction in respect to both task-accomplishment and social-needs.

The principal, on the other hand, is the very opposite of aloof. In fact, he is so non-aloof that he becomes

intrusive; he must know everything that is going on. He is always emphasizing all the things that should be done (high Production Emphasis) but somehow nothing does get done. The principal is considerate, but his consideration appears to be a form of seductive oversolicitousness rather than a genuine concern for the needs of others. Although he preserves an average degree of Thrust, as evidenced by his attempts to move the organization, he nonetheless fails to motivate the teachers, primarily because he himself does not provide an example.

The Closed Climate

The Closed Climate marks a situation in which the group members obtain little satisfaction in respect to either task-achievement or social-needs. In short, the principal is ineffective in directing the activities of the teachers; at the same time, he is not inclined to look out for their personal welfare. The teachers are disengaged and do not work well together; consequently, group achievement is minimal (high Disengagement). The principal does not facilitate the task-accomplishment of teachers (high Hindrance). Esprit is at a nadir, reflecting low job satisfaction in respect to both job satisfaction and social-needs satisfaction. The salient

bright spot that appears to keep the teachers in the school is that they do obtain satisfaction from their friendly relationships with other teachers (average Intimacy).

The principal is highly aloof and impersonal in controlling and directing the activities of the teachers (high Aloofness). He sets up rules and regulations about how things should be done, and these rules are usually arbitrary (high Production Emphasis). But his words are hollow because he himself possesses little Thrust and does not motivate the teachers by setting a good personal example. He is not concerned with the social needs of teachers; in fact, he can be depicted as inconsiderate (low Consideration). This climate characterizes an organization for which the best prescription is radical surgery.

(Abstracted from Halpin, 1966)