

A DEVELOPMENTAL ANALYSIS OF RATING BEHAVIOR

by

KARL W. KUHNERT

B.A., The Pennsylvania State University, 1977

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1981

Approved by:


Major Professor

SPEC
COLL
LD
J668
J4
1981
K84
C.2

Acknowledgements

My sincerest thanks is expressed to the members of my supervisory committee, including Drs. Frank (Skip) Saal, Ronald Downey and Leon Rappoport, for their direction, patience, and above all, uncompromising support in seeing this project through its completion.

I especially would like to thank my thesis advisor, Skip, whose contributions to my experience in graduate school, both intellectually and personally remain unchallenged. Not only were his psychological, practical, and critical literary suggestions throughout this project invaluable toward its completion, but his wit, understanding, and confidence in me along the way served to effectively temper even the most difficult of times.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	i
LIST OF TABLES	iii
INTRODUCTION	1
Development	3
The Organismic-Developmental Orientation	4
Developmental Assumption	4
The Holistic Assumption	5
Rating Behavior	6
Cognitive Structure	8
Cognitive Complexity Perplexity	10
The Pervasiveness of Cognitive Complexity	11
Zajonc's Method	13
Summary	15
METHOD	16
Participants	16
Procedure	16
Cognitive Complexity	16
Performance Ratings	17
RESULTS	19
DISCUSSION	27
Development of Cognitive Complexity	27
Differentiation and Hierarchic Integration	30
Cognitive Development and the Psychometric Characteristics of Ratings	32
CONCLUSION	34
FOOTNOTES	36
REFERENCE NOTES	37
REFERENCES	38
APPENDICES	

LIST OF TABLES

- Table 1: Class Analysis as a Function of Time.
- Table 2: Mean Scores as a Function of Time.
- Table 3: Intercorrelations Among Differentiation, Hierarchic Integration and Halo.
- Table 4: Correlations as a Function of Class, Academic Level and Age.

A Developmental Analysis of Rating Behavior

Performance appraisal judgments have continued to interest industrial psychologists for more than five decades; their widespread use and importance have been recently described in considerable detail (Landy & Farr, 1980; Saal, Downey, & Lahey, 1980). Most of the research on judgmental measures of performance has focused on a limited number of variable classes including ratee characteristics, both psychological and biodemographic types of rating scales and scale format characteristics, and the purposes for which ratings are solicited. Unfortunately, even after all these variables have been considered, the light shed on the rating process has been somewhat less than blinding. Consequently, a more recent trend in rating research has been to concentrate on the rater as an active processor of information, and on how the rater's cognitions interact with other facets of the rating situation.

A rater characteristic that is receiving increased attention is cognitive complexity. Although there is some confusion as to the exact definition of the term, it is generally (Adams-Webber, 1979) held that cognitively complex persons are prone to make finer distinctions among dimensions of complex stimuli, whereas cognitively simple persons make relatively grosser discriminations of the same stimuli. Schneier (1977) examined the effects of both the rater's cognitive complexity and the cognitive demands of various rating scale formats on the psychometric properties of ratings and the rater's perceptions of those ratings. His results suggested a "cognitive compatibility theory" of rating behavior: When the raters' cognitive complexity was compatible with the scale format's complexity, ratings were characterized by less leniency and less range restriction, and raters reported increased satisfaction with, greater confidence in, and a preference for that particular format. Halo was inversely related to

cognitive complexity regardless of rating scale format. Although these results apparently support a cognitive-process orientation to rating behavior, replication of these results has not been forthcoming.

Recent studies (e.g., Bernardin & Boetcher, Note 1; Lahey & Saal, in press) failed to support the relationship linking cognitive complexity and scale format complexity. Using an identical measure of cognitive complexity and a single rating scale format, Bernardin & Boetcher found no significant differences between complex and simple raters with respect to leniency and halo measures. Moreover, Lahey & Saal, using three different measures of cognitive complexity and four rating scale formats, found no systematic differences between complex and simple raters in tendencies to exhibit leniency, halo, and range restriction, nor in raters' confidence in their ratings. Further, their most important finding was the absence of any cognitive complexity x scale format interactions. Since Schneier's (1977) results are becoming widely cited, and are proving instrumental in encouraging other researchers to investigate cognitive capabilities of raters, it behooves us to examine these studies in light of one another, and to design additional empirical studies with an eye toward reconciling seemingly contradictory results.

A different approach, emphasizing process considerations, relies heavily on a developmental framework. Without questioning the underlying assumptions, it has been tacitly assumed that a state of equilibrium, stability and rest is a more realistic picture of the rater than a state of upheaval, uncertainty and change. But the history and development of a rater cannot reasonably be ignored. Landy & Farr's (1980) "process model implies that the rater's experience with ratings affects the validity of those ratings." They further stated: "We know little or nothing about the effects which decisions based on current ratings have on future ratings. Research in this area is long overdue" (p. 101).

The purpose of this paper is to contribute to a reconceptualization of the performance appraisal process from a developmental perspective, namely the organismic-holistic orientation (Werner, 1957). This framework simply states that "organisms are directed toward a series of transformations reflecting a tendency to move from a state of relative globality and undifferentiation toward states of increasing differentiation and hierarchic integration" (p. 126). It is through differentiation and intergration that the criteria for development are established. If raters, with increased experience, display increased differentiation and hierarchic integration in their perceptions of ratees, then development must be acknowledged as a phenomenon that contributes meaningfully to the rating process.

If raters' cognitive complexity does develop over time, and if Schneier's (1977) cognitive compatibility model is accurate, then passage of time should be associated with systematic changes in the psychometric characteristics (e.g., halo) of raters' appraisals. That is, if it can be empirically demonstrated that an individual's complexity develops over time, we shall be in a position to evaluate Schneier's model longitudinally, instead of from the more common cross-sectional perspective. In order to view rating behavior from a developmental perspective, the first step is to define what is meant by the term development.

Development

The word "development" expresses protean meanings. It has been of tremendous heuristic value in various disciplines such as biology, child psychology, economics, anthropology, and so on. In addition, development is often used in non-scientific areas such as literature, art, and drama. Because the word is applicable to a wide range of phenomena, from the nascent economic proposals of Karl Marx (1970) to the tragic heroes of William Shakespeare (1964), it is difficult to agree on a common definition.

Development usually, although not always, denotes a more or less continuous process which involves progressive changes from a simple to a more complex structure or organizational pattern over time. Discussions of development commonly include as essential the ideas of: (1) an organism conceived as a living system; (2) time; (3) movement over time toward complexity of organization; and (4) "hierarchization," or the integration of parts of systems into larger units or "wholes" (Harris, 1957). Moreover, development is often reserved for changes which are not only irreversible, but which are creative modes of organization not previously manifested in the history of the developing system. Development, then, is a concept which champions a way of understanding behavior in a multifaceted manner.

The Organismic-Developmental Orientation

The study described below is grounded in the organismic-developmental framework of Werner and Kaplan (1963). In order to apply the organismic-developmental perspective to industrial psychology, it is first necessary to discuss two of its underlying assumptions.

Developmental Assumption. Werner and Kaplan (1963) have defined development in terms of their orthogenetic principle. Organisms, according to this principle, are directed toward a series of transformations "reflecting a tendency to move from a state of relative globality and undifferentiation toward states of increasing differentiation and hierarchic integration" (p. 126). This central point of Werner and Kaplan distinguishes "genetic" development from other changes that may occur in an organism over time.

The importance of a genetic explanation was noted by B. F. Lamov:

"To understand any mental phenomenon one must view it in the process of development; indeed, that is the only way it can be understood. Development here, of course, is not seen as some monotonous process of mere quantitative changes, but as a process in the course of which contradistinctions emerge and are resolved and qualitative transformation takes place in the system of mental phenomena" (1978, p. 73-74).

As such, genetic development does not merely consist of the continuous refinement or compounding of cognitions and behaviors, but is characterized by the emergence of novel relationships between and among these concepts.

One last characteristic of development concerns the "fate" of genetically earlier modes of functioning after higher functions and forms have emerged. Through the process of genetic development, it is maintained that lower levels of functioning are not lost with the attainment of higher levels; they are retained and often used by the organism. For example, the ability to perform calculus does not submerge or preclude the ability to perform addition. Or, when an individual is confronted with a novel or difficult task, one often returns to a more primitive form of functioning before progressing to higher operations.¹ Werner and Kaplan (1963) refer to this manifestation as the genetic principle of spirality.

The Holistic Assumption. In general, "the holistic assumption is opposed to any view that would treat an element (for example, a moment or momentary experience) as if it possesses a fixed structure and meaning, irrespective of the whole or context of which it is part" (Werner & Kaplan, 1963, p. 132). Thus, the value of relating to material isolated from its context is questionable. For example, researchers in personnel psychology have exhaustively explored the end product of an individual's ratings (e.g., mean rating) or "errors" in ratings (central tendency, halo, leniency, etc.), but have not with the same vigor investigated how a person's ratings relate to their own past or expectations of the future.² The failure has been one of not recognizing that most ratings are all but meaningless for the individual unless they are understood in the context of that individual's past and anticipated future. In short, according to the holistic assumption, "every behavioral act, whether outward bound movement or internalized cognitive operation, gains its significance and

status in terms of its role in the overall functioning of the organism" (Werner and Kaplan, 1963, p. 132).

Taken together, the developmental and holistic assumptions may provide a basis for understanding rating behavior as a dynamic, non-fragmented process. Recognition of the "genetic" changes in raters over a period of time may provide a way of understanding the changing nature of an individual's ratings. Before providing such a framework, however, a brief review of the role of rating behavior in the field of personnel psychology is warranted.

Rating Behavior

From the earliest writings, the problem of how best to judge another person's performance has played a major role in the industrial psychological literature (Wherry, 1952). Today, performance evaluations are, or at least should be, an integral part of any business, company or organization. Employee counseling, selection, merit raises, promotions and training are some of the activities that are potentially affected by these judgmental indicies. Almost every aspect of personnel psychology is dependent on ratings in some form.

The majority of research on judgmental measures of performance has focused on the effect of a fairly wide range of variables on ratings (Landy & Farr, 1980). These have included rater characteristics (e.g., sex, age, race, education, experience, etc.), ratee characteristics (sex, age, race, performance level, etc.), types of rating scales (e.g., graphic, behaviorally anchored, forced-choice, mixed standard, etc.), and scale format characteristics (e.g., definitions, number of response categories, types of anchors, etc.). Almost all of this research has focused on the "content" of the rating situation, and relatively little attention has been devoted to the "process" aspects of performance evaluation.

The other class of judgments that has attracted a great deal of attention among industrial psychologists, and which depends heavily on ratings, is interview judgments, particularly as they contribute to selection decisions. That an extensive literature exists on this topic is not surprising, since it has been estimated that well over 90% of all organizations use interviews at some point in their selection procedures (Landy & Trumbo, 1980; Ulrich & Trumbo, 1965). As was the case in the performance appraisal literature, research on interviews has probed a wide range of topics. Major reviews in the area (Mayfield, 1964, Ulrich & Trumbo, 1965; Wagner, 1949) have focused on the reliability and validity of interview decision, the format of the interview (structured vs. unstructured), and the types of facts that are best assessed in the interview (interpersonal skills, motivation, etc.). A number of researchers have focused on another content variable, the effects of the type of rating scale, on interview judgments (Maas, 1965; Vance, Kuhnert & Farr, 1978). Other researchers have focused on what may be thought of as process variables. Included here would be primacy-recency effects (Farr, 1973), contrast effects (Lancy & Bates, 1973; Wexley, Yukl, Kovacs & Sanders, 1972), the influence of positive and negative information (Hollman, 1972) and the relative importance of several different content dimensions on interviewers' decisions (Hakel, Dobmeyer, & Dunnette, 1970).

A theme, then, which is beginning to pervade much of the literature on rating behavior, in the context of both selection interviewing and performance appraisal, is that rating activities constitute information processing tasks. Raters are exposed to a continuous stream of information with varying degrees of relevance to the purpose at hand over some period of time (generally lengthy for performance appraisal, and of generally short duration with interviews).

The rater's task is to extract from this large body of data those aspects of greatest relevance, to assimilate them, and to arrive at some evaluative judgment of others, typically in the form of a numerical rating. Viewing evaluation or rating tasks as a developing process serves to focus greater attention on the rater than has heretofore been the case. Furthermore, this viewpoint focuses on the active, cognitive processing of the rater rather than on static bio-demographic characteristics. Interesting questions arise as a result of viewing the rater as a processor of information: Does a rater's cognitive structure/organization develop over time? More importantly, is this cognitive restructuring in any way reflected in a rater's ratings? These questions are crucial, since the rater mediates the "incoming" information and the eventual rating decision.

There are extensive literatures regarding the effects of informational characteristics (rater, ratee, context, source, order, etc.) on rating judgments, and other variables that are directly a part of the rating situation (type of scale, format characteristics, purpose, etc.). Prefatory to that literature, however, should have been an understanding--or at the very least a recognition--of the changing, developing nature of a person's ratings; that is, a consideration of the individual as he/she interacts with a changing world. The history of rating behavior is not an auxiliary aspect of the rating situation, but rather forms its very base. Before completing the task of relating development to the rating process, an historical diversion into the literature on cognitive structure is appropriate.

Cognitive Structure

Perceptions reflect a complex blend of our own characteristics and those of the other person. When forming impressions of others, we more or less consciously take note of their physical appearance, including indications of

social role and class, their actions, voice, expressive movements, and all other cues that may be psychologically informative or meaningful. From such evidence, we formulate a fairly detailed perception of what he/she is like. These notions, often implicit, refer to the observer's implicit personality theory (Bruner & Tagiuri, 1954).

Among the first to study implicit theories of personality was Cronbach (1955). He analyzed the correlations of judges' trait ratings for several stimulus persons and was able to show that different judges were employing different underlying dimensions. This implicit theory refers to: (a) the categories that the person employs to describe the range of abilities, attitudes, interests, physical features, traits and values that one perceives in oneself and others; and, (b) the beliefs that the person holds concerning which of these perceived characteristics tend to go together and which do not (Seymore and Jones, 1972). Furthermore, there is evidence that people differ consistently in the traits they use, and in the weights/importance they give to those traits in their perceptions and thoughts about others (Rummetviet, 1960). For example, some people tend to describe other persons in terms of "inner" or psychological traits, while others tend to focus on and describe physical traits (Bieri, 1966; Sarbin, 1954; Wolin, 1956).

If we are to assume that we create and use trait relationships, and that as individuals we differ in the ways we "see" other people, then perhaps our inferences reveal more about us than they do about the stimulus people. Kelly (1955), in his theory of personal constructs, went one step further by contending that "if we are to see a person's psychological processes operating within a system which he/she constructs, we need also to account for the evolution of the system itself in a similarly lawful manner" (p. 77). Kelly argued that each individual evolves a hierarchically organized subsystem of

personal constructs to interpret and predict events. This hypothesis contains direct parallels with the developmental models of both Piaget and Werner. Piaget (1960), for instance, posited that psychological structures evolve through a process of differentiation and reintegration of operational "schemata" at increasingly higher levels of abstraction.

Cognitive Complexity Perplexity. As a result of Kelly's work, a major focus of research has been the formal analysis of conceptual structures, that is cognitive complexity (Mayo & Crockett, 1964; Scott, 1963; Schneier, 1977; Zajonc, 1960). It was Bieri (1955), however, who first introduced into personal construct theory the idea of analyzing the level of development of an individual's "system of cognitive dimensions for construing behavior" in terms of its relative degree of differentiation.

In general, participants who tend to sort figures in a similar way on several constructs are designated as "cognitively simple" (undifferentiated); whereas those subjects who tend to sort figures differently on every construct are designated as "cognitively complex" (differentiated) by Bieri's method of measuring individual differences in cognitive structure. His specific hypothesis was that cognitive complexity, measured by his own index of differentiation, would correlate positively with accuracy in predicting the behavior of other people.

Although Bieri's hypothesis received very little direct empirical support (Adams-Webber, 1979), subsequent researchers continued to redefine cognitive complexity in their own terms, and investigate its relationship to social judgment. Crockett (1965), for example, proposed that the degree of development of an individual's cognitive system with respect to other individuals is reflected in both its level of differentiation, operationally defined in terms of the number of constructs, and its degree of hierarchical integration, operationally defined in terms of both the pattern of relationships among constructs

within subsystems and the extent to which subsystems are interrelated by superordinate constructs. Unfortunately, these variations in definition (hence meaning) have led to a veritable quagmire of terms and phrases in the area. Cognitive complexity, however, has most generally been postulated to mean that persons are prone to employ many dimensions when they perceive and evaluate stimuli, or are inclined to make fine distinctions among dimensions of meaning. On the other hand, cognitively simple persons are thought to make only very gross discriminations and/or employ relatively few dimensions.

The Pervasiveness of Cognitive Complexity. Traditionally, researchers have viewed cognitive complexity from one of two perspectives. Some, such as Werner (1957), have viewed cognitive complexity simply as a pervasive trait encompassing one's entire cognitive field; others have seen it as a situationally determined construct that is specific to a particular domain of thought. The latter position, which is expounded on later in this paper, has been championed by Scott (1963) and Zajonc (1960).

The pervasiveness of cognitive complexity is reflected by a structural tendency whereby an individual consistently construes differing elements in a similar organizational manner. This proposition would be supported by studies finding that individuals who construe one set of elements (e.g., objects) complexly also construe other sets of elements (e.g., people) complexly. A number of researchers (Allard & Carlson, 1963; Bieri & Blacker, 1956); Carr, 1969; Epting, 1972; Hess, 1966; Rhodes, Carr & Jurji, 1968; and Tripodi & Bieri, 1963) have maintained that cognitive complexity is a general concept. Allard and Carlson, for example, provided support for the pervasiveness of cognitive complexity, with complexity being measured by considering both the pattern of construct applications and the descriptive terms written.

Using three stimuli for measuring complexity, (people known personally by the subject, famous people, and geometric designs), they reported correlations between the resultant complexity scores ranging from 0.57 to 0.67 (all $p < 0.001$). Due to the magnitude of the correlations and the amount of variation in the stimulus objects, those authors concluded that there was strong support for the hypothesis of generality, both within the interpersonal realms and between realms.

On the other hand, the strongest support for the specificity of cognitive complexity comes from Zajonc (1960). Zajonc's theory proposed that individuals process or "filter" information differently depending on the situational requirements and the individual's ability or expertise. Zajonc has demonstrated in several studies (1960; Zajonc & Wolfe, 1966) that the cognitive structures an individual uses when he or she expects to receive and then transmit information are different from those employed when one is expecting simply to receive information. Ultimately, one's momentary "cognitive style" is contingent upon the immediate situation.

Although the evidence for cognitive complexity as a pervasive trait vs. a situationally specific characteristic is mixed, both positions can accommodate the assumption that cognitive complexity is a construct that develops over time. That is, cognitive complexity, like performance appraisal, is conceptualized as developing and everchanging in the face of competing alternatives--always different and emerging. Thus, if rating behavior is viewed as a process, cognitive complexity, which is thought to moderate how people rate others, should also have developing properties over time. Most important to this study is an empirical method that examines the developing cognitive properties of raters. The methodological approach used is similar to that of Zajonc (1960). What follows is a brief description of his method, and how it relates to this study.

Zajonc's Method. Zajonc (1960) believed that the cognitive structures an individual uses when he or she expects to receive and then transmit information are different from those employed when expecting simply to receive information. Although Zajonc's theme is not of central importance here, his description and measurement of cognitive structure is. Zajonc's method is unique in that it describes the whole process of one's cognitive structure, from the elicitation of initial information from his subjects to the construction of a set of scores defining various features of cognitive structure.

The task in Zajonc's studies required subjects to review a job application letter from a candidate to a prospective employer. Each subject was issued a booklet, and attached to each booklet were stacks of 52 blank cards marked from A to ZZ. After reading the job application letter, subjects' initial task was to write on each card one characteristic which described their perception of the applicant. They were told to write whatever came to mind, and to note as many characteristics as they believed were necessary to describe the applicant adequately. The number of characteristics attributed to the applicant constituted the first measure, the degree of differentiation demonstrated by each subject.

A second measurement (unity) was based on the argument that some of the characteristics might depend on one another in a functional way; that is, if one changed, the others would change, too. All the cards were placed in alphabetical order, A to ZZ, and subjects were asked to list all the characteristics that would change if characteristic "A" were changed, absent, or untrue of the applicant. Similar lists were constructed for the remaining characteristics. Dependency matrices constructed from these responses permitted calculation of a measure of unity. The unity measure, which served as a measure of hierarchic integration, was based on the following argument:

If we define the dependence of the attribute $A(i)$ on the attribute $A(j)$ as equal to one when a change in $A(j)$ produces a change in $A(i)$, and as equal to zero when a change in $A(j)$ does not produce a change in $A(i)$, then a dependency matrix can be constructed for all attributes of a given cognitive structure, and the total dependency of each attribute may easily be obtained by summing the entries in the appropriate row. To compare the unity of structures of different degrees of differentiation, the measure of unity must be normalized. Given a structure with n attributes, the maximum sum of dependencies in the cognitive structure is $n(n-1)$. Unity then can be viewed as:

$$\frac{\sum \text{Dep}(A(i))}{n(n-1)}$$

where $\text{Dep}(A(i))$ is the total dependency of the i th attribute. (Zajonc, 1960, p. 160).

In short, the greater the number of dependent changes listed by the subject (as a proportion of the total number of adjectives or phrases), the greater the degree of unity.

There are a number of apparent advantages when Zajonc's approach of assessing conceptual structure is compared with that of the more popular grid formats advanced by Bieri (1955). First, Zajonc's approach directs a subject's attention to a specific situation, person, or problem as a starting point, which contrasts with grid procedures that deal with generalized relationships among constructs. Zajonc's procedure reveals ideas and concepts that are relevant to a specific situation. As a result, the method is sensitive to individual developmental changes that occur within a particular context.

Second, the method allows subjects the freedom to subgroup constructs into more discreet clusters. This permits closer examination of constructs that govern other constructs. Zajonc utilized both concepts of differentiation and hierarchic integration in measuring cognitive complexity. Although his term for it was "unity", spirit and context of the measure are synonymous.

with hierarchic integration. This thesis highlights the fact that one must consider both differentiation and integration in cognitive development, since together they characterize developmental changes. Most crucial to the thesis, then, is the ability of Zajonc's method to monitor the development of a rater's conceptual structure over time. In particular, from the vantage point of epistemology and developmental theory construction, Zajonc's method of measuring cognitive complexity provides the criteria for developmental change. That is, as individuals gain increased differentiation and move toward complexity of organization, it becomes possible to conceptualize rating behavior from Werner's (1957) perspective of "genetic" development.

Summary

This introduction has dealt with the reconceptualization of both performance appraisal and cognitive complexity from the framework of a developmental analysis, namely the organismic-holistic orientation of Werner and Kaplan (1963). Considered together, performance appraisal and cognitive complexity may lead to a novel understanding of the rating process. One way to approach this elusive process is to sample the products of thinking at various times in a rater's developmental progression. In this way, knowledge of the structural organization of the individual would allow one to identify the direction of his/her developmental process. Since the dynamics surrounding the development of persons' cognitive structures remain an empirical question, the relationship between such development and other cognitive behaviors (i.e., the rating process) is necessarily unknown.

It has been proposed that percepts do seem to develop toward a state of differentiation and hierarchic integration. If this is a normal state of progression, what are the implications for performance appraisal? This question is of paramount importance, for if it can be empirically demonstrated that

raters' judgments of others develop over time, and that these judgments are reflected in their ratings of those other individuals, then a new understanding of the rating process may be in the offing. Emphasis is therefore placed on theoretically and empirically linking performance ratings to the cognitive development of the individual.

So far, my inquiries have been progressively narrowing from the broad study of development and its many meanings, by way of certain preoccupations with Heinz Werner, to the specific research problems associated with performance appraisal. The following is an empirical inquiry into specific relationships that exist between cognitive development and the rating process. There are three questions of particular interest:

- 1) Do individuals' cognitive structures develop over time? .
- 2) How, and in what direction(s) does development proceed?
- 3) What are the relationships (if any) between measures of cognitive development and a psychometric property (halo) of actual performance ratings?

Method

Participants

Participants in this study were 45 students from three psychology courses representing different academic levels: introductory (n=14), upper class (n=20), and graduate (n=11). Students from the introductory course received required experimental credit for participating, while students in the upper class and graduate courses were volunteers.

Procedure

Cognitive Complexity. Measures of differentiation and hierarchic integration were obtained at three different times over the course of a 15-week semester. The method for collecting the data was similar to Zajonc's (1960;

see page 13). In Zajonc's experiments, applicants read a job application letter; in the present study, students were asked to provide a list of characteristics or qualities that best described their psychology instructor. The total number of characteristics constituted the differentiation score.

Participants were then asked to list those characteristics that would change if the first characteristic were changed, absent, or untrue for the instructor. Similar lists were created for the eventuality of change in each of the remaining characteristics. Dependency matrices were constructed from these responses, and the measure of hierarchic integration was derived from this matrix using the following formula:

$$\frac{\sum \text{Dep } A(i)}{n(n-1)},$$

where $A(i)$ is the number of characteristics dependent on the i th characteristic, and n is equal to the number of characteristics. Instructions for obtaining these measures appear in Appendix I.

Performance Ratings. Students provided actual performance ratings of their psychology instructor. These ratings were gathered twice, once at the beginning and once at the end of the semester. The ratings and measures of cognitive development were obtained during the same sessions, and the rating scales were administered to the subjects in one of two conditions, either prior to acquiring measures of differentiation and hierarchic integration, or after these measures were taken. This procedure controls for possible order effects in the data.

The particular scale used for this investigation was a behaviorally anchored rating scale (BARS) which was developed for a previous research project (Kirkeide, Note 1) involving teacher evaluations. This scale was developed by undergraduate psychology students using the method suggested by Smith and Kendall

(1963). Each of nine dimensions considered to be important to teaching effectiveness were separately rated on a seven-point scale which contained specific behavioral statements to be used as referents to guide the rater. The rating scale and instructions appear in Appendix II.

A great deal of time and interest has been devoted to the development of techniques and procedures for assessing the veracity and general psychometric qualities of rating scales. Unfortunately, a number of factors (i.e., conceptual confusion, lack of agreement regarding operational definitions, and an overall absence of systematic inquiry into the statistical indices of rating errors) have limited progress in the research into the psychometric quality of ratings (Saal, Downey, & Lahey, 1980). In addition, Saal et al, have noted the limitations of various research designs for providing information leading to the assessment of rating quality. In particular, the design proposed here, a rater x dimension design, "is simply incapable of yielding any rater x ratee interaction, or any ratee main effect" p. (420). As a result, it is not possible to pinpoint the source of a rating error. Without a complete Rater x Ratee x Dimensions matrix, (in which all the raters evaluate all the ratees on all the behavioral dimensions) each alternative strategy "yields data that are inherently limited regarding possible analyses" (p. 420). In other words, measures of leniency and range restriction that can be derived from this design are confounded with respect to source of "error".

For example, leniency, a shift in mean ratings from the midpoint of the rating scale in the favorable or unfavorable direction (Bernardin, Lashells, Smith & Alveres, 1976), may reflect "error" on the part of rater or actual behavior on the part of the ratee. The extent to which obtained ratings discriminate among different ratees in terms of their respective performance levels (Motowidlo & Borman, 1977) cannot be evaluated since this measure is

contingent upon having multiple ratees. Therefore, measures of leniency and ratee dispersion were not obtained.

One common rating index, however, halo, which is consistently understood as a rater's failure to discriminate among conceptually distinct and potentially independent aspects of a ratee's behavior, could be evaluated in the context of this developmental research. Halo was assessed for each rater individually by calculating the standard deviation of each rater's ratings across the nine dimensions on the scale (Bernardin & Walter, 1977). This is a common statistical technique for measuring halo, given this particular research design (Saal, Downey, & Lahey, 1980). Relationships were examined by correlating halo, as measured standard deviations, with the cognitive measures of differentiation and hierarchic integration.

Results

The results of this study were analyzed to assess whether cognitive complexity developed over the semester, and to determine if there were any corresponding changes in selected characteristics of the students' ratings over the same 15-week period. Analyses were performed separately for each of the three classes, and a second group of analyses combined the three classes. Differentiation, hierarchic integration, and halo scores were used as data points in separate analyses of variance (ANOVAS), with time (session) as a repeated measure. In addition, intercorrelations among differentiation, hierarchic integration, and halo were computed in order to show relationships among the measures. The raw data for each subject appear in Appendix III.

Table 1 contains the ANOVA results reported separately by class. Results for the introductory class reveal no significant change for any of the measures. The advanced class showed significant decreases in differentiation,

Table 1
Class Analysis as a Function of Time

<u>Introductory Class (n=14)</u>					
	Session 1	Session 2	Session 3	F	p
Differentiation	11.86	10.29	11.86	1.18	n.s.
Integration	.2795	.3078	.2809	<1	n.s.
Halo	.7897	---	.7984	<1	n.s.
<u>Advanced Class (n=20)</u>					
	Session 1	Session 2	Session 3	F	p
Differentiation	11.30	9.20	9.20	4.26	<.05
Integration	.2029	.2743	.2400	3.09	<.06
Halo	.9097	---	.9029	<1	n.s.
<u>Graduate Class (n=11)</u>					
	Session 1	Session 2	Session 3	F	p
Differentiation	12.55	10.09	9.18	1.49	n.s.
Integration	.1592	.3024	.3092	8.82	<.01
Halo	1.0736	---	.9302	<1	n.s.

and a marginal increase in hierarchic integration over time. In the graduate class, only hierarchic integration changed (increased) significantly over the three sessions. Halo remained unchanged across sessions for all three classes.

The classes showed a similar pattern of change in differentiation and hierarchic integration scores from session 1 to session 2. In all cases differentiation decreased and hierarchic integration increased. Therefore, in order to increase the power of the statistical test, classes were combined and the analyses were repeated. Using this strategy, changes in differentiation and hierarchic integration from session 1 to session 2 reached significance. Halo again remained unchanged. Mean scores on all three measures as a function of session appear in Table 2.

Post hoc (Newman-Keuls) tests of significance were performed in order to detect specific differences between sessions. The Neuman-Keuls tests demonstrated a decrease in differentiation scores between the first and second session and the first and third session. Differences between the second and third sessions were not significant. Hierarchic integration significantly increased from time 1 to time 2 and from time 1 to time 3. Again, changes in scores between the second and third sessions did not reach significance.

For the purpose of comparing levels of differentiation and hierarchic integration across different studies, ranges of differentiation and integration scores were obtained from previous research. Comparing the present results with Zajonc's (1960) revealed similar ranges in both differentiation and hierarchic integration scores. The range of mean (collapsed across sessions) differentiation scores for the present sample was 9.76 to 11.78 and the mean hierarchic integration scores were .2161 to .2697. The range in differentiation scores for Zajonc (1960) was 5.12 to 7.90 and integration scores ranged from .222 to .309.

Table 2
Mean Scores as a Function of Time

<u>Differentiation^a</u>				
Session 1	Session 2	Session 3	<u>F</u>	<u>p</u>
11.78	9.76	10.02	4.78	<.05

<u>Hierarchic Integration^a</u>				
Session 1	Session 2	Session 3	<u>F</u>	<u>p</u>
.2161	.2916	.2697	7.55	<.01

<u>Halo</u>				
Session 1		Session 3	<u>F</u>	<u>p</u>
.9113		.8770	<1	N.S.

^aSession 1 scores are different from both Sessions 2 and 3.
Sessions 2 and 3 are not different from one another.

The levels of differentiation in the present study appear somewhat higher than in Zajonc's work. This may be explained by procedural differences in the measurement of differentiation. Zajonc had students describe a job applicant based on information contained in an application letter rather than asking them to describe a known person (someone they are involved with). As a result, the number of characteristics (differentiation) generated for each participant in Zajonc's study may have been limited. Levels of hierarchic integration do not seem to be affected by these differences. Given these comparisons, the modifications in Zajonc's original testing method that were employed in this study may have had an effect on the differentiation scores. However, since this study views differentiation and integration as specific, contextual constructs, these differences don't represent serious methodological problems.

Intercorrelations among the variables (Table 3) show that the measures of differentiation were significantly correlated over the three experimental sessions; so were the measures of hierarchic integration. These results suggest that, although both differentiation and hierarchic integration changed over the course of the study, the students' scores changed in a relatively consistent manner.

When hierarchic integration was correlated with differentiation within experimental sessions, significant negative correlations were found, except in the third stage of the study. That is, students who exhibited high degrees of differentiation in the first two sessions also showed low degrees of integration in those sessions. The interrelationships between differentiation and hierarchic integration have not been clarified in the literature. A possible theoretical link between these two concepts is discussed below.

Halo at time 1 was significantly correlated with halo at time 2; however, these scores were not related to either differentiation or hierarchic integration in any of the three sessions. Moreover, since halo did not change

Table 3

Intercorrelations Among Differentiation, Hierarchic Integration and Halo

	<u>Differentiation</u>			<u>Hierarchic Integration</u>			<u>Halo</u>	
	Time 1	Time 2	Time 3	Time 1	Time 2	Time 3	Time 1	Time 3
Time 1	-----	.23*	.35**	-.33**	-----	-----	n.s.	n.s.
Time 2		-----	.63**	-----	-.32**	-----	n.s.	n.s.
Time 3			-----	-----	-----	n.s.	n.s.	n.s.
Time 1				-----	.55**	.33**	n.s.	n.s.
Time 2					-----	.58**	n.s.	n.s.
Time 3						-----	n.s.	n.s.
Time 1							-----	.37**

* $p = .06$ ** $p < .01$

significantly from session 1 to session 3, a systematic relationship between changes in differentiation and hierarchic integration and changes in rater's halo was precluded. In other words, halo was not related to cognitive complexity, nor to the development aspects of cognitions.

Pearson product moment correlations of differentiation, hierarchic integration and halo with class membership were used to check for possible ratee effects (i.e., individual differences in instructors). In addition, these measures were correlated with academic level and age of the students. These analyses were prompted by Zajonc and Wolfe's (1968) finding that both differentiation and complexity (integration) were positively related to an individual's position in the organization--the higher the position, the higher the differentiation and integration. To facilitate these analyses, differentiation, integration and halo scores were summed across sessions in order to obtain a single index for each subject on each measure. Correlations appear in Table 4.

Zajonc's results were not supported; nor was there any evidence for ratee effects. The only significant correlation was obtained between differentiation and age. This result could reflect the impact of the experience of raters on rating behavior, however, neither class nor academic level, which are direct measures of experience in academic settings, correlated with differentiation. Consequently, the significance of the relationship between age and differentiation remains equivocal.

Finally, correlations were performed to test for reliability in the ratings over the semester. Specifically, raters' ratings on nine dimensions from time 1 were correlated with each of their ratings at time 2. Eight of the nine correlations were significant ($p < .01$) and ranged from .38 to .62. The

Table 4
Correlations as a Function of Class, Academic Level
and Age

	Class	Academic Level	Age
Differentiation	-.10	-.14	-.31*
Integration	-.13	-.21	.02
Halo	.24	.23	-.01

*p < .05

non-significant correlation obtained on dimension nine which dealt with exams were not given by any of the instructors during the first three weeks of classes (prior to the first experimental session). These results add support for the reliability of the present rating scale and strengthen the argument that ratings did not systematically differ over experimental sessions. The ratings did not reflect changes in raters' cognitive development.

Discussion

Development of Cognitive Complexity

The results of the present study support the idea that raters' cognitive complexity changes over time. More importantly, it appears the cognitive complexity not only changes, but also exhibits developing properties. Evidence for development comes from the fact that in the combined analyses (introductory, advanced, and graduate students analyzed together), raters' differentiation and hierarchic integration scores changed significantly over the course of the study. Since participants were asked to produce adjectives describing their instructor during the first week of classes, supposedly when familiarity with the instructor was minimal, it was surprising to find raters generating more adjectives (hence higher differentiation) in the first session than in any other session. Hierarchic integration, as expected, increased at session 2, but showed no change between session 2 and session 3.

These findings can be interpreted from two different theoretical positions. This paper has been grounded in the organismic-holistic framework of Werner and Kaplan (1963). According to their analysis, psychological development involves a progression from relatively undifferentiated functioning with little specialization toward greater differentiation and the emergence of complex subsystems separating various aspects of behavior (hierarchic integration). The present

results seem only to support one aspect of this framework; although hierarchic integration was found to increase over time, differentiation actually decreased.

The apparent discrepancy between these results and Werner's hypothesis may be reconciled by referring to the "pervasiveness" vs. specificity issue noted earlier in the text. For Werner, differentiation and hierarchic integration represent organismic phenomena, global in nature, encompassing all areas of psychological functioning. Subsequent researchers have expanded this concept of differentiation (Witkin, 1962, 1965; Watchel, 1972, Zajonc, 1960). In particular, Zajonc (1960) conceptualized differentiation as it relates to a single cognitive structure where cognitive structure is defined as an "organized subset of the given cognitive universe in terms of which the individual identifies and discriminates a particular object or event" (p. 159). Thus, "the number (N) of attributes constituting a given cognitive structure is taken to reflect its degree of differentiation" (p. 160). The important difference between these researchers' views is simply that in Werner's framework differentiation represents the entire universe of functioning while Zajonc defines it in terms a particular domain in that universe.

This difference in theoretical conceptualization of differentiation has led to different hypothesis regarding developmental changes in the construct. Zajonc (1960), for instance, assumed that other factors may operate simultaneously to reduce differentiation within a given cognitive structure, "namely, the tendency to select attributes that are consistent with one's position" (p. 165). In other words, for Zajonc, the differentiation of the cognitive universe may increase over time but the differentiation within a structure does not necessarily correspond. With respect to the present study, support can be found for this position. For example, students may need to gather as much initial information as they can about an instructor in order to perform

well in class (preparing for exams, discussion, etc.). Once information is combined (hierarchically integrated) into expectations, however, there may be less need for additional descriptors of behavior, and some of these may lose significance and eventually "get dropped" from the repertoire of adjectives. It is argued, then, that rater development proceeds in the direction of increased meaning, and that this development may be accompanied by either increases or decreases in differentiation.

Referring again to Zajonc's (1960) analysis, it is expected that the degree of differentiation should be greater for individuals that are forced to deal with incongruent information. He shows that the receivers of contradictory information demonstrate a more pronounced level of differentiation than do receivers of congruent information. With respect to this study, it could be argued that after the third week of classes (after session 1) students' exposure to conflicting or incongruent information was minimal permitting the "other factors" to reduce differentiation. Without accompanying increases in hierarchic integration, this argument would become moot. Hierarchic integration, however, did increase, thereby maintaining the criteria for development.

In order for development to proceed, increases in the hierarchical integration of information is a necessary condition. This is true whether one considers the entire universe of functioning (a la Werner) or simply subsets of that universe (Zajonc's interpretation). Both theoretical interpretations of cognitive structure require an increase in the integrative aspects of functioning under any conditions of development. The crucial point is that both differentiation and hierarchic integration characterize development (Kelly, 1955; Lewin, 1951; Piaget, 1932, Werner, 1957). Focusing on one or the other of these different processes results in missing the nature of the synthesis

that characterizes psychological development, and accordingly provides an incomplete view of the organism (Lerner, 1966).

Differentiation and Hierarchic Integration

Harvey, Hunt and Schroder (1961) conceptualized differentiation and hierarchic integration as processes "by means of which the organism breaks down his environing world into parts and relates them in ways relevant to his motive system and existence" (p. 22). For Piaget (1960), concepts develop through the interdependent operation of accommodation and assimilation, which represent another perspective on the differentiation/hierarchic integration interaction. Although these examples suggest that differentiation and hierarchic integration are theoretically related to one another, their exact statistical relationship has yet to be established (Schroder, Driver & Struefert, 1967; Bieri, 1961).

Lewin's (1951) field theory proposed an expected relationship between differentiation and hierarchic integration. In order to do this, he distinguished between two types of integration. "Simple dependence" refers to the interrelationships between parts of single cognitive structures, whereas hierarchic organization refers to the complicated interdependence of the parts of the whole system. He proposed that the simple dependence of a structure would decrease as its differentiation increased (Bieri, 1961). If this conceptualization is true, it could help to account for the negative relationship between differentiation and hierarchic integration found in this study. Conceptualizations of the relationship between these two major characteristics of the entire cognitive universe remain problematic.

One reason that researchers have been unable to theoretically and empirically link differentiation and integration has been suggested by Harvey, Hunt

and Schroder (1961). They explained that there is often the implicit assumption that hierarchic integration must be preceded by differentiation. Although this assumption is logically appealing, it is plausible that differentiation and hierarchic integration could occur simultaneously, and thereby render attempts to predict their interrelationships futile. In any event, the interrelationships between differentiation and hierarchic integration are less than clear, and a critical appraisal of this issue is lacking in the literature on cognitive structure.

The notion that cognitive structure can develop and change over time is central to this study. Development, however, was not demonstrated between session 2 and 3. Since there are no data or theories to explain these results, a number of hypothesis are proposed. In almost any psychological experiment, fatigue or boredom can be a factor. The problem may be particularly acute in the method used in this study because students were asked to generate adjectives descriptive of their instructors a total of three separate times. Although the mean scores in differentiation decreased from session 1 to session 2 the mean scores did not continue to decrease from session 2 and session 3. In addition, performance on the more demanding task increased as reflected in the scores for hierarchic integration, thereby diminishing the possibility of a fatigue effect.

If fatigue can be ruled out as an explanation for the lack of change from session 2 to session 3, others must be sought. It may be that the time span from session 2 to session 3 was too short for development to proceed. Although the time span between session 1 and session 2 was the same as that between session 2 & 3, it could be argued that development is not a linear

process, and that further development requires additional time, events, exposure, etc. In other words, there is no reason to believe that development occurs at equal time intervals, let alone along the specific intervals chosen for this study. Finally, the instructor's performance in class may not have produced enough variability (contradictory information) in order for perceptions to develop beyond the ninth week. It could be argued that development of perceptions of instructors occurs only during the first few weeks of exposure, and that later time is spent confirming these early conceptions. Whatever explanation is preferred, future research along the contours of a developmental analysis should be directed toward identification of the parameters (duration, magnitude, etc.) of development.

Although other researchers have suggested that a person's cognitive complexity can develop, this idea has been addressed primarily in the context of child psychology, emphasizing the development of language skills and social experience (Adams-Weber, 1979). Very little attention has been given to the notion of cognitive development in adults. These data suggest, however, that adults' cognitive complexity is not static, and may reflect or influence a person's developing perception of others. As stated earlier, these perceptions are of critical importance to industrial psychology, since they initiate and guide the rating process.

Cognitive Development and the Psychometric Properties of Ratings

Although cognitive complexity seems to develop even over a short time, these results show that there is no corresponding change in the halo associated with raters' ratings. This finding supports the results reported by Lahey & Saal (in press) who suggested that cognitive complexity does not explain the psychometric properties of ratings based on rating scales containing various levels of discriminations between and within dimensions.

The tempting response is to suggest a moratorium on cognitive complexity research, or to simply add one more study to the growing list of those that question the validity of Schneier's (1977) findings (e.g., Bernardin & Boetcher, Note 1; Bernardin & Cardy, Note 2; Vance & Kuhnert, Note 3). These reactions, however, may be somewhat premature.

It is important to investigate the reason(s) that the development of a rater's perception of a ratee was not reflected in this property of the information elicited by performance appraisal instruments. Why don't the process aspects of ratings seem to correspond to this end product of those ratings? Since the rater's task is to extract from a large body of information those aspects of greatest relevance (according to the rating scale), assimilate them, and to arrive at some evaluative judgment in the form of a numerical index, one potential explanation lies in the items or dimensions tapped by the rating scale. These aspects of ratee behavior may not reflect those aspects most relevant to the individual rater. Although contemporary methods of rating scale construction such as Smith & Kendall's (1963) behaviorally anchored rating scale procedure attempt to maximize correspondence between scales and raters, these methods rely on the aggregation of information over a number of raters, and may not reflect any one rater's "real" perceptions. The cognitive developmental perspective, which allows the rater to characterize a ratee in the manner that he/she chooses, and further allows for changes in characterizations, may afford better assessment of the raters' perceptions than traditional rating scales. It may be the case that standard rating scales, and traditional indices of rating quality (i.e., halo), may not only be insensitive to an individual rater's development, but may also be inappropriate means for evaluating the process aspects of rating behavior.

This is not to diminish the work of countless researchers who have analyzed the content of ratings and their usefulness for appraising performance, but simply to provide alternative means for investigating raters' ratings. In order to incorporate process considerations into rating research, the criteria must be expanded beyond current accepted practice. This does not necessarily imply that the methods needed to examine the developmental processes of raters would be practical to implement in the context of a functional organizational setting. It is acknowledged that such an individualized approach to ratings could produce administrative nightmares. For the purposes of psychological research and expanding our knowledge of the rating process, however, the developmental perspective offers alternatives to standard practice.

These issues constitute the "cutting edge" of this paper. That is, the essence of the cognitive complexity-rating scale issue lies directly in whether one accepts the end-products of ratings as criteria for the analysis of rating behavior, or whether process considerations (e.g., the developmental perspective proposed here) may represent a more promising analytical perspective, at least for theoretical purposes.

Conclusion

Industrial psychology during the past 50 years has not directed much attention toward developmental consideration. Although research on leadership, motivation, job satisfaction, etc. have attended to process aspects, (e.g., opponent-process model of satisfaction) the ultimate goal has been to demonstrate predictive efficacy. The assumption has been that if emerging theories manifest predictive efficacy proper developmental theory must prove corroborative (Gergen, 1977). Attempts to understand the qualities of a good leader, or any other industrial psychological concern, cannot await a complete theory of human development. This self-fulfilling separatism is

perhaps characteristic of any paradigmatic discipline (Kuhn, 1962). As a result, concurrent advances in different fields may often lead to disparate conclusions, even when based on similar data. Because researchers initially tend to disregard others with whom they disagree, this separatism may lead to severe antithetical developments among research disciplines. As long as researchers, even disparate in discipline, confine themselves to one theoretical base, there are limited alternatives in the selection of the strongest data set which may confirm or disconfirm a theoretical perspective (Mitroff, 1976).

In order to compare theoretical orientations, and gain support for one of them through empirical observation, there must be competing alternatives. In addition, the testing of a single theory requires competing hypotheses within the theory. Theory construction, therefore, at least recognizes rudimentary considerations of others' views as well as one's own. For this reason, research efforts should be committed to theoretical and methodological pluralism (Feyerabend, 1968).

With these considerations in mind, this thesis was designed to emphasize developmental consideration in the field of industrial psychology. This study was based on the axiom that the history of the rater should not be ignored--and demonstrated that experience does make a difference. Determining just how cognitive development influences ratings is a matter for future discourse. On a broader level, the impact of raters' cognitive development on the interactions with ratees and rating scale formats needs to be articulated and studied. It would seem that a developmental reconceptualization is necessary if a "true" process model of rating behavior is to be advanced. From the standpoint of theory building and metatheory, the developmental framework may provide a way of understanding rating behavior in all its glory.

Footnotes

1. A similar dynamic characterizes Alerfer's ERG model of motivation (Aldefer, 1972).
2. There is some research along these lines that has investigated the effects of whether ratings are to be used for research purposes or for administrative decisions (Warmke & Billings, 1979).

Reference Notes

1. Bernardin, H. J., & Boetcher, R. The effects of rater training and cognitive complexity on psychometric error in ratings. Paper presented at the annual meeting of the American Psychological Association, Toronto, Ontario, Canada, 1978.
2. Bernardin, H. J., & Cardy, R. L. Cognitive complexity and appraisal effectiveness: It makes no never mind. Unpublished manuscript, 1980.
3. Vance R. J., & Kuhnert, K. W. Cognitive complexity and rating behavior. Paper presented at the annual meeting of the Midwestern Psychological Association, St. Louis, MO., 1980.

REFERENCES

- Adams-Webber, J. R. Personal Construct Theory: Concepts and Applications. New York: John Wiley & Sons, 1979.
- Alderfer, C. P. Existence, relatedness, and growth. New York: Free Press, 1972.
- Allard, R. & Carlson, E. R. The generality of cognitive complexity. Journal of Social Psychology, 1963, 59, 73-75.
- Bernardin, H. J., Lashells, M. B., Smith, P. C., & Alveres, K. M. Behavioral expectation scales: Effects of developmental procedures and formats. Journal of Applied Psychology, 1976, 61, 75-79.
- Bieri, J. Cognitive complexity-simplicity and predictive behavior. Journal of Abnormal and Social Psychology, 1955, 51, 263-268.
- Bieri, J. Complexity-simplicity as a personality variable in cognitive & preferential behavior. In D. W. Fiske & S. R. Maddi (Eds.) Function of Varied Experience, Homewood, IL: Dorsey Press, 1961.
- Bieri, J. Cognitive complexity and personality development. In O. J. Harvey (Ed.), Experience, structure, and adaptability. New York: Springer, 1966.
- Bieri, J. and Blacker, E. The generality of cognitive complexity in perception of people and inkblots. Journal of Abnormal and Social Psychology, 1956, 53, 112-117.
- Bruner, J. S., and Tagiuri, R. Person perception. In G. Lindzey (Ed.), Handbook of social psychology, vol. 2. Reading, MA: Addison-Wesley, 1954.
- Carr, J. E. Cognitive complexity: Construct descriptive terms versus cognitive processes. Psychological Reports, 1965, 16, 133-134.

- Crockett, W. H. Cognitive complexity and impression formation. In
Brendon A. Maker (Ed.), Progress in experimental personality research,
vol. 2. New York: Academic Press, 1965.
- Cronbach, L. J. Processes affecting scores on understanding of others
and assumed similarity. Psychological Bulletin, 1955, 52, 177-193.
- Epting, F. R. The stability of cognitive complexity in constructing
social issues. British Journal of Social and Clinical Psychology,
1972, 11, 122-125.
- Farr, J. L. Response requirements and primary-necessary effects in a
simulated selection interview. Journal of Applied Psychology, 1973,
57, 228-232.
- Feyerabend, P. K. How to be a good empiricist: A plea for tolerance
in matters epistemology. In P. H. Nidditch (Ed.), The philosophy
of Science. Oxford University Press, 1-12.
- Gergen, K. J. Stability, change and eliance in understanding human development.
In N. Datan & H. Reese (Eds.), Life-span developmental psychology: Dialecti-
cal perspectives on experimental research. New York: Academic Press, 1977 .
- Hakel, M. D., Dobmeyer, T. W., & Dunnette, M.D. Relative importance of three
content dimensions in overall suitability ratings of job applicants' resumes.
Journal of Applied Psychology, 1970, 54, 65-71.
- Harris, D. B., The concept of development. Minneapolis, MN: University
of Minnesota Press, 1957.
- Harvey, O. J., Hunt, D. E. and Schroeder, H. M. Conceptual systems and
personality organization. New York: Wiley, 1961.
- Hess, H. F. Additional evidence for the generality of cognitive com-
plexity. Perception and Motor Skills, 1966, 23, 857-858.
- Hollmann, T. D. Employment interviewers' errors in processing positive and
negative information. Journal of Applied Psychology, 1972, 56, 130-134.

- James, W. The principles of psychology. New York: Holt, 1890.
- Kelly, G. A. The Psychology of personal constructs. New York: Norton, 1955.
- Kuhn, T. S. The structure of scientific revolutions. Chicago: University of Chicago Press, 1962.
- Lahey, M. A., & Saal, F. E. Evidence incompatible with a cognitive compatibility theory of rating behavior. Journal of Applied Psychology, in press.
- Landy, F. J., & Bates, F. Another look at contrast effects in the employment interview. Journal of Applied Psychology, 1973, 58, 141-144.
- Landy, F. J. & Farr, J. L. Performance rating. Psychological Bulletin, 1980, 87(1), 72-107.
- Lomov, B. F. Psychological processes and communication. Soviet Psychology, Fall 1978, SVIT(1), 3-22.
- Lerner, R. M. Concepts and theories of human development. Phillippines: Addison-Wesley Publishing Co., 1976.
- Lewin, K. Field theory in social science. New York: Harper, 1951.
- Marx, Karl. Economic and philosophical manuscripts of 1844. London: Lawrence & Wishart Ltd., 1970.
- Maas, J. B. Patterned scaled expectation interview: Reliability studies on a new technique. Journal of Applied Psychology, 1965, 59, 431-433.
- Mayfield, E. C. The selection interview: A re-evaluation of published research. Personnel Psychology, 1964, 17, 239-260.
- Mayo, Clara W. and Crockett, W. H. Cognitive complexity and primacy-recency effects in impression formation. Journal of Abnormal and Social Psychology, 1964, 68, 335-338.

- Mitroff, I. The subjective side of science: A philosophical inquiry into the psychology of the Apollo Moon Scientists. Amsterdam: Elsevier Scientific Pub. Co., 1974.
- Motowidlo, S. J., & Borman, W. C. Behaviorally anchored scale for measuring morale in military units. Journal of Applied Psychology, 1977, 62, 177-183.
- Piaget, J. The psychology of intelligence. New York: Harcourt-Brace, 1960.
- Rhodes, R. J., Carr, J. E. and Jurji, E. D. Interpersonal differentiation and perceptual field differentiation. Perceptual and Motor Skills, 1968, 27, 172-174.
- Rommetviet, R. Selectivity, intuition and halo effects in social perception. Oslo: Oslo University Press, 1960.
- Saal, R. E., Downey, R. G., & Lahey, M. A. Rating the ratings: Assessing the psychometric quality of rating data. Psychological Bulletin, 1980, 88(2), 413-428.
- Sarbin, T. R. Role theory. In G. Lindzey (Ed.), Handbook of Social Psychology, vol. 1. Cambridge, MA: Addison-Wesley, 1954.
- Schneider, C. E. Operational utility and psychometric characteristics of behavioral expectation scales: A cognitive reinterpretation. Journal of Applied Psychology, 1977, 62, 541-548.
- Schroeder, H. M., Priver, M. J. and Streufert, S. Human information processing. New York: Holt, Rinehart and Winston, 1967.
- Scott, W. A. Cognitive complexity and cognitive balance. Sociometry, 1963, 26, 66-74.
- Seymore, R. W., Jones, E. E. Changes in social perception as a function of the personal relevance of behavior. Sociometry, 1957, 20, 75-85.

- Shakespeare, William Hamlet: Second quanto, 1604-1605. Oxford, Charendon Press, 1964.
- Smith, P. C., & Kendall, L. M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. Journal of Applied Psychology, 1963, 47, 149-155.
- Tripodi, T., & Bieri, J. Cognitive complexity as a function of own and provided constructs. Psychological Reports, 1963, 26, 13.
- Ulrich, L., & Trumbo, D. The selection interview since 1949. Psychological Bulletin, 1965, 63, 100-116.
- Vance, R. J., Kuhnert, K. W., & Farr, J. L. Interview judgments: Using external criteria to compare behavioral and graphic scale ratings. Organizational Behavior and Human Performance, 1978, 22, 279-294.
- Vance, R. J., & Kuhnert, K. W. Cognitive complexity and rating behavior. Presented at the Annual Meeting of the Midwestern Psychological Association, St. Louis, Missouri, May 1980.
- Wachtel, Paul L. Cognitive style and style of adaption. Perceptual and Motor Skills, 1972b, 35, 779-785.
- Wagner, R. The employment interview: A critical review. Personnel Psychology, 1949, 2, 17-49.
- Warmke, D. L., & Billings, R. S. Comparison of training methods for improving the psychometric quality of experimental and administrative performance ratings. Journal of Applied Psychology, 1979, 64, 124-131.
- Werner, H. Comparative psychology of mental development. New York: International Universities Press, 1957.
- Werner, H., & Kaplan, G. Symbol formation: An organismic-developmental approach to language and the expression of thought. New York: Wiley, 1963, 3-11.

- Wexley, K. N., Sanders, R. E., & Yukl, G. A. Training interviewers to eliminate contrast effects in employment interviews. Journal of Applied Psychology, 1973, 57, 233-236.
- Wherry, R. J. The control of bias in rating. Personnel Research Board Rep. 922. Washington, D.C.: Department of the Army, Personnel Research Section, February, 1952.
- Witkin, H. A. Psychological differentiation and forms of pathology. Journal of Abnormal Psychology, 1965, 70, 317-336.
- Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. Psychological differentiation. New York: Wiley, 1962.
- Wolin, L. R. An analysis of the content of interpersonal perception. Unpublished doctoral dissertation, Cornell University, 1955.
- Zajonc, R. B., & Wolfe, D. M. Cognitive consequences of a person's position in a formal organization. Human Relations, 1966, 19, 129-150.
- Zajonc, R. The process of cognitive tuning in communications. Journal of Abnormal and Social Psychology, 1960, 61, 159-167.

NAME: _____
AGE: _____
YEAR IN SCHOOL: _____

This exercise is designed to gather information about some of the ways in which you view other people and how you use certain adjectives to describe others. This exercise does not take very long to complete and you are urged to fill it out carefully and as completely as possible.

Instructions for Differentiation and Dependency Measures

Part One

On each of the pages in the BLUE booklet in front of you, write one characteristic word (or phrase) which describes your professor. Put down whatever comes to mind as there are no "correct" or "incorrect" answers. You do not have to fill in all of the pages provided but enough to adequately describe your instructor, in your opinion. If you need more space, additional pages are available.

Part Two

It is possible that some of the characteristics that you have listed are related to one another in such a way that if one changes, others would also change. For example, suppose you had listed both honest and trustworthy as characteristics. If the characteristic of being honest were changed (i.e., if the person becomes dishonest), you may no longer believe the person to be trustworthy. So being honest and trustworthy are dependent on one another.

In the GREEN booklet in front of you, specify the characteristics you would expect to change if each of the characteristics in the blue booklet were changed. On the first page of the GREEN booklet, list the characteristics that would change if the first characteristic in the blue booklet were changed. On the second page of the GREEN booklet, list the characteristics that would change if the second characteristic in the blue booklet were changed. On the third page of the GREEN booklet,

Continue this process until you have completed a page in the GREEN booklet for each of the characteristics in the blue booklet.

Please identify each of the blue and green booklets with the name listed on the front page of the instruction sheet. Although all information gathered here will be kept strictly confidential, it is necessary to identify each persons' booklets. No names will be associated with the data collected here. We only need your names for identification purposes.

Thank you very much for your cooperation.

Appendix II

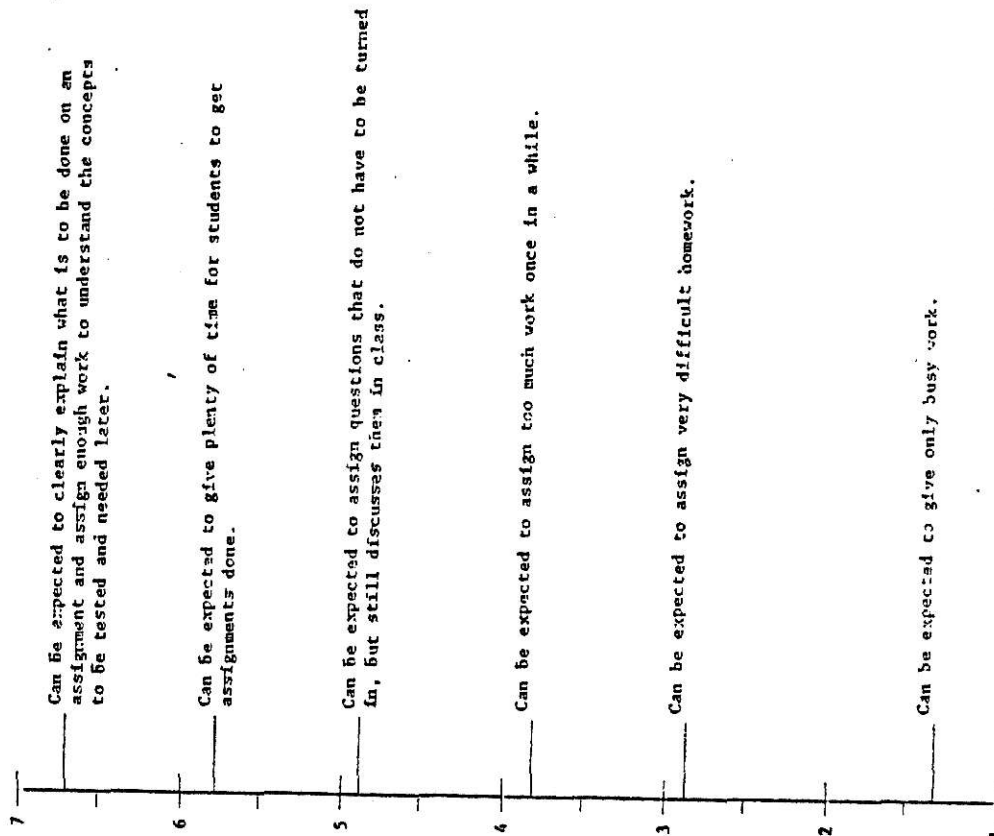
The following pages include nine categories concerning various aspects of teaching with definitions provided for each one. The categories and definitions resulted from meetings held with groups of college students. Below each category is a seven-point rating scale with behavioral statements located at various point along the scale. All of the statements were written by students and each statements location on the scale was determined by student's evaluation of the level of performance it best represented.

Please read the definition for each category, carefully, and the behavioral statements shown on the right of the accompanying scales. Then compare those behaviors of your instructor which apply to each particular category against the behaviors on the scales. Finally, use these behaviors as references or aids in helping you to determine the rating that you feel your instructor deserves. That is, based upon the behaviors that you have seen, rate your instructor according to the type of behavior that you would expect of him/her when compared to the behavior on the scale. It is possible that your instructor has never exhibited any of the behaviors shown on the scale. However, based upon those behaviors that you have seen exhibited by your instructor and which are applicable to the category your rating, rate him/her on the level of performance you would expect from him/her relative to the statements on the scales.

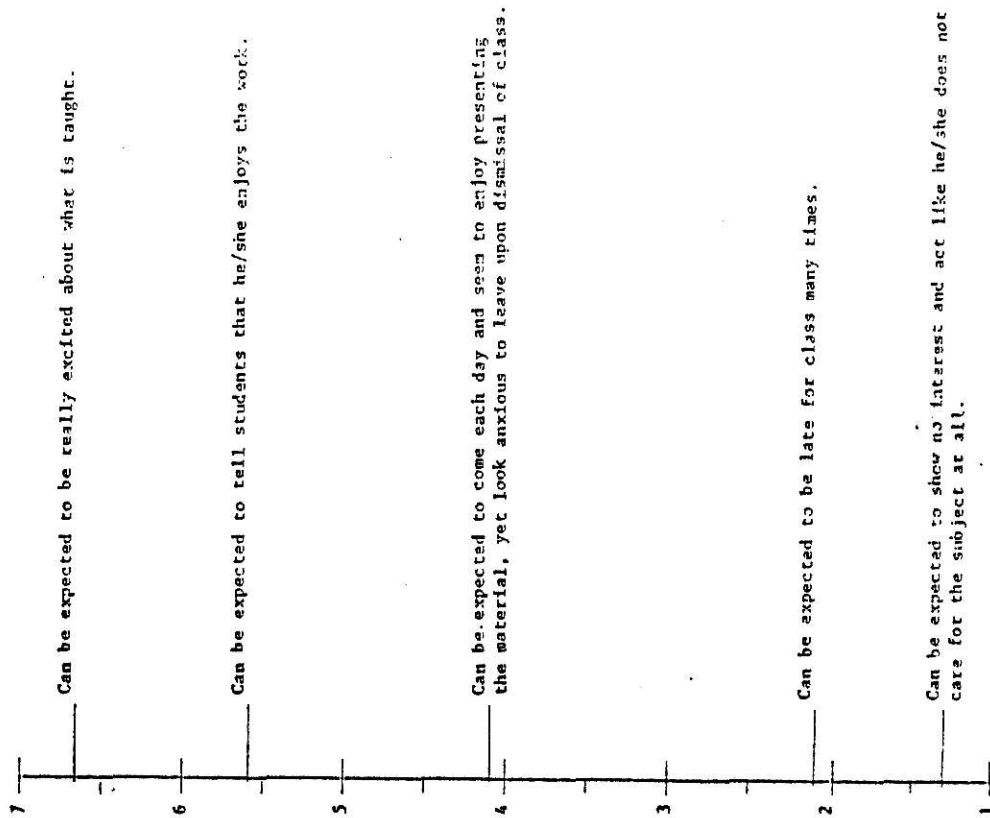
Please make a single mark anywhere on the vertical line line. It is not necessary to make a mark only where a statement is located. The statements are merely references against which you are to compare the actual observed behavior of the instructor.

At the end of the set of scales there is one additional five-point scale which you should fill out. After rating your instructor on all the categories decide how sure you are of these ratings and circle the one number which best describes how sure you are of all the ratings. In other words how certain are you of the total set of ratings that you have given.

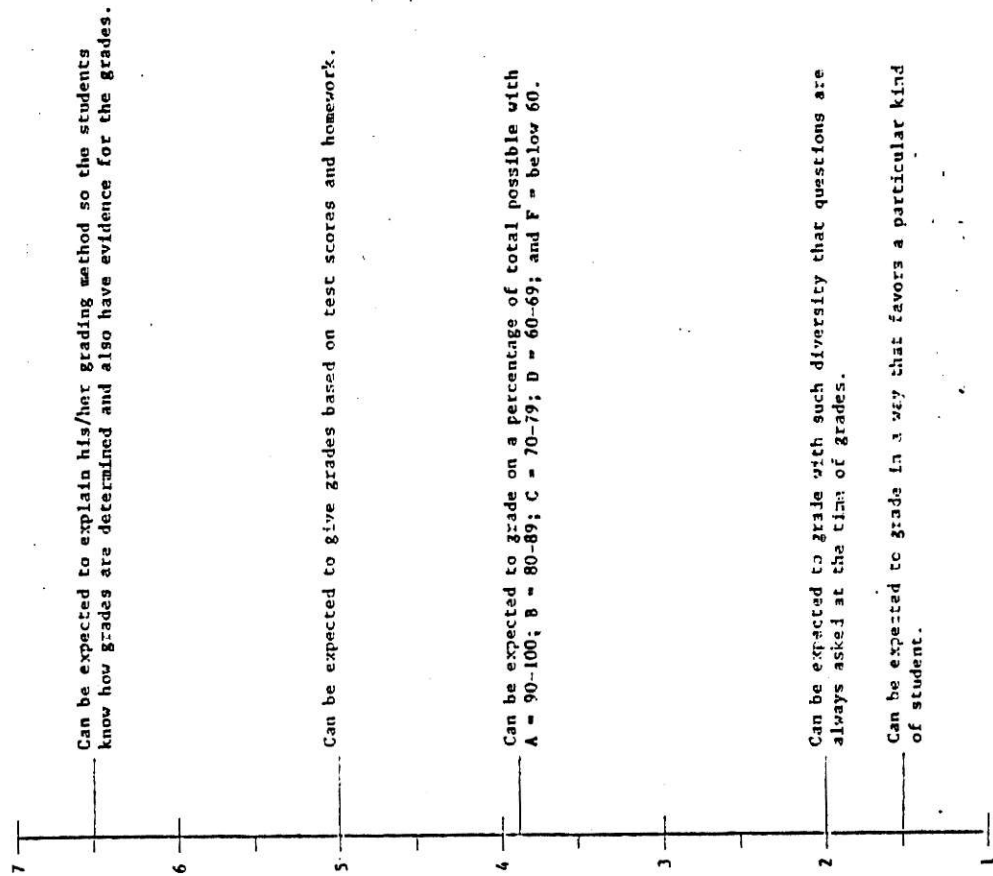
ASSIGNMENTS: Extent to which the instructor is clear on what is to be done, avoids assigning excessive amounts, and provides assignments which contribute to the understanding of the subject matter rather than just providing busy work.



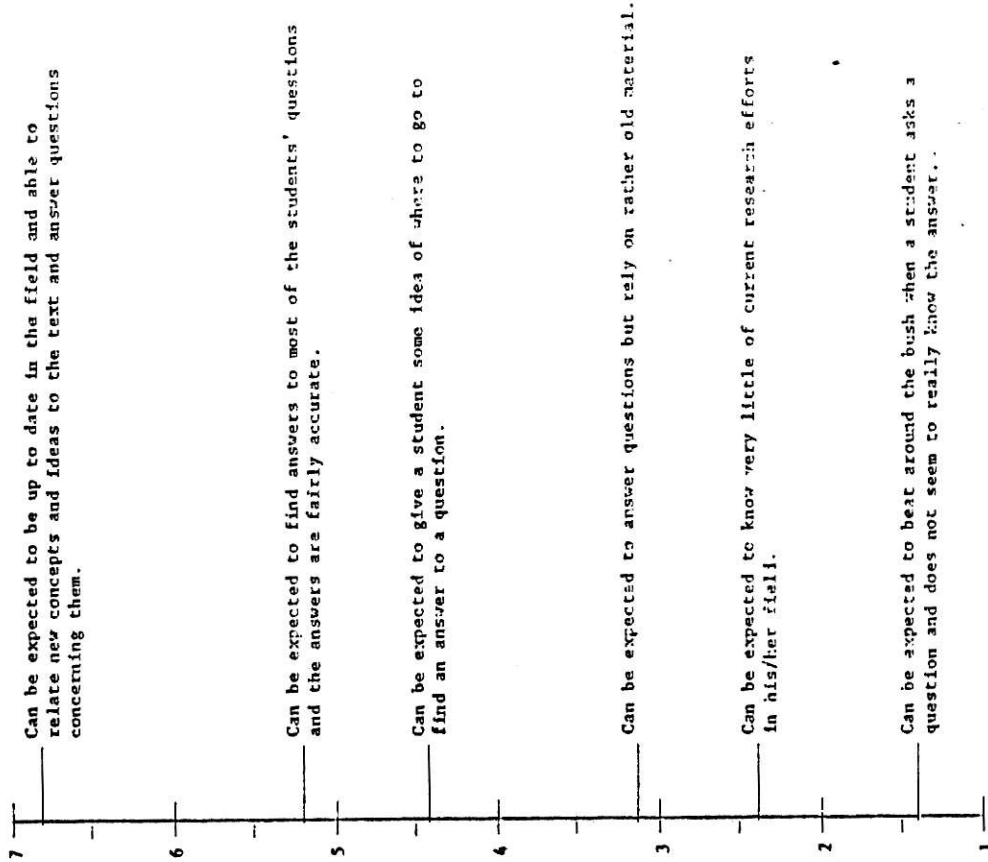
ATTITUDE TOWARDS SUBJECT: Extent to which the instructor shows personal interest in the material and displays a positive attitude towards teaching the subject.



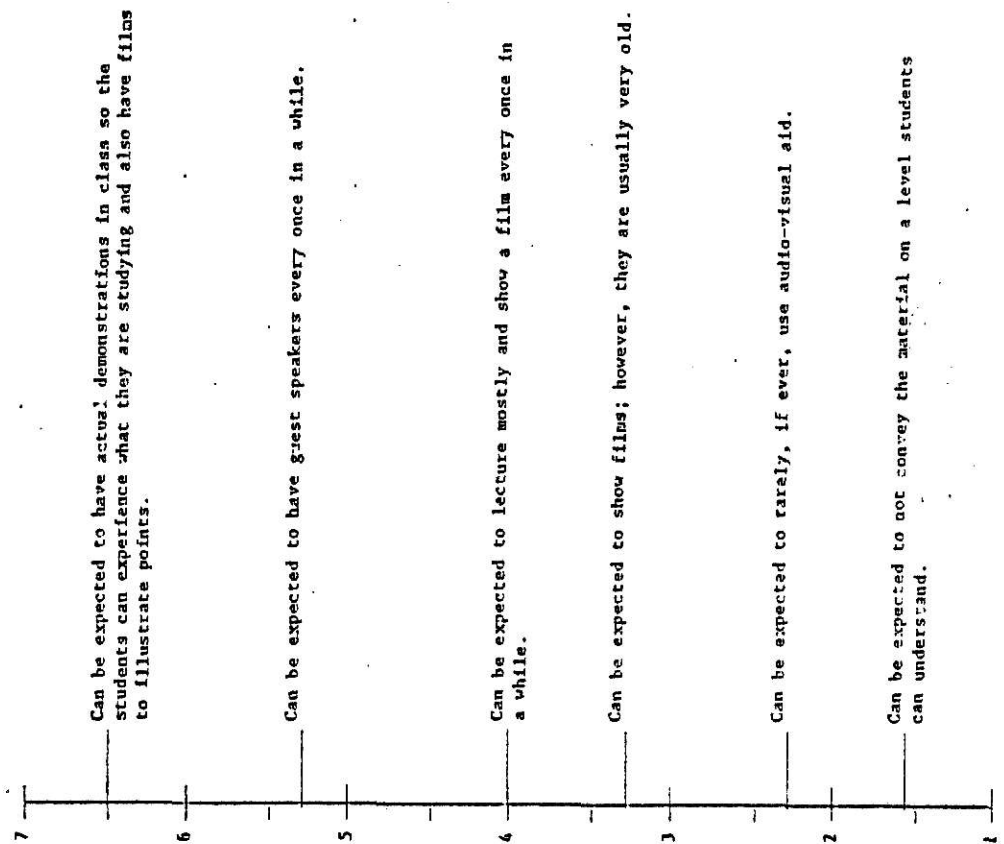
GRADES: Extent to which the instructor's grading practices remain consistent and free of confusion and are also fair.



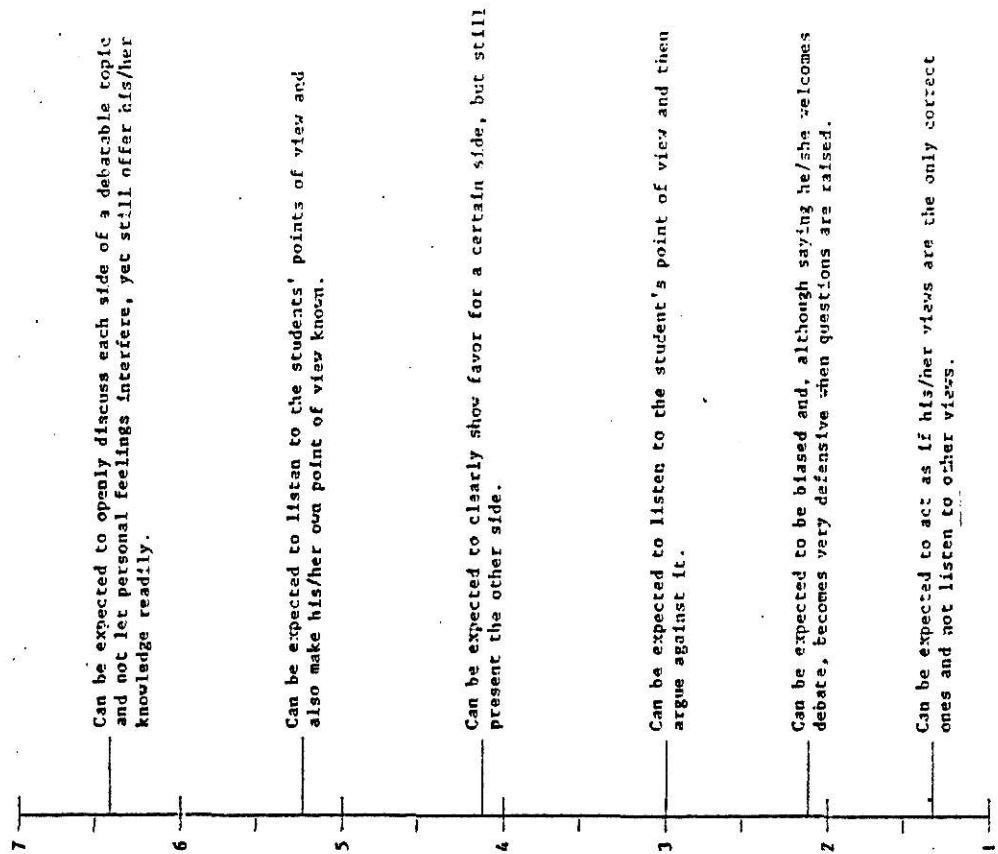
INSTRUCTOR KNOWLEDGE: Extent to which the instructor is aware of current material related to the course or to his/her field and is able to accurately answer or direct the student to specific sources that will answer questions concerning the subject matter.



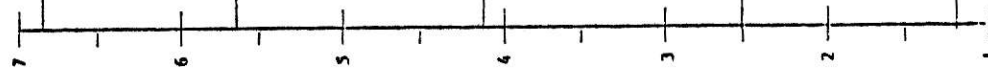
MANNER OF PRESENTATION: Extent to which the instructor's methods of presentation and use of audio-visual aids help emphasize and clarify important points; ability to present material clearly and concisely on a level students can understand.



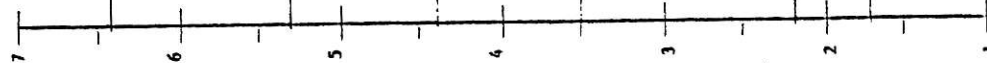
OBJECTIVENESS: Extent to which the instructor remains objective and presents a fair treatment of all points of view on controversial or debatable topics.



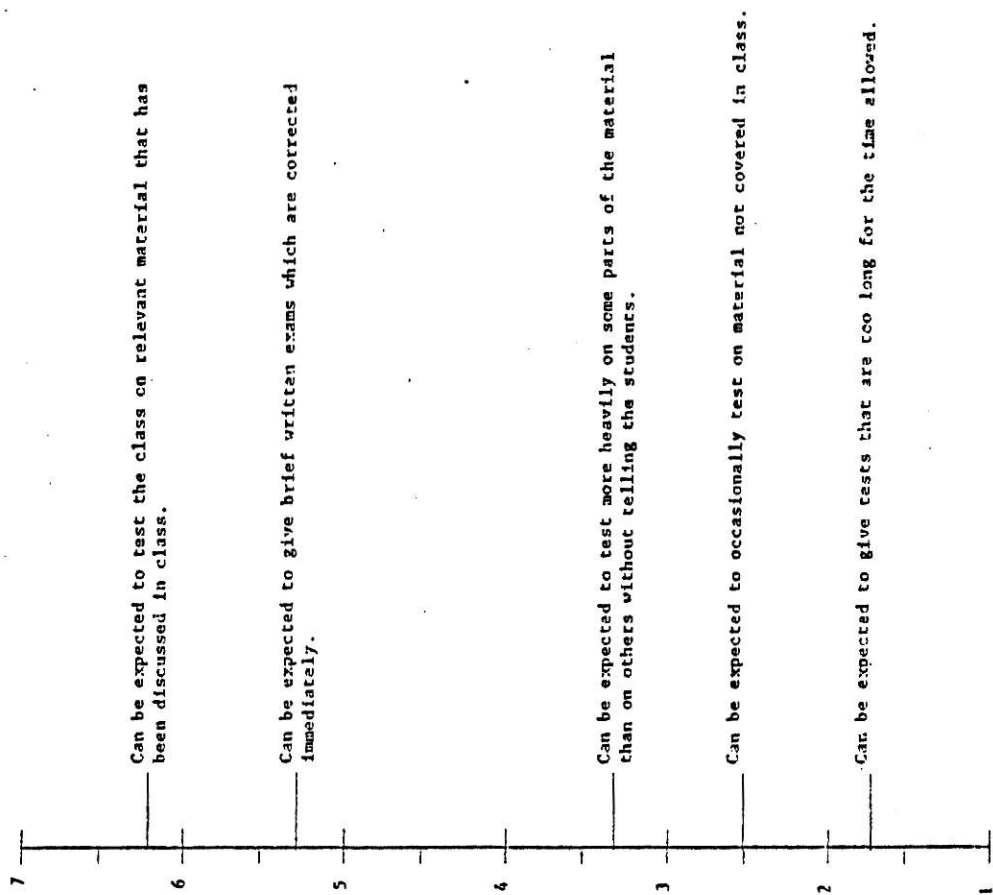
STUDENT-TEACHER RELATIONS: Extent to which the instructor shows a true, sincere concern for the welfare of the students through such things as dependability, availability for help, and consideration of student feelings; establishing rapport with the students.



ORGANIZATION: Extent to which the instructor arranges the subject matter and course objectives in an orderly and logical sequence for thorough coverage.



TESTS: Extent to which the instructor writes clear, unambiguous questions that relate to and are representative of in-class material and outside readings which were stressed adequately in class.



Appendix III

		Raw Scores							
		Differentiation			Hierarchic Integration			Halo	
S#		1	2	3	1	2	3	1	2
I N T R O D U C T O R Y	1	9.00	7.00	14.00	.333	.452	.363	.511	.577
	2	14.00	19.00	20.00	.214	.290	.214	.546	.000
	3	9.00	10.00	6.00	.319	.289	.027	.656	1.117
	4	13.00	11.00	10.00	.276	.182	.178	1.061	.774
	5	9.00	8.00	10.00	.667	.607	.400	.229	1.744
	6	8.00	9.00	11.00	.250	.264	.200	.888	.354
	7	26.00	12.00	12.00	.040	.227	.386	.458	.486
	8	11.00	12.00	15.00	.427	.235	.343	1.151	1.253
	9	15.00	8.00	10.00	.405	.536	.556	.888	1.279
	10	11.00	10.00	9.00	.190	.211	.208	.219	.157
	11	11.00	12.00	14.00	.291	.402	.368	.550	.565
	12	11.00	7.00	9.00	.118	.191	.250	1.197	.336
	13	10.00	11.00	11.00	.244	.263	.346	.687	1.330
	14	9.00	8.00	15.00	.139	.161	.095	1.014	1.206
A D V A N C E D	15	6.00	6.00	5.00	.200	.333	.067	.589	.717
	16	16.00	7.00	10.00	.213	.357	.233	1.000	.762
	17	10.00	8.00	10.00	.100	.304	.200	.800	.989
	18	9.00	9.00	9.00	.069	.236	.111	1.530	1.264
	19	11.00	7.00	7.00	.136	.143	.214	.668	1.247
	20	7.00	7.00	10.00	.310	.381	.322	1.359	.721
	21	12.00	10.00	7.00	.167	.233	.262	1.244	.731
	22	12.00	9.00	10.00	.083	.083	.056	.234	.118
	23	17.00	11.00	11.00	.136	.309	.218	1.523	1.394
	24	19.00	7.00	8.00	.137	.405	.429	.898	.731

Appendix III (cont'd)

S#	<u>Differentiation</u>			<u>Hierarchic Integration</u>			<u>Halo</u>	
	1	2	3	1	2	3	1	2
25	11.00	16.00	9.00	.146	.263	.319	.786	1.167
26	10.00	13.00	12.00	.167	.282	.189	.973	.755
27	11.00	13.00	13.00	.309	.283	.205	.797	1.043
28	7.00	8.00	9.00	.500	.321	.375	1.263	1.262
29	20.00	13.00	13.00	.084	.273	.263	1.161	1.179
30	12.00	8.00	5.00	.485	.250	.167	.724	.645
31	9.00	7.00	7.00	.139	.167	.119	.891	1.264
32	11.00	6.00	10.00	.218	.316	.222	.083	.579
33	9.00	9.00	13.00	.222	.292	.263	1.085	.951
34	7.00	10.00	6.00	.238	.256	.567	.533	.540
35	11.00	20.00	14.00	.264	.211	.198	.632	.425
36	35.00	8.00	12.00	.029	.232	.235	1.957	1.373
37	9.00	7.00	7.00	.024	.310	.429	1.397	.705
38	4.00	5.00	3.00	.167	.450	.333	1.732	1.027
39	10.00	8.00	9.00	.067	.446	.194	.882	1.516
40	6.00	10.00	7.00	.023	.089	.143	.963	1.374
41	13.00	11.00	10.00	.128	.812	.133	.612	.520
42	9.00	7.00	8.00	.333	.405	.339	.726	.664
43	19.00	15.00	14.00	.076	.167	.385	1.322	1.450
44	8.00	6.00	4.-0	.393	.533	.667	1.197	.437
45	14.00	14.00	13.00	.247	.302	.346	.387	.740

G
R
A
D
U
A
T
E

A DEVELOPMENTAL ANALYSIS OF RATING BEHAVIOR

by

KARL W. KUHNERT

B.A., The Pennsylvania State University, 1977

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1981

Schneider's (1977) cognitive compatibility theory of rating behavior, which states that when a rater's cognitive complexity (ability to differentiate among multiple dimensions of complex stimuli) is congruent with the complexity of rating scale formats, ratings will be psychometrically superior, has received a great deal of attention (e.g., Dunnette & Borman, 1979; Landy & Farr, 1980). Other researchers, however, have been unable to replicate these results (Bernardin & Boetcher, Note 1; Lahey & Saal, in press; Vance & Kuhnert, Note 2).

Without questioning the underlying assumptions, it has been assumed that a state of equilibrium, stability and rest is a more realistic picture of the rater than a state of upheaval, uncertainty and change. Landy & Farr's (1980) "process model implies that the rater's experience with ratings affects the validity of those ratings." They further state: "We know little or nothing about the effects which decisions based on current ratings have on future ratings". In other words, the history and development of a rater cannot be ignored.

The purpose of this paper is to suggest that the performance appraisal process may be profitably examined from a developmental perspective, namely the organismic-holistic orientation of Werner (1957). This framework states that "organisms are directed toward states of increasing differentiation and hierarchic integration" (p. 126). Interestingly, these concepts have also been commonly used to define cognitive complexity. Therefore, it is through differentiation and integration that the criteria for cognitive development are established. If raters gain increased differentiation and hierarchic integration in their perceptions of ratees with increased experience, then development must be acknowledged as a phenomenon that contributes meaningfully to the rating process. If raters' cognitive complexity (defined in terms of both differentiation and hierarchic integration) does develop over time, and if Schneider's (1977) cognitive compatibility model is valid, then passage of time should be

associated with systematic changes in the psychometric characteristic (e.g., halo) of raters' appraisals.

Cognitive complexity measures were obtained from 45 students in three psychology classes representing three different academic levels (introductory upper, graduate) at three times during a 15-week semester. Adopting the method described by Zajonc (1960), differentiation was based on the number of adjectives used by students to describe their psychology professor and integration on the degree to which those adjectives are interdependent. In addition, ratings of instructors' performance were obtained (at the beginning and at the end of the semester) from each student using a nine-dimension behaviorally anchored rating scale developed according to the method described by Smith and Kendall (1963). Halo was assessed for each rater by calculating the standard deviation of ratings across the nine performance dimensions.

Repeated measures ANOVAS were used in order to determine if differentiation, hierarchic integration and halo measures changed over time. Intercorrelations among these measures are also reported. The results indicate that, although both differentiation and integration do seem to develop even over a short time, there is no corresponding change in the halo inherent in raters' ratings. This finding casts doubt upon Schneier's (1977) "compatibility" hypothesis.

More important, perhaps, is that raters' perceptions were shown to develop systematically over an admittedly short period of time. Why this development was not reflected in the ratings is an important area for future inquiry. In any case, from the standpoint of theory building and metatheory, the developmental framework may provide a way to understand rating behavior in its multifarious glory.