The Use of high - inference Measures To Study Classroom Climates: A Review

Rudolfo Chávez Chávez
California State College, Bakersfield

Review of Educational Research
ABSTRACT

This paper reviews the impact high-reference measures had on studying classroom climates. Early literature of classroom climates from the late twenties through the early sixties is reviewed. It is noted that the study of classroom climates had two distinct, unrelated beginnings. As a result, the use of high-inference measures to study classroom climates was not prevalent in the literature until the mid-sixties, whereas the study of classroom climates using other measures was well established by the early sixties. The review then discusses several high-inference measures developed in the last 30 years. Also, pertinent classroom climate literature that incorporates specific high-inference measures has been reviewed.

After reviewing much of the current classroom research of the time, Withall and Lewis (1963) concluded that:

Much of early research of classroom learning was set in the stark framework of the given conditions in an educational situation . . . . It seems that the researchers of the 1920’s and the mid-1930’s sought to analyze the process of teaching in terms of the conditions brought to it by the teacher as a professional worker. She [ was presumed to be devoid of any personal needs, purposes, idiosyncrasies. The other actors in the classroom, i.e. the learners, were even more taken for granted. (p. 709, emphasis added)

More recently, John A. Goodlad (1979) stated that “too many researchers are preoccupied with research on single instructional variables that rarely account for more than 5% of the variance in student outcomes. Too few [researchers] study the complex phenomena of schooling in their natural environment, developing the needed new methodologies instead of seeking to adapt the old” (p. 347). Contrary to these remarks that learners were taken for granted and that the natural environment of schooling was not being studied, there has been a sustained research effort of studying learners’ interactions in day-to-day classroom activities in the natural environment. This review will discuss that research effort in two parts.

The first part is a review of the early literature of classroom climates from the late twenties through the early sixties. It should be noted that according to the literature reviewed, classroom climate research using high-inference measures had two distinct, unrelated beginnings. (Rosenshine & Furst, 1971, define a high inference measure as a rating system that requires an observer to make an inference from a series of classroom events using specific constructs, such as satisfaction, cohesiveness, etc.) The use of high-inference measures to study classroom climates was not prevalent in the literature until the mid-sixties, whereas the study of classroom climates with low-inference measures was well established by the early sixties. (Rosenshine & Furst, 1971, define a low-inference measure as a rating system that classifies specific, denotable, relatively objective classroom behavior and is recorded as frequency counts by an observer.) Nonetheless, the early classroom climate research using low-inference measures is relevant because it serves as a cornerstone for the development of a conceptual framework of classroom climate research. The second part of this review is a discussion of several high-inference measures that have been used in the last 30 years, with particular attention given to the My (lass Inventory (MCI) and the Learning Environment Inventory (LEI).

Early Literature of Classroom Climates

Researchers who took an early interest in classroom behavior were social psychologists. Their basic interest was in the interactions between student and student and between student and teacher
Research emphasizing these interactions seems to have begun with the work of Dorothy Thomas (1929). Thomas mentioned that the study of classroom behavior consisted largely of descriptive accounts (i.e., case histories and diary records). She stated, however, that those accounts presented certain difficulties as material for scientific analysis:

The data obtained in such records are, at their best, objective, in the sense that they deal with certain verifiable facts, but they are selective, inconsistent, and usually incomparable with other records. This is due to the tremendous complexity of any social behavior act and the consequent recording of different elements of these complex acts at different times. At their worst, these records are such an intermixture of fact and interpretation as to be utterly worthless from the scientific point of view. Even at their objective best, the selection and emphasis are more or less dependent on the recorder. The control of this sort of error in our social data is one of the first problems claiming our attention. In other words our data must become independent of our observers within a small and predictable range of error. (p. 3)

To accomplish this, she concluded that the study of social behavior could best be made by evolving indices of an individual’s overt actions involving other persons compared with overt actions involving objects (or abstractions) and the self (Thomas, 1929). Three general techniques were developed: observing a child for a given period in the nursery school and recording the child’s overt social behavior each time it occurred; recording specific social situations made by the child within the nursery school; and creating a psychological test situation with limited social and material stimuli instead of the nursery school for recording data. Because of the high standards set by Thomas (i.e., accuracy and objectivity), research many years later would have similar standards (cf. Medley & Mitzel, 1963).

An important influence on the study of classroom climates was the work of Lewin, Lippitt, and White (1939) and Lippitt (1940, cited in Withall, 1949). These studies examined the effect of three leadership roles and the concomitant group climates by observing the behavioral patterns of four “clubs” of five 10- or 11-year-old boys. Each club was examined under three leadership conditions: democratic, autocratic, and laissez-faire. In addition to matching the groups to control for individual differences, leaders were rotated to control for treatment variations. The data recorded were (a) the interaction within each group, (b) the interaction between the leader and the group, (c) the overt expression of aggression, and (d) the productivity in club projects.

The findings of both studies were similar. Lippitt (1940) concluded the following:

1. Different leadership styles produced different social climates and resulted in different group and individual behaviors.
2. Conversation categories differentiated leader behavior techniques more adequately than social behavior categories.
3. Autocratic leadership elicited either an aggressive rebelliousness toward the leader or an apathetic submission to the leader.
4. Leadership style was the primary factor in producing climatological differences and club personnel were of secondary importance.

For Lewin et al. (1939), the following conclusions stood out above all others:

1. Aggressive behavior was either very high or very low under authoritarian conditions, extremely high under laissez-faire conditions, and intermediate under democratic conditions.
2. Productive behavior was higher than or as high in authoritarian climates when the leader was present as in democratic climates but much lower when the leader was absent, moderately high and independent of the leader’s presence or absence in the democratic climates, and lowest in the laissez-faire climates.

Shortly after the Lewin et al. and Lippitt studies came investigations to determine the influence of teachers’ classroom personality on students’ behavior and students’ classroom behavior on each other (Anderson & Brewer, 1945; Anderson & Brewer, 1946) and the development of a method to observe students’ and teachers’ behaviors simultaneously (Anderson, Brewer, & Reed, 1946). These researchers developed 23 teacher behavior categories and 7 student behavior categories. More important, they divided the behavior into socially integrative or dominative behavior. According to the researchers, the former behavior promoted the interplay of differences and advanced “the psychological processes of differentiation, [that] facilitate[d] the emergence of originals” (Anderson & Brewer, 1945, p. 9). The latter behavior was “characterized by rigidity or inflexibility of purpose, by an inability or an unwillingness to admit the contribution of another’s experience, desires, purposes, or judgment (Anderson & Brewer, 1945. p.9). Socially integrative behavior (among teachers and students) was flexible, adaptive, objective, scientific, cooperative; dominative behavior, on the other hand, was seen as stifling someone else’s behavior coupled with a resistance to change and consistent with bigotry and autocracy (Anderson & Brewer, 1945).

The Anderson et al. (1946) study illustrated that students’ as well as teachers’ behaviors could be categorized and “it demonstrat[ed] further the validity of this method of studying teachers’ classroom personalities” (p. 153). However, because of the limited population (data were collected on four teachers and four students) it was impossible to estimate the reliability of any of the scores proposed for comparing different teachers, classes, or occasions in the same class. Moreover, it was recognized that the validity of the measures of teacher behavior would have been higher if the proposed categories were internally consistent (Anderson & Brewer, 1945).

As the fifties approached, classroom climate research became theoretically and empirically oriented. The streams of thought captured for this orientation were Lewin’s (1936) field theory, Murray’s (1938) need-press model, Thelen’s (1950) educational dynamics model, and others. Hypotheses were derived from analysis of time-lapse pictures, recordings (stenographic and mechanical), and observations in the classroom by Sensitive and trained educators using newly developed measures, which were often compared with the results of standardized tests (cf. Medley & Mitzel, 1963; Withall & Lewis, 1963). The earliest of the several researchers to follow was Withall (1949). In his study, he renamed the interactions between students and students and students and teachers as the social emotional climate. This group phenomenon was defined as

a general emotional factor which appears to be present in interactions occurring between individuals in face to face groups. It seems to have some relationship to the degree of acceptance expressed by members of a group regarding each other’s needs or goals. Operationally defined, it is considered to influence: (1) the inner private world of each individual; (2) the esprit de corps of a group; (3) the sense of meaningfulness of group and individual goals and activities; (4) the objectivity with which a problem is attacked; and (5) the kind and extent of interpersonal interaction in a group. (pp. 348—349)
This definition illustrates phenomenological activities that are emotional and intellectual on the one hand, and individual and social on the other, where all activities are interactive within the classroom. However, Withall (1949) did not seem to believe students’ interactions, as suggested by earlier studies, to be as important as the teacher’s interactions. He suggested that it should be possible to measure the socioemotional climate in terms of teacher behavior alone and developed the following seven categories to encompass all types of statements that teachers use in classrooms: “(a) learner-supportive, (b) acceptant and clarifying, (c) problem-structuring, (d) neutral, (e) directive or hortative, (f) reproving or deprecating, (g) teacher self-supporting” (p. 349). Withall perceived these categories to lie on a continuum, whereby a continuum from problem-centeredness to person centeredness could exist, or from objectivity to subjectivity, or from learner centeredness to teacher-centeredness. Therefore, Withall believed categories a, b, and c to be learner-centered and categories e, f, and g to be teacher-centered. The patterns of verbal behavior elicited by the teacher being observed would determine whether the teacher was learner- or teacher-centered or another of the seven mentioned categories. For example, if the ratio of category c outweighed the proportions of categories a and b combined or categories e, f, and g combined, then the teacher was said to be more problem-centered than learner- or teacher centered (Withall. 1949).

Withall (1949) found significant relationships among different measures of group process, pupil reactions, expert ratings, and styles of problem-solving activity. In addition, Withall and Thelen (1949) validated Withall’s Climate Index. The validation of the instrument involved (a) a comparison of the categories of the Climate Index with Anderson and Brewer’s (1946) Integrative-Dominance Ratio secured on the same data; (b) pupils’ pencil-and-paper assessment of their classroom situation on the basis of seven standardized questions; and (c) the use of an electrical graph for recording pupils’ positive and negative feelings expressed by their pushing one of two buttons while being exposed to experimentally varied “learner-centered” and “teacher-centered” social emotional climates. Moreover, Withall (1949) found that different teachers produced different climates with the same group of pupils.

Although the following two studies, also from the early fifties, are not directly related to classroom climate research, their results are a necessary link.1 Bovard (1951) argued that a group-centered classroom climate is more amenable to more students than a leader-centered classroom climate. He defines group-centered as having “student-to-student verbal interaction . . . fostered by a number of specific techniques, such as seating students in a circle, and deflection of teacher-directed questions back to the group” (p.215). He defines leader-centered as having “student- to-student conversations..., politely but firmly limited, and verbal interaction channeled between teacher and individual student” (p. 215). For the study, two college psychology classes served as the experimental and control groups. Each class had the same teacher, course content, examinations, and text, and the same relative amount of time devoted to lectures, discussion, and role playing. Also, the two classes were matched on number of veterans, grade point averages of the previous semester, and scores on the Otis Self-Administering Test of Mental Ability (Form A) (Bovard, 1951). Bovard used two techniques to measure for differences:

(a) outside observation of class discussion, and (b) the administration of an “affect scale.” He found that 61% of student remarks were directed to other students in the group-centered class, whereas only 10% of student remarks were directed to other students in the leader-centered
class. Also, the average level of affect was greater in the group-centered than in the leader-centered class.

The findings, however, suffer methodological error. To illustrate, Bovard mentioned in passing that students in the control and contrast classes knew about each other. One assumption may be made that students may have had the opportunity to converse about the class and its professor. Consequently, discovering that the professor was changing behavior from one class to another may have led the students to deduce that a prescribed role set was being fostered in the classroom.

This could influence the students (consciously or unconsciously) to act unnaturally or play up to the prescribed role, resulting in the guinea pig effect (see Isaac & Michael, 1978, pp. 62—63).

In spite of the above error, the implicit theme of Bovard’s study was to determine if Negro student veterans, Jewish and Catholic students, and students from a wide range of socioeconomic backgrounds would develop a cohesive group by allowing verbal interaction among them to occur. Bovard found this to be true. Bovard (1951) assumed “from the present evidence that the amount of social interaction in the classroom will influence the individual student’s perception, feelings, and interpersonal relations, and perhaps even (the student’s) personality development” (p. 223). Moreover, because of the racial, denominational, and socioeconomic implications of this study and regardless of its methodological error, Bovard’s study is a pinnacle study which investigates classroom climate and the qualitative effects it may have on its student members.

Buswell’s (1953) study also had an indirect influence on classroom climates. Her study is important to this review because it considered social relationships in the classroom. Her purpose was to determine whether or not those students who were accepted by their peers differed in certain achievements from those who were rejected. She hypothesized that no relationship between the social structure of a classroom group and the students’ achievement in some of the basic elementary school subjects would exist.

Buswell (1953) studied a group of 286 kindergarten children and a group of 321 fifth grade children. Several standardized tests (Stanford-Binet, Iowa Test of Basic Skills, Ohio Social Acceptance Scale) and a sociometric test to determine the best liked peers were used. She found that the highly accepted group (according to the sociometric test) was significantly higher than the rejected group in mean achievements (i.e., the standardized test). Her primary conclusion “is that when we consider a classroom of boys and girls in either the lower or upper grades, it may be said that in general those who are succeeding in their school work will also be succeeding in their social relationships with their peers” (p. 51). Though Buswell’s work does not discuss classroom climates per se, it does consider and discuss the students’ social interactions in the classroom and the resulting social system that may form a classroom climate (Getzels, 1969; Withall & Lewis, 1963).

The following studies are more closely related to a trend that developed throughout the late fifties and early sixties in classroom climate research in which the concomitant classification of nonverbal behavior and classroom social structure would be measured. Based on the research efforts of Cornell, Lindvall, and Saupe (1952), Medley and Mitzel (1958a, 1963), and Withall (1949), Medley and Mitzel (1958b) constructed an omnibus instrument termed “OScAR” (Observation Schedule and Record).
The OSCAR evolved by modifying and combining items constructed by Cornell et al. (1952) and Withall (1949) on the basis of two techniques developed by Medley and Mitzel (1963). The techniques reduced the difficulty of the judgments required, which would increase observer behavior, and redefined Cornell et al.’s and Withall’s categories in simpler terms. This would decrease the amount of training for the use of OScAR (Medley & Mitzel, 1958b, 1963). Also, there was a separate process for scoring and a process for observation of the classroom activities.

The OScAR has 14 scoring keys that can be categorized into three scales. The scoring keys are (a) time spent on reading, (b) problem-structuring teacher statements, (c) autonomous administrative groupings, (d) pupil leadership activities, (e) freedom of movement, (f) manifest teacher hostility, (g) supportive teacher behavior, (h) time spent on social studies, (i) disorderly pupil behavior, (j) verbal activities, (k) traditional pupil activities, (l) teacher’s verbal output, (m) audiovisual materials, and (n) autonomous social groupings. The three scales are (a) emotional climate, (b) verbal emphasis, and (c) social organization (Medley & Mitzel, 1958b, 1959; see Medley & Mitzel, 1959, 1963, for a visual representation of the OScAR).

Medley and Mitzel’s original intention for the OSCAR was to obtain some indices of objective data (i.e., classroom behavior) to solve “practical problems such as how to select students likely to become successful teachers, how to screen out those who cannot get along with children, and what ought to be the content of teacher training” (1958b, p. 91). Curiously, the effect of the emotional climate scale overshadowed the other two scales of the OScAR. Morrison (1961), using the OScAR, found that emotional climate was an important factor of teacher competence and the rapport between teacher and pupil. In another study, Bowers and Soar (1961) found a correlation between the emotional climate of the OScAR and the Minnesota Teacher Attitude Inventory that was consistent with the relationship between teacher effectiveness and classroom behavior found by Medley and Mitzel (1959). As shown by Medley and Mitzel and others working with the OScAR, concomitant classification of nonverbal behavior and classroom social structure was possible. Their studies indicated that the emotional climate scale was more reliable when compared with the scoring keys of the OScAR than were the verbal emphasis scale and the social organization scale. This illustrates that the emotional tone of classroom climates is a most important dimension in classroom research.

Hughes’s (1959) study follows this trend of classifying nonverbal behavior and classroom social structure, with a direct implication to classroom climate research. She developed a comprehensive set of categories for the classification of teacher behavior that are reminiscent of the categories developed by Withall (1949). Her contribution to classroom climate research was the development of social categories that focused on social interaction in an elementary school setting. The following are the seven categories with a prescribed percentage range: (a) controlling function, 20—40%; (b) imposition, 1—3%; (c) facilitating, 5—15%; (d) content development, 20—40%; (e) personal response, 8—20%; (f) positive affectivity, 10—20%; and (g) negative affectivity, 3—10%. Hughes’s implication from these percentage allotments appears to be that the quality of classroom life vis-à-vis social interaction will affect teacher and student behavior positively.

Thus far, the studies cited have directly dealt with classroom climates or have had an influence on classroom climate research. The studies beginning with Thomas (1929) and ending with Hughes (1959) have all used low-inference measures of one type or another. Of interest, most of the measures
used have been cumbersome (e.g., Cornell et al.’s, 1952, COCD, or Medley and Mitzel’s, 1958b, 1959, OSCAR), or have lacked a methodological framework, which could result in misinformation of reported data (e.g., Bovard, 1951).

One other low-inference measure, the Flander’s Interaction Analysis System (Amidon & Hough, 1967), will be summarized. The Flander’s Interaction Analysis System is important to this review because it was the most sophisticated technique for the observation of classroom climates developed in the early sixties. It was unique at that time because it preserved a certain amount of information regarding the sequence of behavior (Medley & Mitzel, 1963). This system records the sequence of behavioral events, focuses on teacher influence, distinguishing between “direct” and “indirect” influence according to the 10 items listed in Table I.

Flander’s Interaction Analysis System was used extensively throughout the sixties and seventies in its original form as well as in modified forms (see Amidon & Hough, 1967; Flanders, 1970; Furst, 1967; LaShier, 1967; Rosenshine & Furst, 1973) and also modified for bilingual interaction analysis (see Hughes & Harrison, 1971; Townsend, 1976; United States Commission on Civil Rights, 1973: Zamora, 1974). It is beyond the scope of this review to summarize the results of the mentioned studies that used the Flander’s Interaction Analysis System. Suffice it to say that this analysis system and many other low-inference measures reflecting a train of thought (i.e., behaviorism) prevalent throughout the sixties had an impact on how to describe teacher and class behavior.

### TABLE I

**Flanders Interaction Analysis System**

<table>
<thead>
<tr>
<th>Teacher talk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indirect influence</strong></td>
</tr>
<tr>
<td>1. Accepts feelings: Accepts and clarifies the feeling tone of the students in a nonthreatening manner. Feelings may be positive or negative. Predicting or recalling feelings is included.</td>
</tr>
<tr>
<td>2. Praises or encourages: Praises or encourages student action or behavior. Jokes that release tension, not at the expense of another individual, nodding head or saying, “Urn hum?” or “Go on” are included.</td>
</tr>
<tr>
<td>3. Accepts or uses ideas of student: Clarifying, building, or developing ideas suggested by a student. As teacher brings more of his own ideas into play, shift to category 5.</td>
</tr>
<tr>
<td>4. Asks questions. Asking a question about content or procedure with the intent that a student answer.</td>
</tr>
<tr>
<td><strong>Direct influence</strong></td>
</tr>
<tr>
<td>5. Lecturing: Giving facts or opinions about content or procedure; expressing own ideas, asking rhetorical questions.</td>
</tr>
<tr>
<td>6. Giving ideas: Directions, commands, or orders to which a student is expected to comply.</td>
</tr>
</tbody>
</table>
7. Criticizing or justifying authority: Statements intended to change student behavior from nonacceptable to acceptable pattern; bawling someone out: stating why the teacher is doing what he is doing; extreme self-reference.

8. Student talk-response: Talk by students in response to teacher. Teacher initiates the contact or solicits student statement.

9. Student talk-initiation: Talk by students that they initiate. If calling on student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he or she did, use this category.

10. Silence or confusion: Pauses, short periods of silence, and periods of confusion in which communications cannot be understood by the observer.

---

HIGH-INFERENCE MEASURES/CLASSROOM CLIMATES

In summary, research on classroom behavior began with the early work of Dorothy Thomas, who was interested in the consistency of observers’ interpretation of classroom behavior. She used three techniques in which accuracy and objectivity were the sine qua non of her research. A few years later, Lewin et al. and Lippitt found that leadership styles affected classroom behavior. Anderson et al. also found that a teacher’s classroom personality affects classroom behavior. Withall renamed that classroom behavior social emotional climate and developed seven categories to describe it. Next, Bovard and Buswell contributed indirectly by influencing applied research. Finally, the advent of the sixties brought a high sophistication to low-inference measures, such as Medley and Mitzel’s OScAR, Hughes’s instrument based on Withall’s seven categories, and the Flander’s Interaction Analysis System. There are differences in the terms applied by all these researchers to the behaviors within the classroom. Nevertheless, all are referring to highly similar, and in many cases identical, phenomena of the classroom climate, which can be measured and are important to the study of classroom climate literature. Following is a discussion of high-inference measures and their subsequent use for classroom climate research.

High-Inference Measures

In 1950—51 the Activities Index (AI) was constructed at the University of Chicago (Stern, 1970). The AI was employed in a number of studies at the University of Chicago for assessing student personality (Stern, Stein, & Bloom, 1956). As descendents of the AI the environment indexes were developed, which in short were based on “the various aspects of environment in high schools, colleges, and evening colleges which help to give them their unique cultural atmospheres” (Stern, 1970, p. 14). Four environment indexes, based on Murray’s (1938) Need-Press Model, were developed: (a) the College Characteristics Index (CCI; Pace & Stern, 1958); (b) the High School Characteristics Index (HSCI; developed in 1960 and reported by Stern, 1970); (c) the Evening College Characteristics Index
(ECCI; developed in 1961 and reported by Stern, 1970); and (d) the Organizational Climate Index (OCI: developed in 1963 and reported by Stern, 1970).

These Indexes were high-inference measures. The indexes were self-administered questionnaires requiring approximately 15 minutes to complete. The CCI, HSCI, and ECCI were similar in that each consisted of 300 items distributed among 30 scales of 10 items each (Stern, 1970). The student’s task on the OCI, HSCI, or ECCI was to answer “true” or “false” on the 300 items covering a wide range of topics about the high school or college. The scales were constructed to be able to infer continuity and consistency in otherwise discrete events (Stern, 1970). If, for example, students were assigned seats in classrooms, if attendance records were kept, if faculty members saw students outside class only by appointment, if there was a prescribed form for all term papers, if neatness counted, among other things, then it was probable that the “press” at this school emphasized the development of orderly responses on the part of students: “These conditions establish the social climate or atmosphere of a setting. Students’ perceptions of this climate or ‘learning environment’ were useful in predicting achievement and in contributing to the understanding of educational processes” (Moos, 1979b, p. 81; cf. Stern, 1970, and Walberg, 1976). Furthermore, other high-inference measures were developed to assess the climate of schools or classrooms, Sinclair (1969) developed the Elementary School Environment (ESES) instrument consisting of 100 statements about elementary school conditions, processes, and activities to obtain a measure of student perceptions of the educational environment (cf. Nielsen & Kirk, 1974).

Another high-inference measure, also based on the Murray Need-Press Model, was the Classroom Environment Scale (CES; Trickett & Moos, 1970). The CES has nine dimensions with 10 items per dimension. The dimensions or subscales are (a) involvement, (b) affiliation, (c) support, (d) task orientation, (e) competition, (f) order and organization, (g) rule clarity, (h) teacher control, and (i) innovation. Also, three different forms of the CES were developed: The Real Form (Form R) asks teachers and students how they perceive the current classroom social environment; the Ideal Form (Form I) asks people how they conceive of an ideal classroom environment; and the Expectations Form (Form E) asks prospective members of a class what they think the social milieu they are about to enter is like (Moos, 1979). The CES was reliable in measuring several types of interaction that occur within a classroom. For example, Trickett and Moos (1970), on discovering that significant amounts of variance resided in the student by classroom interaction, suggested that if the prediction of behavior is to improve, systematic data about the context of behavior will be a prerequisite. Hence, by using the CES Trickett and Moos (1973) proposed they could capture “the psychosocial environment” of . . . junior. . . and high school classrooms by asking teachers and students to report their perception of various aspects of the classroom” (p. 94). Trickett and Moos (1973) found that the “CES did indeed discriminate among classrooms on the nine dimensions and in so doing gave a differentiated picture of the environment of diverse classrooms” (p. 100).

Another study employing the CES illustrates well the range of this high-inference measure. Moos and Moos (1978) hypothesized that “involvement, affiliation, and support are positively related and competition and teacher control are negatively related to average class grades” (p. 264). They found “substantial relationships” between student and teacher perceptions of the classroom environment and mean class grades. For student perceptions, all three of the relationship dimensions (i.e., involvement, affiliation, and teacher support) had significant correlation with mean grades. Rule clarity
and teacher support had a significant negative correlation with mean grades. For teacher perceptions, involvement had a significant positive correlation and teacher control had a significant negative correlation with mean grades: “Thus, both students and teachers perceived classrooms with high average final grades to be higher in involvement and lower in teacher control” (p. 265). In summary, they found that classes with absenteeism were seen as high in competition and teacher control and low in teacher support. Students in these classes felt they were often clock-watchers, that they needed to be careful about what they said, that it was easy to get into trouble in the class, that passing the class was relatively difficult, and that the teacher was fairly strict, and so on (Moos & Moos, 1978; cf. Moos, 1979a).

The CCI, HSCI, Ed, OCI, ESES, and CES are only a few of the many and diverse high-inference measures that evolved from the first measures (i.e., Al, CCI). For example, Halpin and Croft (1963) constructed the Organizational Climate Description Questionnaire, which portrayed the organizational climate of the elementary school through the self-reports of its teachers and principals. Walberg and Thomas (1974) developed a Likert-Teacher Questionnaire and a parallel Observation Rating Scale to investigate open education cross culturally. Steele, House, and Kerins (1971) developed the Classroom Activities Questionnaire (CAQ), which measured the prevalence of activities in the affective and cognitive domains of the classroom. The CAQ was originally a low-inference measure but was modified to a high-inference measure by adding a 4-point Likert scale to the various statements of the CAQ. The findings using the CAQ (with the high-inference modification) suggested that if further replication occurred, empirical ways of identifying the structure of curriculum and instruction were possible (Steele, Walberg, & House, 1974; Walberg, House, & Steele, 1973).

In summary, several more recent studies have focused on assessing the characteristics and impacts of social environments of classrooms (see Moos, 1979a, 1979b; Nielsen & Kirk, 1974; Randhawa & Fu, 1973; Walberg, 1979). However, these studies are not directly related to classroom climate research and therefore beyond the scope of this review. All the high-inference measures mentioned gave an indication of the high-inference measures that have been developed and have been successfully used to assess the characteristics and impacts of the social environments within classrooms.

The My Class Inventory (MCI) and the Learning Environment Inventory (LEI), two high-inference measures, have been used extensively. The MCI and LEI can be machine scored by key punching the item responses and adding the appropriately keyed answers on a computer. Also, both instruments never mention the teacher and therefore do not pose a threat as do some other instruments that explicitly focus on teacher characteristics and behavior (e.g., Flander’s Interaction Analysis System), and it is possible to omit some scales that are not of interest to the researcher (see Chavez, 1980). Moreover, because of the MCI’s and LEI’s simplicity, over 200 American, British, Canadian, French, German, Swedish, Indian, Australian, and New Zealand investigators have used and/or adapted the inventories to fit their specific research needs and interests (Anderson, 1973; Anderson & Walberg, 1974; Haertel, Walberg, & Haertel, 1981; Walberg, 1979).

The My Class Inventory (MCI)

The MCI is a reworked version of the LEI (to be discussed) and is used for elementary level students from 8 to 12 years old. The MCI consists of five nine-item scales in a simple agree-disagree format appropriate for younger children. The five scales are friction, competition, difficulty, satisfaction,
and cohesiveness. Friction is defined as the extent of disagreement, tension, and antagonism in the classroom. Competition is the extent to which the students perceive an atmosphere of competition in the classroom. Difficulty is the extent to which students perceive class activities to be difficult. Satisfaction is the extent to which students like their classroom. Cohesiveness is the extent to which the classroom develops a feeling of intimacy as a result of student interactions (Anderson, 1973).

My Class Inventory (MCI) studies

In contrast to the several studies that have used the LEI as a high-inference measure, only a few studies have used the MCI. Walberg (1969b) first mentioned using the MCI; however, the source within the article was secondary and was mentioned only in passing. Using the MCI as one of several evaluation instruments, Walberg (1970) studied the urban-suburban busing program for desegregating the Boston public schools. He found that the urban elementary school children bused to suburban schools around Boston did not gain in arithmetic and reading achievement more than those school children that remained in urban schools; however, there were significant differences between the two groups on perceptions of their classes. For example, those children that were bused from urban to suburban schools (grades 3—6) rated their classes more satisfying than those children who remained in the urban schools (Anderson, 1973; Walberg, Sorenson, & Fischbach, 1972).

The MCI was also used to learn more about the measures of social environments of the elementary school by examining their ecological correlates—school size, the ratio of boys and girls in the school, and parents’ socioeconomic status (SES)—plus the impact of those variables on the elementary school students’ perceptions (Walberg et al., 1972). The MCI used in this particular study varied from Walberg’s (1970, elaborated in Anderson, 1973) in that the cohesiveness and friction scales were dropped and a goal direction scale was added. The results using univariate regression relationships on the Walberg et al. (1972) study revealed (a) effects of school SES composition on SES group perceptions of competition, (b) effects of SES composition on boys’ and girls’ perceptions of difficulty, and (c) effects of school sex composition on the satisfaction interaction (Walberg et al., 1972). In conclusion, the principal finding demonstrated that “it was possible to detect [effect on the ambience of school as perceived by different kinds of students, particularly with regard to competition and difficulty” (p. 147). This was especially true for lower SES students, which was disturbing to the authors because of the various desegregation plans prevalent at that time.

Talmage and Hart (1977) used the MCI to determine what curriculum issues within mathematics education were distinct from as well as interactive with instructional concerns. Thus, the researchers believed that they “must study the learning environment in addition to [achievement; and must define the specific instructional approaches to fit a variety of school settings” (p. 346). It was hypothesized that a conscious awareness due to the characteristics of investigative teaching (by teachers who had participated in a National Science Foundation [mathematics] Elementary Mathematics implementation program) could be transferred to the classroom. The MCI was used as an independent variable to determine the amount of transfer that could be measured by changes in the classroom learning environment.

It was found that students who were taught mathematics by teachers participating in the NSF program perceived their classes as significantly more cohesive than did students whose teachers were not NSF participants (Talmage & Hart, 1977). Also, the researchers found that the four remaining MCI
scales were affected more by grade level than by group. Less friction was perceived at the end of the school year than at the beginning in both experimental and control classes. Competitiveness dropped significantly over time in the upper grades, while a smaller decline was reported in the primary/control classes. Of interest, the primary/experimental class showed a gain in competitiveness. Also, difficulty seemed to decrease over time. The researchers determined from the results that such mathematics approaches should consider differences in the learners along grade level (Talmage & Hart, 1977).

Talmage and Walberg (1978) used the MCI’s five scales with several other variables as part of an evaluation study to select basal reading series for adoption in a large school district and to explore other factors possibly related to reading achievement. Of interest, the several variables incorporated in this study (see p. 192) did not account for significant differences. Nevertheless, competitiveness in the learning environment proved the only other variable to predict posttest reading scores after pretest effects were removed: the higher the competitiveness perceived by students, the lower the reading achievement scores. Talmage and Walberg concluded that competitiveness was the crucial variable for reading progress in the study. Moreover, they confirmed that if they had not used a high-inference measure (i.e., MCI) as well as other proficiency measures, valuable information about aspects of the natural setting affecting reading achievement would have gone undetected.

Chavez and Cárdenas (1980) modified the MCI to employ all the climate scales except difficulty. In addition, instead of using nine statements with yes-no response options to define each scale, they used five statements, each with a 5-point Likert scale. Many of these statements were reworded for clarity following the pilot study. The study population included sixth grade students in bilingual (Spanish/English) bicultural education programs and nonbilingual education programs in New Mexico. Data collected from students were age; sex; ethnicity; type of classroom (bilingual or nonbilingual); third and fifth grade California Test of Basic Skills (CTBS) achievement scores in language arts, reading, and the CTBS battery total. The purpose of the study was two-fold: to provide data about students’ perceptions (via the modified MCI) of their classrooms; and to determine whether affectivity, as perceived by the students, would enhance the educational progress of students in bilingual/bicultural and nonbilingual classrooms.

Chavez and Cárdenas (1980) found that Chicano and non-Chicano students would achieve at higher levels in language arts, reading, and battery total achievement when the students’ perceptions of the satisfaction and competition scales were high. Also, Chicano students needed to perceive twice the amount of satisfaction to experience the same achievement level in reading and battery total as non-Chicanos, and twice the amount of competition to experience a similar achievement level in language arts as non-Chicano students. Finally, the lower the degree of friction perceived, the higher students’ level of achievement would be. Interestingly, classroom climate variables, via the modified MCI, were significant only as interacting variables within a stepwise regression. This high-inference measure “tended to reflect the inclusiveness of the classroom and its climate, illustrating that the classroom is a dynamic interchange of affective and cognitive forces forged with personalized and socio-psychological phenomena” (p. 313).

Fraser and Fisher (1982) also modified the MCI. They deleted 7 of the original 45 items, which greatly improved the reliability of each of the five scales. When the 7 items were removed, the reliability estimate was found to be .63 for cohesive ness, .68 for friction, .68 for difficulty .76 for satisfaction, and
The modified MCI was administered to 2,305 seventh grade students in 100 science classrooms in 30 schools throughout Australia. Also, two cognitive and one affective outcome measures were used as pre-posttest indicators to determine relationships between classroom environment (climate) and learning outcomes. The purpose of this study was to determine the relationships between students’ learning outcomes and their perceptions on the modified MCI. Fraser and Fisher found that the high-inference scales of the MCI differentiated significantly among the perceptions of students in different classrooms, and that the internal consistency and discriminant validity of the scales were satisfactory.

The Learning Environment Inventory (LEI)

In the late sixties a series of research and evaluation studies were enacted to determine the success of Harvard Project Physics (HPP). HPP was an experimental course implemented in various secondary school physics classes across the United States. HPP used a variety of new instructional media emphasizing the philosophical, historical, and humanistic aspects of physics (Walberg & Anderson, 1968). To determine the efficacy of HPP, Walberg (1969b) devised an instrument called the Classroom Climate Questionnaire (CCQ). The CCQ’s dimensions were based on Hemphill and Westie’s (1950) Group Dimensions Descriptions Questionnaire (GDDQ), which measured general characteristics of adult behavior. The items of the GDDQ, however, were “inappropriate for the classroom group, but they did suggest a number of dimensions possibly related to learning. On the basis of [those] dimensions and others hypothesized to be relevant, 90 items were written for the Classroom Climate Questionnaire” (Walberg, 1969b, p. 444). Despite the validity of the CCQ, psychometric studies showed that the items were unreliable and redundant; thus, work began on the LEI.

According to Walberg (1969b), the Getzels and Thelen (1960) theory of the class as a social system had proven useful in earlier research. The Getzels-Thelen theoretical treatise suggested that in classrooms, personality needs and role expectations interacted to form a climate in which group behavior, including learning, could be predicted. Hence, the Getzels-Thelen (1960) treatise was used as a theoretical guide for constructing the new LEI.

The LEI was designed to measure the social climate of learning in a classroom as perceived by the pupils (Anderson, 1973). The original form of the LEI contained 14 scales, and in the 1969 revision a 15th was added. The climate dimensions or scales are (a) cohesiveness, (b) diversity, (c) formality, (d) speed, (e) environment, (f) friction, (g) goal direction, (h) favoritism, (i) cliqueness, (j) satisfaction, (k) disorganization, (l) difficulty, (m) apathy, (n) democraticness, and (o) competitive ness.

In selecting the 15 climate dimensions, an attempt was made to include as scales only concepts previously identified as good predictors of learning, concepts considered relevant to social psychological theory (Getzels and Thelen’s, 1960, theory), concepts exemplifying useful theory and research in education, or concepts intuitively judged relevant to the social psychology of the classroom. (Anderson, 1973, p. 4).

Each of the scales contains seven items. To insure homogeneity of item content, all 34 items that had been misclassified by the judges were either revised or replaced. Six of those items with poor scale correlations were modified also. All items described typical school classes, and the respondent
expressed his or her agreement or disagreement with each item on a 4-point Likert scale. In addition to measuring the social climate of learning in a class as perceived by its pupils, the LEI was devised to assess a student’s perceptions of his or her class. That is, if the concern was for variables such as pupil sex, self-concept, or personality, then the individual scores could be used to assess the perceptions of the individual student in his or her classroom.

Learning Environment Inventory (LEI) studies

The first studies to be reviewed used data from 1-IPP. The subsequent studies reviewed used various and distinct data bases. Walberg and Anderson (1968) investigated whether individual students’ satisfaction with the climate of a classroom would enhance learning. Their hypothesis that student achievement and interest in the subject could predict structural and affective aspects of classroom climate was confirmed; the climate measures of the personal relations between class members did predict learning. The variables that accounted for 12 correlations with the criteria (Hennon-Nelson IQ, Physics Achievement Test, Science Process Inventory, Semantic Differential for Science Students, and Pupil Activity Inventory) were personal intimacy, friction, and satisfaction: “Thus, it is not the identification with the group that correlates with learning but the perception that the class is personally gratifying and without hostilities among the members” (Walberg & Anderson, 1968, p. 148).

The following year Anderson, Walberg and Welch (1969) explored potential determinants of the social climate itself in an effort to gain insight into the manner in which classroom climates evolved. Anderson et al. (1969) administered the LEI to students of HPP where the teachers of those students belonged to one of three groups: inexperienced experimental group, experienced experimental group, and experienced control group. It was hypothesized that (a) students would perceive significant differences in the learning climate of the three teacher groups; (b) students’ perceptions of the two experimental groups would be similar when compared to the control; (c) the experimental classes would be perceived as more diverse than the control; and (d) students would perceive the experimental course to be less difficult because of the diversity of instructional approaches. Anderson et al. (1969) found significant differences in the climates of classes receiving different course and teacher treatments and that teacher experience with a course had less effect on climate than the course itself. These findings supported hypotheses a and b. The data also supported hypotheses c and d. Of interest in this study was that “course effects appear to account for appreciably more variance than teacher selection and experience in teaching a course” (p. 324).

Because the above study found that course effects as perceived by students via the LEI accounted for more variance than the teacher and his or her experience in teaching, Walberg (1969b) questioned if aside from course effects, social environments influenced classroom learning. After correlating the LEI scales with six posttests and an IQ test taken by the students, Walberg asserted that the “measures [i.e., LEI] of the social environment of learning as perceived by students predict learning criteria before and after relevant control variables such as posttest and IQ test are statistically removed from the criteria” (p. 448). Moreover, in replicating his own work, Walberg (1969a) affirmed that student perceptions of the classroom learning environment could be measured reliably with a high-inference measure such as the LEI, and that environmental measures were valid predictors of learning.
Another study using HPP data was conducted by Walberg and Ahlgren (1970). They investigated the predictability of the environment scales from a variety of teacher, student, and class characteristics. Six cognitive tests (pre and post), the LEL, seven personality scales, an IQ test, and a biographical questionnaire were administered within a span of a year to obtain some measurable characteristics. In brief, based on the various correlational relationships that could result from all the above criteria, “the present study show that the classroom climate [be predicted from a number of variables (i.e., cognitive test, personality measures, etc.)]” (p. 165). They noted that it was not their purpose to introduce complex interpretations of all the significant findings until this study could be replicated; however, the study’s data did imply (because of the various measures that were used) that “cognitive and non-cognitive learning might be affected by manipulation of variables that affect classroom climate” (p. 165). Moreover, positive information on predictors of classroom climate, the possibility of manipulating the predictors and demonstrating a valid way for assessing climate, replicating and extending the method into other causes and classrooms all became viable considerations.

Thus far, this review has shown by the Walberg and Anderson (1968), Anderson et al. (1969), and Walberg (1969a, 1969b) studies that high school social environments of HPP demonstrated validity of student perceptions of their classrooms, vis-à-vis the LEI, in predicting cognitive, affective, and behavioral learning, and that the social environments of classes randomly assigned to different physics classes differed significantly. Furthermore, in a correlational study Walberg and Ahlgren (1970) found that the social climate in high school HPP classes could be predicted from cognitive and noncognitive pretests, student biographical characteristics, teacher experience, and other variables. Hence, there was research evidence that the use of a high-inference measure, the LEI, was indeed useful for investigating various occurrences in high school classrooms, cognitive and otherwise, with reliability.

Because the above research from Harvard Project Physics had established that students’ perceptions could be measured reliably using the LEI, researchers initiated studies in other areas with various types of classrooms under diverse conditions. Anderson (1971) studied the effects of course content and teacher sex on the social climate of learning. Differences in four high school classroom climates (science, mathematics, French, and humanities) were found. The data illustrated that French and humanities classes were “easier going” compared to the “hard” sciences. No relationship was found between teacher’s sex and pupil’s perceptions of the learning climate in the classes.

In an evaluation of mathematics achievement, O’Reilly (1975) examined the relationship between certain social and personal characteristics of pupils, the climate (as measured by the LEI) of 9th and 10th grade mathematics classes, and students’ mathematics achievement. Data were collected from 1,100 students in 48 classes equally divided between the 9th and 10th grades from 12 secondary schools. He found that the scales of the LEI accounted for 67% of the variance in achievement. Variables such as parents’ education were related to the LEI scales of goal direction, difficulty, and apathy; “dislike for schooling” was related to the favoritism, apathy, and democratic scales; and student achievement was positively related to the cohesiveness, environment, satisfaction, difficulty, and democratic scales. Five scales were negatively related: friction, favoritism, cliqueness, disorganization, and apathy (O’Reilly, 1975). Because of these data, O’Reilly (1975) suggested that a supervisory program for teachers could be designed for monitoring classroom climates and to assist those teachers to change the climate accordingly to affect achievement positively.
Most studies using the LEI had been done in urban or suburban settings; Randhawa and Michayluk (1975) changed that by examining the learning environments of classrooms in grades 8 and 11 in rural as well as urban settings. They studied 47 classrooms from the rural setting and 49 from the urban setting with an approximately even distribution between 8th grade and 11th grade students. Mathematics, science, social studies, and English courses were represented. Randhawa and Michayluk found that course content did not affect the learning environment of the classrooms. However, “results did indicate that course content affected only the cohesiveness of the classroom significantly” (p. 270). They also found that “mathematics and social studies classrooms had significantly higher perception means on cohesiveness than English classrooms” (p. 271). In addition, “rural classrooms were characterized by cohesiveness, cliqueness, disorganization, and competition; whereas urban classrooms were characterized by environment, difficulty, and satisfaction” (pp. 272—273). Based on these findings, Randhawa and Michayluk asserted that “urban classrooms offer better physical environments, are intellectually stimulating, and seem to meet the needs of the learners in such a way that they [the students] perceived their learning experiences sufficiently satisfying. Urban classes would seem to have better learning orientation than the rural classes” (p. 277).

As noted, the LEI had been used for a variety of research and evaluative purposes. All the same, cross-cultural generalizability of students’ perceptions to the social environment of learning had been omitted. Walberg, Singh, and Rasher (1977) corrected that omission by conducting an experiment in the State of Rajasthan, India. The LEI (translated into Hindi) and IQ measures were correlated with the mean end-of-course achievement scores of 166 groups of “studious” and “nonstudious” members of 83 general science classes and 134 similar groups in 67 social studies classes, which were randomly sampled. Walberg et al. found that the patterns of the correlations for general science and social studies were generally comparable. Also, multiple correlations of IQ plus all the LEI scales went considerably beyond the correlations with IQ alone in both the general science and social studies groups.

In another study, the LEI was used to investigate if the stability of student perceptions of the classroom climate would change over a short time (Lawrenz, 1977). For this investigation, Lawrenz (1977) proposed three questions: “(I) Are class mean scores for the LEI scales stable over time? (2) Are individual student scores for the LEI scales stable over time? and (3) Do individual students who take the LEI more than once have different perceptions of the classroom environment than their classmates who take it only once?” (pp. 78—79). In response to these questions, Lawrenz found that LEI class mean scores were stable over time; LEI individual scores were stable over time; and no difference was found to exist between students who took the LEI once and those who took the LEI twice. Lawrenz concluded that student perceptions of the classroom environment as measured by the LEI were consistent over time; moreover, previous research on the classroom learning environment was strengthened because the implicit assumption of stability had been verified.

Coattailing on the Lawrenz (1977) study, which verified classroom climate stability over short periods, Welch (1979) contended that stability in educational environments over longer periods had not been investigated. Welch’s (1979) three fold purpose was (a) to determine if students’ perceptions, using the LEI, of the learning environment in secondary science and mathematics classes changed during a 4-year period (1972—76); (b) to determine if students perceived a different learning environment in science and mathematics classes; and (c) to determine if learning environments of the junior high differed from those of the senior high.
Welch’s (1979) study (based on a stratified random sampling technique of 15 western states) included 563 classes in 1972 and 558 classes in 1976 with similar subject matter treatments in those 4 years (i.e., science and mathematics). Teachers administered a modified version of the LEI (10 scales instead of IS) in early spring 1972 and again in 1976 to all students. Class means from the LEI were obtained and submitted to a three-way multivariate analysis of variance. The main effects due to year (1972—76), subject (science, mathematics) and grade level (junior, senior high) indicated that the 10 LEI scales taken together over the 4 years were significant. Welch suggested that time, curriculum, and age would have an effect on the students’ perceptions of the learning environment. Also, there were significant multivariate interactions between grade (junior, senior high) and subject (science, mathematics), suggesting that on some LEI scales the junior-senior high perception changes were different for science and mathematics. However, “no interactions were noted for the effects of time—the primary concern of this study” (p. 179). Other results suggested that a more conservative environment and more student satisfaction were emerging between 1972 and 1976. Math classes were perceived to be goal-directed, difficult, and democratic, whereas science classes were disorganized, formal, and cliquish with students perceiving more friction and favoritism. Next, student- and activity-centered environments were more prevalent in junior high classes than in senior high classes. Finally, the LEI results of this study lend “strong support to the claims by teachers and administrators that the educational scene is shifting; [is] the perceptions of students seem to reflect the gathering momentum of a back-to-basics movement” (p. 177).

Other important studies reviewed are the investigations by Fraser (1978, 1979). Fraser (1978) modified the LEI for use in individualized junior high school classrooms. The modified LEI formed a battery of nine classroom climate scales suitable for science classrooms and the seventh grade (see pp. 128—131 for detailed description of modification). In addition, the LEI provided dependent and independent variables to be evaluated in the individualized seventh grade classroom. The LEI as a dependent variable would reveal differences in student perceptions of the classroom learning environment. The LEI as an independent variable would predict student cognitive and affective learning outcomes. Using the LEI as a dependent variable, Fraser (1979) found the experimental science students’ perceptions of learning environment 6 months after the beginning of the school year significantly more favorable than control students’ perceptions on several dimensions. Using the LEI as an independent variable, Fraser (1979) found that the experimental science students experienced more favorable changes in attitude toward science than control students. “It was also found that all dimensions of the learning environment considered together accounted for a significant increase in variance accounted for in five of the science learning posttests” (p. 232).

The final two studies reviewed speak well of the research sophistication that has evolved using high-inference measures for studying classroom climates. Cort (1979) combined the MCI, the LEI, and the Classroom Activities Questionnaire (CAQ, discussed earlier) to create a relatively new high-inference measure. Scales from the MCI, LEI, and CAQ were incorporated to form the My Social Studies Class (MSSC), which was used as one of the nine instruments for a large-scale evaluation of an upper elementary social studies program, Man: A Course of Study (MACOS). The nine instruments produced three types of variables: teacher process, student process, and climate variables. This review will deal only with the climate variables and their relation to the other variables.
Cort (1979) found that climate was positively related to teacher attitudinal characteristics (teacher process variables) and to how students perceived emphasis and activities (student process variable) in their classes:
The more a class perceived itself as interested in problem solving and as creative ..., the better the perceived climate. The lower teachers scored on traditional educational views and on approval of controlling behavior, the better the classroom climate. The less the class was perceived as traditional (emphasis on grades, right answers, facts, individual work, and so forth), the better the climate. (p. 254)

Though this study dealt more with the evaluation of MACOS, the unique combination of the LE! and MCI, high-inference measures that focus on learning climate, and the CAQ, a high-inference measure that focuses on cognitive processes, is worth reporting for two reasons: (a) By his obvious interest in students’ perceptions of MACOS and the climate it may produce, Coil (1979) implicitly emphasized that serious evaluation studies must employ as one of its evaluation components students’ perceptions and the interaction among other evaluation components (in this case, the MACOS evaluation used nine instruments divided into three evaluation components: teacher process, student process, and climate); and (b) the combination of the high-inference measures (i.e., LEI, MCI, and CAQ) illustrated the flexibility the LEI and MCI have for being modified and the usefulness these measures have in determining the effects the learning climate may have on students.

The final and essential study to be reviewed reflects the consistency of the LEI as a high-inference measure as well as the importance of investigating classroom climates. The study is a meta-analysis by Haertel, Walberg, and Haertel (1981). They focused on the “predictability of end-of-course cognitive, affective, and behavioral learning from mid-course social-psychological perceptions [LEI measure] with and without statistical control for beginning-of-course measures or ability or both” (p. 28). The quantitative synthesis yielded 12 LEI studies of 10 data sets reporting 734 correlations calculated from a total of 17,805 students in 823 classes from the United States, Canada, Australia, and India. Three questions were proposed:

1. Which perceptions [of the 15 subscales within the LEI] are most productive?
2. What learnings [cognitive affective, or behavioral] are most predictable? and
3. How does the predictability vary across such factors as grade levels of students, subject matter, and methodological characteristics of the [12] students? (p. 28)

In response to these questions Haertel et al. (1981) found the following:
I. The positive subscales were cohesiveness, satisfaction, task difficulty, formality, goal direction, democracy, environment, and competition; the negative subscales were friction, cliqueness, speed, apathy, favoritism, disorganization, and diversity.
2. Perceptual aspects of the social-psychological environment of learning were consistent in the direction of their relationship to cognitive, affective, and behavioral learning outcomes, with or without statistical controls for ability, pretest, or both.
3. Correlations differed significantly across perceptual scales, units of analysis, nations, and grade levels.

It is important to note that the final differences need to be investigated further. Nevertheless, the theoretical plausibility and the incremental predictive validity of the LEI scales when measuring student
to subgroup to class to school as units of analysis (Fraser, 1981) offer a positive approach to future research and evaluation procedures (Haertel et al., 1981).

In summarizing the second part of this review, research using high-inference measures for investigating environmental phenomena began with the Activities Index. From the Activities Index, several “environmental indexes” were developed based on Murray’s Need-Press Model. These environmental indexes were self-administered questionnaires asking high school and college students about various dimensions existing within those institutions. The responses to the various dimensions were regarded as “student perceptions” of the climate or learning environment and were useful in predicting achievement and contributing to the understanding of educational processes.

Once launched, other high-inference measures were constructed and employed to study school and classroom climates. This review illustrated the diversity of high-inference measures in conjunction with the efficaciousness of those measures for studying the characteristics and impact of learning climates in classrooms. Two well-established high-inference measures are the MCI and the LEI. The MCI, although not that prevalent in the research literature, has had substantive use. The LEI has been used in an inexhaustible number of settings. The two measures were described, and the results of several studies were reviewed chronologically, which illustrated how the early classroom climate research using high-inference measures affected itself cumulatively during almost three decades. Moreover, an outstanding feature of both the MCI and LEI is the flexibility of the scales, which were manipulated as dependent or independent variables. When the researchers manipulated the scales as dependent variables, they could determine the effects different curricula would have on students and if, indeed, the students perceived classroom climates differently. When the researchers manipulated the scales as independent variables, they could determine which scale variables would best predict cognitive and/or affective learning among students. Finally, in correlational conjunction with other variables (e.g., subject matter, achievement, rural-urban, cross-cultural, etc.), the MCI and LEI, in particular, consistently exemplified flexibility, longitudinal stability, and reliability.

**Conclusion**

This review documents the historical genesis of classroom climate research and high-inference measure research. It also shows that classroom climate research using low-inference measures was well-established by the early sixties, but classroom climate research using high-inference measures did not become prevalent until the mid-sixties. Essentially, this review brings together disparate stands of research that had remained isolated. High-inference measures developed in the past three decades were reviewed, culminating with a discussion on two well-established high-inference measures used extensively to study classroom climates—the My C/ass Inventory and the Learning Environment Inventory.

The MCI and LEI have indeed established a strong research tradition in the last decade (Fraser, 1980, 1981; Walberg & Haertel, 1980). This tradition is appropriately illustrated by the recent meta-analysis conducted by Haertel et al. (1981). The meta-analysis poignantly shows that the predictability of students’ cognitive, affective, and behavioral outcomes are related to students’ perceptions of psychosocial characteristics in classrooms. In juxtaposition to the above, Fraser (1981) affirms that “the strongest research tradition during the last decade of classroom environment [synonymous with “classroom climate”] research has involved investigation of the predictability of students’ cognitive and
affective learning outcomes from their perceptions of classroom environments” (p. 46). Fraser (1981) gives particular emphasis to the use of classroom environment variables in curriculum evaluation. He strongly suggests that curriculum evaluation should more often include classroom environment dimensions as criterion and predictor variables. In addition, “classroom environment information can provide a valuable basis for reflection upon and subsequent improvement of curriculum and classroom practice” (p. 70).

The various classroom climate studies reviewed in this paper culminated with Haertel et al.’s meta-analysis. This coupled with the noteworthy insights of Fraser, elucidates the validity of classroom climate research using high-inference measures. Also, this review contradicts the notion that school climate research per se will reveal valid climate variables (Anderson, 1982) because a variety of classroom climates, positive and negative, may exist in the same school. Moreover, this review lends support to the idea that the effective schools research movement should incorporate the use of high-inference measures to study classroom climates in identified effective schools. Once implemented, this approach may assist policy-makers to avoid research problems now evident in the effective schools movement (Cuban, 1983).

References


Getzels, J. W., & Thelen, H. A. (1960). The classroom group as a unique social system. In N. B. Henry (Ed.), The dynamics of instructional groups, sociopsychological aspects of teaching and learning:


RUDOLFO CHAVEZ CHAVEZ, Assistant Professor, California State College, School of Education, 9001 Stockdale Highway, Bakersfield, CA 93311-1099. Specializations: Classroom climate research, teacher training research, bilingual/multicultural education, elementary education.

1 Cornell, Lindvall, and Saupe’s (1952, cited in Medley & Mitzel, 1963) study was also a necessary link; however, the primary source was unavailable. Consequently, this study is not reviewed in the paper. Cornell et al. attempted to measure differences in classrooms as a means of characterizing differences of school systems (Medley & Mitzel, 1963). The instrument developed was the Classroom Observation Code Digest (COCD), which contains eight dimensions: (a) differentiation, (b) social organization, (c) initiative, (d) content, (e) variety, (f) competency, (g) and (h) classroom climate, as reflected in the behavior of the teacher and the behavior of the pupil, respectively. A pilot study was conducted using the COCD. Thirty-two classrooms, visited three times by six observers in teams of two, recorded the behaviors independently for 30 minutes. The reliability coefficients indicated that high objective observer agreement existed, but not how well the scores discriminated different classrooms from one another.