A Method for Knowledge Engineering in Clinical Decision Making

Sheila S. Giere
Utah State University

A METHOD FOR KNOWLEDGE ENGINEERING

IN CLINICAL DECISION MAKING

by

Sheila S. Giere

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Psychology

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1989

ACKNOWLEDGEMENTS

This study was conducted under the direction of Drs. Richard Baer and Joseph Ferrara. I would like to express my sincere appreciation to Dr. Baer for his patience and support and, in particular, for the benefit of his ever-positive perspective. I would also like to thank Dr. Ferrara for his encouragement, assistance, and unfailing sense of humor.

I would also like to thank Dr. Alan Hofmeister for his invaluable assistance in the design of this study, Dr. Keith Checketts for his assistance with the statistical analysis, and Dr. Michael Bertoch for his helpful suggestions.

<div align="right">Sheila S. Giere</div>

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

A Method for Knowledge Engineering

in Clinical Decision Making

by

Sheila S. Giere, Doctor of Philosophy

Utah State University, 1989

Major Professors: Richard Baer, Michael Bertoch

Department: Psychology

The purpose of this study was to validate the problem behavior evaluation section of an expert system computer program, Class.BD. Class.BD was developed to assist special education personnel in determining whether students qualify for special education services as behaviorally disordered/severely emotionally disturbed students.

The subjects were six individuals from the state of Utah who regularly 1) work with behaviorally disordered/severely emotionally disturbed students and 2) participate in multidisciplinary assessment teams. Three of the subjects were special educators, and three were school psychologists.

Specifically, this study investigated the impact of five behavioral factors on the subjects' ratings of the

seriousness of problem behaviors. The five behavioral factors were 1) the severity or nature of the problem behavior, 2) the frequency with which the problem behavior occurs, 3) the duration over which the problem behavior has been occurring, 4) the generality of the problem behavior or the number of school environments the behavior occurs in, and 5) the percentage of the student's peers who engage in the same behavior. For each behavioral factor, three levels of that factor were determined: high, moderate, and low. Problem behavior descriptions were developed by the researcher, each of which presented the five behavioral factors at a predetermined combination of levels. Of 65 problem behavior descriptions, 33 described externalized problem behaviors and 32 described internalized problem behaviors. Subjects were asked to rate the seriousness of each problem behavior description on an 11 point scale, where 1=mild and 11=severe.

The results showed high levels of agreement among subjects on ratings of seriousness of problem behaviors. There was also high agreement between the subjects' ratings and ratings generated by the Class.BD expert system. Thus, Class.BD was validated. Further, the subjects gave highly similar ratings to descriptions of externalized and internalized problem behaviors.

The results also indicated that the severity of the

problem behaviors had the most impact on subjects'
ratings. Subjects discriminated three levels of severity
but only two levels of frequency, duration, generality,
and percentage of peers.

Finally, the results provided support for the use of
analysis of variance as a viable method ofknowledge
engineering, i.e., extracting information about how
experts make decisions. Its superiority over traditional
interview methods is discussed.

(342 pages)

A METHOD FOR KNOWLEDGE ENGINEERING

IN CLINICAL DECISION MAKING

by

Sheila S. Giere

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Psychology

Approved:

_____   _____
Co-Chairman   Committee Member

_____   _____
Co-Chairman   Committee Member

_____   _____
Committee Member   Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

1989

CHAPTER I

STATEMENT OF THE PROBLEM

Need Statement

The special education classification category
"behaviorally disordered/seriously emotionally disturbed"
(BD/SED) is reserved for those students who, by virtue of
their emotional or behavioral problems, are unable to
profit from regular education and who need special
education intervention. Special education services for
such students were mandated by Public Public Law 94-142,
the Education for All Handicapped Children Act of 1975.

In discussing prevalence estimates of the number of
children requiring special education, Balow (1979) noted
that all children exhibit emotional and/or behavioral
problems at some point during their years in public
school. He estimated that 20 to 30% of all public school
students show signs of these problems at any given time.
For the vast majority, the problems are transient or not
severe enough to interfere with ability to function
academically. For an estimated 2 to 3% (Balow, 1979),
however, the problems are severe enough to interfere with
academic progress.

Many of the present definitions/guidelines provided
by state offices of education as well as by the federal

government lack the specificity that would allow school personnel to discriminate between those students whose problems are transient, situational, or related to lack of effective behavior management and those students whose problems are long term, pervasive, and serious enough to require special education services (Executive Committee of the Council for Children with Behavioral Disorders, 1987).

Class.BD, an artificial intelligence expert system, was developed to reduce the problems associated with classification of students with behavioral disorders for the purpose of special education placement. Expert systems are computer programs that are designed to replicate the decision-making processes used by knowledgeable and experienced humans. The Class.BD computer program was designed to provide a second opinion regarding the classification of a student as BD/SED. One section of the program contains rules and facts that are used to evaluate the degree of seriousness of the student's behavior. These rules assign weights to factors of a student's problem behaviors (e.g., severity, duration) as well as to behavior-checklist or rating-scale scores and other sources of information. By combining the weights assigned to each factor of the student's problem behavior, checklist scores, and other sources of information, a "certainty factor" for the

seriousness of behavior problems is calculated. A certainty factor is a numerical value between 0 and 100 that is an index of belief that the conclusion is correct (Shortliffe & Buchanan, 1984). Prior to this study, the rules by which Class.BD calculated certainty factors were based on the consensus of the members of a Utah State Office of Education (USOE) task force on criteria for classification in the special education category "behaviorally disordered/seriously emotionally disturbed" (see Appendix A). The rules needed to be verified against practices currently being applied by educators in the field.

Problem Statement

Expert systems are a promising tool for assisting in the identification and classification of children in need of special education services. However, if they are to be useful in assisting with complex clinical decisions, such as the determination of whether a student's behavioral problems are serious enough to require special education intervention, the means by which practitioners make complex clinical decisions must be assessed.

In the particular case of Class.BD, this kind of clarification was needed to verify the nature of relationships between the factors. Also, there was a need to determine whether the weights that had been

assigned to the factors used to evaluate problem behavior and the rules used to combine these weights resulted in realistic, defensible conclusions. Information was needed on the relationship between the weights assigned to each factor of a problem behavior in the Class.BD computer program and the clinical impressions of persons who regularly make special education placement decisions concerning students with behavioral problems. It remained to be seen whether individuals who make such decisions would be in agreement with one another concerning their ratings of specific behaviors as well.

Class.BD assigns weights to the dimensions of problem behavior based on the magnitude of the dimension (i.e., number of days/weeks/months it has been occurring, number of times per day/week/month it occurs). These weights are then combined to obtain a final certainty factor. Certainty factors can range from 0 to 100 and provide a numerical index of the degree of certainty that a student's behavior problems are serious enough to warrant special education intervention. Prior to the present study there was no objective evidence to indicate that the factors determined by the USOE task force were actually useful cues for persons evaluating the magnitude of a problem behavior. It remained to be demonstrated whether each of these factors is important in the decision-making process. What was needed was an

investigation that tested the validity of the rules for determining the seriousness of behavior problems suggested by the USOE task force and incorporated into the CLASS.BD program against the rules used by experienced field-based decision makers. Thus, the investigators were able to look at 1) which of the five specific factors were important to the field-based decision makers when considerating the seriousness of students' problem behavior, 2) whether there was consistency among the field-based decision makers in their ratings on the seriousness of problem behaviors, and 3) whether there was consistency among ratings of field-based decision makers and the certainty factors generated through the problem behavior evaluation section of the Class.BD expert system.

CHAPTER II

REVIEW OF THE LITERATURE

Expert Systems

Expert systems are a subclass of artificial intelligence, a broad area of computer science. Artificial intelligence, in general, is that branch of computer science that is concerned with designing computer programs that simulate human characteristics. Expert systems are a type of artificial intelligence in which programmers attempt to develop computer programs that generate high levels of accurate performance in problem- solving tasks. Tasks appropriate for expert systems are those that require years of specialized training for human beings to achieve similarly high performance levels (McCoy & Levary, 1988). An expert system is programmed to present the user with a series of questions. Users type in their answers on the computer keyboard; the answers are then compared to the information in the computer program knowledge base (i.e., rules in the program). The programs are designed to use the input from the user to arrive at a conclusion or problem solution.

The majority of the early expert system applications were in complex scientific areas in which objective data are used in determining recommendations. For example,

PROSPECTOR (Duda, Gaschnig, & Hart, 1979) generates expert advice on finding ore deposits based on geological data. MYCIN (Shortliffe, 1976) generates advice on diagnosis and therapy of infectious diseases. MACSYMA (Martin & Fateman, 1971) provides assistance to mathematicians in the area of symbolic computations associated with applied analysis. In these and other such programs, objective scientific data are employed, and complex areas of human expertise are simulated.

With the recent development of personal-computer expert system software, expert system technology has become available to and more practical for practitioners in a wide variety of fields. Expert system technology has been applied to solving problems in business and industry such as "diagnosis of engine failures, tax planning, and feasibility analysis of cases in union disputes about seniority" (Olson & Reuter, 1987, p. 152).

The staff of the Technology Division of the Developmental Center for Handicapped Persons at Utah State University has developed a number of expert systems for use in classification of and program development for students with handicaps. These systems are based on the rules and regulations provided by the federal government and the Utah State Office of Education (USOE) and include the following systems: Class.LD2 (Ferrara & Hofmeister, 1984) is an expert system developed to assist in

determining whether students qualify for services because of a learning disability. 'Mandate Consultant' (Parry, 1986) is an expert system designed to assist special education personnel in ensuring that they have met the regulatory requirements for IEP development according to the rules put forth in Public Law 94-142.   Class.IH (Ferrara, Williams, & Giere, 1987) is designed to assist multidisciplinary assessment teams in determining whether a student qualifies for special education services because of an intellectual handicap.   Class.PH (Gold & Peterson, 1988) is designed to assist in identifying and suggesting interventions for students with physical handicaps.   While developing the knowledge bases for these programs required some knowledge-clarification activities, the majority of the rules for identifying particular student problems were based on empirical findings and straightforward rules (i.e., specific cutoff scores on standardized tests, presence or absence of specific symptoms).   For the Class.BD program (Ferrara, Baer, Althouse & Reavis, 1988), the rules and regulations are not as easily interpretable, and an objective base for determining program rules is not available.

Current Practices in Knowledge Acquisition
for Expert Systems.

The most popular means of determining rules and important facts has been through a series of intensive interviews with a person identified as an expert in the field (Hayes-Roth, Waterman, & Lenat, 1983). The knowledge engineer (the individual responsible for determining the rules to be programmed into the system) using this method may also observe the expert performing the specific task and ask him/her to talk through the process of performing the task. Audio- and videotapes of the expert talking through and performing the task are also often employed. Though this method is the one most frequently reported in the literature on expert system development, it is also an extremely time-consuming and often ineffective means of extracting knowledge from an expert. That the expert made a specific statement or carried out a specific task while involved in a demonstration does not necessarily mean that this information is relevant to performance in the area of expertise. And, even if the information is relevant to task accomplishment, this does not mean that the information is important to task accomplishment (Hoffman, 1987). Also, if applied to a task that is not usually verbalized, it is possible that the problem-solving process will be distorted in that the expert will report

what he/she thinks he/she should report rather than what he/she actually does.

Another aspect of this method to which Hoffman (1987) raised objections is the proposal that knowledge engineers use only one expert in order to avoid contradictions. Rather than considering contradictions undesirable, Hoffman (1987) asserted that "disagreements should be used as clues about the basic research that might be needed to fill in the knowledge gaps..." (p. 60). On this same issue, Triggs (1988) noted that we cannot necessarily rely on an expert to detect all of the problems in a developed expert system. That individual may be impressed with the expert system and may attribute more expertise to the computer program than is warranted. Triggs (1988) also asserted that "the expert systems' 'mode of operation' will almost certainly not conform exactly to the process used by any one expert" (pp. 716-717).

A number of authors have suggested other means for extracting knowledge from human experts. McCoy and Levary (1988) suggested incorporating procedures used in human performance modeling, an area of industrial psychology concerned with the study of how humans interact with machines. The suggested techniques include function analysis and task analysis for performance of each job task. While the inclusion of such methods

would, as the authors asserted, be of value in adding structure to the process of knowledge acquisition and in reducing interviewing time, the methods appear appropriate only for those tasks in which performance is readily observed and the important rules are easily verbalized.

In an article outlining suggested methods for knowledge acquisition based on experimental procedures used in cognitive science, Olson and Reuter (1987) suggested two general methods of extracting knowledge from human experts: direct and indirect. Direct methods include interviews, questionnaires, protocol analysis, interruption analysis (interrupting the expert when the knowledge engineer has a question), and drawing closed curves (drawings indicating relationships between objects). Indirect methods are designed to elicit information that experts themselves have not been able to verbalize and include such methods as a) Multidimensional scaling, a method used on data that are assumed to have come from stored representations of physical n-dimensional space; b) Johnson hierarchical clustering, a method for which the assumption that each item is a member of a cluster or not a member of a particular cluster of items is essential; c) general weighted networks, a method for assessing networks of associations between items; d) ordered trees

from recall, a method in which recall trials are used to investigate how data are produced by a subject; and e) repertory grid analysis, a method that uses a combination of dialog between the subject and experimenter, a rating session, and an analysis that clusters the items rated and the dimensions on which the items were rated.

In an article outlining possible methods for extracting knowledge, Hoffman (1987) suggested a number of methods. Among these are a) Method of Familiar Tasks, which involves an analysis of the tasks the expert usually performs; b) Structured Interviews, where the expert is queried regarding knowledge of facts and procedures; c) Limited Information Tasks, in which the expert is asked to perform a familiar task under circumstances in which some of the usually available information is not presented; d) Constrained Processing Tasks, where the expert is asked to perform a familiar task under time or other constraints; and e) Method of "Tough Cases," in which a familiar task is performed using data from a case that is difficult for the expert. The methods presented by Hoffman (1987) appear to be best suited for those knowledge domains where the expert's performance is easily observed and verbal reports of underlying rules are relatively accessible.

In an article concerned with methods of extracting

implicit knowledge, Berry (1987) suggested two possible methods: protocol analysis and machine induction. Protocol analysis involves observing an individual performing the task and asking him/her to provide a running commentary while carrying out the task. Berry (1987) found this method objectionable in that the results are often incomplete, and the act of producing a running commentary can affect the way a task is actually carried out. Machine induction involves entering a number of examples of different types of decisions from the domain and using a computer program to apply an inductive algorithm to discover the simplest set of rules to describe those examples. This method is useful for those domains in which documented cases are available. And the accuracy with which the algorithm can be applied depends on the representativeness of the cases used in developing it. In addition, as Berry (1987) notes, the rules that are induced will not necessarily be the same as the ones a human expert uses.

The drawback in using these indirect techniques is that their use involves underlying assumptions about the form of the representation of objects and their relations (Olson & Reuter, 1987); i.e., the assumption that the physical representations/models, such as lists, decision trees, networks, and so forth, are appropriate ways to look at human knowledge. Such prior assumptions can lead

to grossly distorted conclusions about the nature of the subject matter domain, as well as about the ways in which human beings actually use and store information.

In summary, research is needed to develop methodologies for extracting implicit knowledge from human practitioners (Berry, 1987). While the authors of the above mentioned articles provide suggestions for how this might be accomplished, little actual research has been done on the question of how to elicit experts' knowledge and inference strategies (Hoffman, 1987).

<u>Knowledge Representation in Expert System</u>
<u>Computer Programs</u>.

In the vast majority of expert systems, the knowledge base is written such that inferences are programmed as "if...then" rules and facts are programmed as "object-attribute-value" rules (Olson & Reuter, 1987). For many knowledge domains in which expert systems have been developed, such facts and "if...then" inferences have an empirical bases. Without those empirical bases, however, it is inappropriate to assume that such facts and inferences exist for a given domain.

A few authors have proposed alternative means for best analyzing the data used to determine the knowledge base of an expert system. In an article discussing the cognitive science research regarding how human beings

process uncertain information, Hink and Woods (1987) suggested first developing a computer program that is capable of automatically acquiring objective information through design and implementation of a usage log. As in the machine-induction method suggested by Berry (1987), the data base would contain detailed records of each consultation that could be used to test the validity and appropriateness of the knowledge represented in the system. The authors suggested that one way to incorporate both the objective and subjective information provided by the domain expert is through the use of contingency tables. The knowledge engineer asks the domain expert to fill in 2 x 2 contingency tables based on the expert's knowledge and experience. For example, consider the simple rule, "If A, then B." The domain expert is asked to judge the probability of each of the following four statements being true: 1) If A, then B; 2) if A, then not B; 3) if not A, then B; and 4) if not A, then not B. The values in each cell are relative estimates of the expert's confidence in the truth of the "if...then" rule. This analysis may be appropriate and advantageous for verifying knowledge in many areas of expertise, particularly those in which "if... then" rules or dichotomies on each dimension of each variable are a realistic representation of those variables.

In a study of the ways judges weigh and combine

items of information in the process of expert decision making, Triggs (1988) reported the use of the conventional analysis of variance to analyze the information rendered by experts. The main effects are taken to be indicative of linear-cue utilization on the part of the expert, and the interaction effects are indicative of configural-cue utilization. Triggs (1988) noted that while many expert judges claim to use configurations of cues extensively, the ANOVA has not provided strong support for this; data indicate that only a small percentage (5-8%) of the systematic variance can be attributed to configural cue usage. In two particular studies, however, the researchers found that 21% and 33% of the variance was associated with configural-cue usage. In two studies comparing the configural-cue usage of experts versus that of novices, the findings indicate that experts use configural cues to a large degree, and novices almost never use configurations of cues in making their judgments. It appears, then, that when writing expert systems such that the functioning best resembles expert judgment, one cannot assume that the factors that go into making that judgment are independent. Domain experts tend to consider relationships between factors, and these relationships have an impact on their judgment processes and final decisions. Triggs does not directly suggest using the ANOVA results to program the knowledge

base; rather, he considers the ANOVA results to be a basis for further interviews with the domain expert and to assist in understanding that expert's performance.

In summary, it is clear that the way in which information is best represented and used in an expert system knowledge base depends a great deal on the nature of the particular knowledge domain. The number of possible outcomes, the number of variables, and the nature of those variables all need to be considered when determining the way in which the information is best analyzed. It is also clear that very little objective research has been conducted to date on validating different ways of analyzing the data developers' input into expert systems.

Problems in Identifying Students With

Behavioral Disorders.

Problems in Definition.

A frequent complaint of practioners charged with identifying students with behavioral disorders for the purpose of special education placement is that the definitions and guidelines provided by the federal government in Public Law 94-142 as well as those provided by individual state offices of education are ambiguous and subject to widely disparate interpretations.

In reviewing definitions of BD/SED of departments of educationof various states, Epstein, Cullinan, and Sabatino (1977) identified the presence or absence of 11 components: disorders of emotion/behavior, interpersonal problems, learning/achievement problems, deviation from norms, chronicity, severity, etiology, prognosis, specification of factors that would exclude a student from classification, special education needed, and certification (approval of classification by some individual or group or determined through specific assessment procedures). The authors found little agreement as to what constituted a behavioral disorder. In addition, they noted that a medical and psychiatric rather than an educational orientation was frequently present and that the terminology for specifying how to

assess a student's problems was vague and not very useful.

In a follow-up study, Cullinan, Epstein, and McLinden (1986) compared the definitions used in 1976 to those used in 1982. They found significant increases in the number of states that emphasized chronicity and severity and that specified rules for excluding certain students from classification. A statistically significant decrease was observed in the the number of states that included "deviation from the norm" as a component, thus indicating less concern for the social significance of behavior problems.

In a similar study, Mack (1980) compared the definitions found in state special education regulations with the definition provided in Public Law 94-142 (Federal Register, 1977; provided in Appendix B). She found that the definitions used in 12 states addressed all the criteria outlined in P.L. 94-142. The definitions used in only 35 of the states included the modifier "over a long period of time and to a marked degree." In 40 of the state definitions the factor "adversely affects education performance" was specified. Over two-thirds of the definitions failed to mention socially maladjusted children at all; of those who did, 2 states included socially maladjusted children in the same definition with emotional disturbance, and 3 states

defined socially maladjusted separately for special education classification purposes.

In a study to assess whether definitional ambiguity, confusion, and lack of consensus all exist within a single state, Swartz and Mosley (1986) surveyed the directors of special education in 75 school districts in Illinois. They found considerable variation in the definitions used from one school district to the next. Thirty-five percent of the respondents indicated that their districts had no procedural guidelines by which to identify behavioral disorders in children. With regard to definitions for the behaviorally disordered category, 24% of the respondents used the Illinois statutory definition, 18% used a modification of that definition, 35% used a locally developed definition, 4% used the P.L. 94-142 definition, and 12% had no official definition on record. The authors concluded that it is "arguable that the federally defined category, seriously emotionally disturbed, and the Illinois category are conceptually or in practice the same" (p. 11).

In summary, it appears that a student who is identified as behaviorally disordered/seriously emotionally disturbed in one school district might not qualify for special education in another state, in another school district within the same state, or even within another school in the same district.

Lack of Appropriate Identification Procedures.

In a survey designed to ascertain occurrence of agreement in procedures used in identifying students with behavioral disorders/serious emotional disturbance across states, Greenburg (1983) surveyed 23 special education administrators in 11 states. Because of confusion reported to him by local special education administrators regarding the definition and identification of behaviorally disordered/seriously emotionally disturbed students, Greenburg (1983) sought to obtain information as to the specific nature of the problems encountered by various special education directors. The problems most frequently reported by the survey respondents were related to vagueness in the definition, particularly associated with lack of guidance as to what constitutes a "severe" behavior problem, and the means to discriminate between "emotionally disturbed" and "socially maladjusted." The respondents found the definitions used in their districts unclear and subject to diverse interpretation. Greenburg (1983) also inquired as to the means by which students with serious emotional disturbance were identified in each school district. He found that there was a variety of personnel involved, depending on the local education agency (LEA). However, there was a great deal of similarity in the data

considered from one LEA to the next and in the general consistency in the process used. Because the survey was open-ended and the responses in narrative form, the findings may not represent the entire spectrum of problems encountered; a more carefully planned study might have yielded more specifics in identification of problems. Also, it cannot be assumed that because the data collected and processes followed were similar across LEAs that what occurred were valid practices. As Greenburg (1983) noted, "The consistency with which local planning units use criteria for differentiation should not serve to condone the practices in view of the debate concerning whether or not such differentiation can be authenticated" (p. 28).

Even when specific procedures are outlined and measures to be used are specified, the guidelines are not always used. In a study to determine what information is used to make placement decisions about students with behavioral disorders, McGinnis, Kiraly, and Smith (1984) evaluated the files of 45 behaviorally disordered students classified as needing special education services in a local educational agency (LEA) in Iowa. The LEA had, as a matter of policy, six specific assessments that were considered essential in classifying a student as having a behavioral disorder. These were personality, observation-structured, observation-clinical

interpretation, anecdotal record, standardized behavior-rating scale, and affective assessment. Even though these measures were specified, the researchers found very little evidence that they were actually used. Rather, the most prevalent information found in the student files were observers' general impressions of the student's behavior. The second most frequently found information was a student's family/environmental history. The authors concluded that "the emphasis appears to be placed on subjective rather than objective data" (p. 245).

In summary, the present definitions provide little guidance to multidisciplinary teams for determining how to assess students for identifying behavioral disorders and for defining precisely what constitutes a behavioral disorder. What is needed then, is a coherent, defensible set of guidelines that can be applied consistently by school personnel to identify those students who, by virtue of their emotional/behavioral problems, require special education intervention.

Class.BD

In an attempt to develop a coherent, defensible definition of behavioral disorders that can be applied consistently and be useful to multidisciplinary assessment teams, the Utah State Office of Education established a task force to develop such a definition. The task of this group was to define those factors that should be considered in determining whether a student can be classified as behaviorally disordered/seriously emotionally disturbed (BD/SED) and to specify decision rules for considering those factors in making BD/SED classification decisions. Based on the recommendations of this task force, the staff of the Technology Division of the Developmental Center for Handicapped Persons at Utah State University developed an artificial intelligence expert system, Class.BD (Ferrara et al., 1988), that incorporates the factors and decision rules developed by the task force. Some recommendations of the task force were also incorporated into a recent revision of the State Board of Education Special Education Rules (Utah State Office of Education, 1988). A copy of the USOE rules for classifying students with behavioral disorders can be found in Appendix C.

The definition provided by the USOE defines a BD/SED student as "one whose behavior or emotional condition over a long period of time and to a marked degree

adversely affects his/her educational performance" (p. A-29). Thus, two conditions must be met in order to classify a student as BD/SED: demonstration that the behavioral problem(s) are serious enough to warrant a special education placement and demonstration that the behavioral problems have adversely affects the student's educational performance.

Specific factors programmed into the Class.BD computer program include a number that identifies adverse effects on educational performance and a number that determines the seriousness of the student's behavior problems. Factors considered in assessing adverse effects on educational performance include the student's academic achievement, grades, and citizenship. Nine factors are considered in evaluating the seriousness of a student's behavioral problem. Five of these factors relate directly to a description of the student's problem behavior(s): 1) severity, 2) frequency of occurrence; 3) duration (length of time over which the problem has been consistently occuring), 4) percentage of the student's peers that exhibit comparable behavior, and 5) generality or the number of school settings in which the problem behavior occurs. Three sources of information external to the school environment are also considered: social service agency reports, law enforcement agency reports, and parent reports of problem behavior outside of school.

A final factor considered is the student's scores on standardized behavioral checklists or rating scales.

Another behavioral factor programmed into the Class.BD expert system is whether the behavior problem is "externalized" or "internalized." According to the USOE BD/SED classification guidelines (see Appendix C), externalized problem behaviors are those where the student acts out against someone or something in his/her social environment and usually involves excesses of behavior. Internalized problem behaviors are those that usually involve behavioral deficits, with the child reacting to his/her problems by withdrawing from the social environment. The certainty factors assigned to externalized and internalized problems are the same but have been included in the program to draw the attention of school personnel to the fact that internalized problem behaviors require attention. As some authors have noted, there is a tendency for classroom teachers to overlook students with internalizing problems (Walker, Reavis, Rhode, & Jenson, 1985).

Class.BD is programmed to combine the certainty factors for each behavioral factor using a "variance" model. Under this model, the certainty that a student's problems are serious enough to warrant special education can range from 0 to 100. An example of how individual certainty factors combine is as follows: A student sets

fires on the school campus once per day. The certainty factor for severity (setting fires) is high, 20, subsuming 20% of the variance; 80% of the variance remains to be subsumed by other factors. The certainty factor for a frequency of "once per day" is 15; this factor will subsume 15% of the remaining variance, or 15% of 80 = 12. The combined certainty factors for severity (20) and frequency (15) then is 20 + 12 = 32. The certainty factors for each subsequent variable entered into the Class.BD computer program are combined in the same manner, with each subsuming a percentage of the remaining variance. The variance model allows for combining an infinite number of factors in any order without ever exceeding a total certainty factor of 99.

To summarize, a number of factors have been identified as important in determining whether problem behavior(s) are serious enough to warrant a BD/SED classification. However, we cannot assume that these factors are taken into consideration in a consistent manner across practioners in the field. Nor can we assume that the presence of each of these factors actually influences how special education decision makers evaluate the seriousness of a student problem behavior. We need to know if human decision makers use these factors and dependencies (or configurations of cues) in

the BD/SED classification decision-making process and, if so, how and under what circumstances.

CHAPTER III

METHOD

Purpose and Objectives

In 1988, the Utah State Office of Education (USOE) revised its guidelines for assessing behaviorally disordered/seriously emotionally disturbed students. As a part of the revision process, the USOE formed a task force that suggested considering five behavioral factors in determining if problem behaviors exhibited by students are serious enough to warrant a behavioral disorders classification. Their ideas have been programmed into an expert system computer program, Class.BD. Although the guidelines suggested by the BD/SED task force represent its best thinking about how behavior problems should be considered in making classification decisions, there was no evidence that their ideas reflected the actual decision-making practices of field-based decision makers. The purpose of this study was to address this issue by validating the BD/SED Task Force/Class.BD decision rules against decision rules used by knowledgeable field-based decision makers. Through the validation process important information was gained regarding 1) what rules field-based decision makers use in assessing student behavior problems, 2) the degree to which different field-based decision makers use similar rules in

assessing student behavior problems, and 3) the degree to which rules used by field-based decision makers in assessing behavior problems are the same as those suggested by the BD/SED task force and incorporated into the Class.BD expert system.

The objectives of this study were two fold. The first objective was to investigate a means for clarifying complex clinical decision-making practices. That is, the methods used in this research provide a means for studying how human decision makers use individual variables and associations between the variables in making complex clinical decisions.

The second objective was to attempt to verify a portion of the rules programmed into the knowledge base of the computer program Class.BD. Specifically, this study addressed the rules for weighing the five factors of problem behavior (severity, frequency, duration, generality, and percentage of peers exhibiting similar behavior) and how those weights combine to yield an overall index of seriousness of a problem behavior.

Method

Population and Sample

The population was public school personnel in the state of Utah who were regularly involved in making

behaviorally disordered placement decisions. Three special education personnel and three school psychologists experienced in working with multidisciplinary assessment teams and making behaviorally disordered classification decisions served as subjects. Subjects were included based on the recommendations of 1) the person who works for the USOE and is in charge of BD/SED programs across the state, 2) the person in charge of the Program Administrative Reviews (PARs) under contract to the USOE, and 3) an individual designated to work on the PAR team.

The task of the Program Administrative Review team was twofold. The first was to assess the degree of compliance of agencies serving the educational needs of handicapped students with state and federal regulations concerning classification of students with handicaps. The second was to assess the individualized educational plans for those students. Each school district in Utah as well as other agencies serving the educational needs of handicapped students and receiving 94-142 funds is subject to a PAR review once every three years.

A master list of individuals who, in the estimation of these three people, were particularly good at making BD/SED classification decisions was requested. Subjects who participated in this study were chosen from this list based on willingness to participate and meeting

additional criteria.

It had been recommended that persons involved in developing a knowledge base should not be involved in the verification of that knowledge base because their prior involvement may bias their evaluations (O'Keefe, Balci, and Smith, 1987; Geissman & Schultz, 1988). As a result, persons involved in the BD/SED task force that determined the variables to be considered in a behaviorally disordered/seriously emotionally disturbed classification decision were not eligible to participate as subjects in this study.

Additional criteria included 1) the individual must have interacted in a professional capacity with at least five behaviorally disordered students in the past school year; 2) the individual must have had at least five years experience working in special education or school psychology; 3) the individual must have held a master's degree in special education, school psychology, or a related field; and 4) the individual must not have been cited by the USOE for any ethical violations. The qualifications of persons on the master list were checked against the five criteria. Three individuals from each discipline were selected to participate. Table 1 lists the occupations, years of experience, number of behaviorally disordered students assessed/taught during the 1987-88 school year, and the highest academic degree

attained for each subject.

Table 1.

Subject Characteristics

	Occup-pation	Years Experience	# BD/SED Students	Degree
Subject 1	Teacher	6	5	Educ. Spec.
Subject 2	Special Ed. Coordinator	6	35	Masters
Subject 3	Special Ed. Coordinator	10+	100+	Masters
Subject 4	School Psychologist	6	40+	Masters
Subject 5	School Psychologist	17	50+	Ph.D.
Subject 6	School Psychologist	15	50+	Ph.D.

Problem Behavior Descriptions

Sixty-five descriptions of problem behaviors were developed by the researcher. Each presented five behavioral factors the BD/SED task force deemed important in assessing behavior problems (severity, frequency, duration, generality, and percentage of peers). Each factor was presented at three possible levels: high, moderate, and low. While one factor varied across the three levels, the other four factors were all presented at one of the three levels. Thus, there were 33 possible combinations. Thirty-three problem behavior descriptions

were developed describing externalized problem behaviors and 32 problem behavior descriptions were developed describing internalized problem behaviors. The USOE defines externalized problem behaviors as those that are directed toward the environment, i.e., the student acts out. Externalized problem behaviors usually involve behavioral excesses. The USOE defines internalized problem behaviors as those where the student withdraws from the social environment and that usually involve behavioral deficits. One possible combination was omitted (internalized, low severity, low frequency, low duration, high generality, and low percentage of peers) because it was found to be logically impossible to have high generality on a behavior that almost never occurred (less than once per month) and had been occurring for less than one month. The combinations of levels of factors used to develop the problem behavior descriptions are shown in Table 2. All of the problem behavior descriptions can be found in Appendix E. An example follows of an externalized problem behavior where severity was at a high level, frequency was at a moderate level, duration was at a high level, generality was at a high level, and percentage of peers was at a high level:

> Mike fights with other students on the playground and in class; these fights usually result in injury to the other student (black eyes, etc.). Mike picks fights with other students an average of four times per week. This has been going on for the last 6 months. Mike has picked fights with other students

in each one of his classes; he has not done this in the lunch room. Approximately 2% of the students in Mike's grade in his school initiate fights.

Table 2.

List of Combinations of Levels of Factors

	Externalized							Internalized				
	S	F	D	G	%			S	F	D	G	%
1.	H	M	M	M	M		34.	H	M	M	M	M
2.	M	M	M	M	M		35.	M	M	M	M	M
3.	L	M	M	M	M		36.	L	M	M	M	M
4.	M	H	M	M	M		37.	M	H	M	M	M
5.	M	L	M	M	M		38.	M	L	M	M	M
6.	M	M	H	M	M		39.	M	M	H	M	M
7.	M	M	L	M	M		40.	M	M	L	M	M
8.	M	M	M	H	M		41.	M	M	M	H	M
9.	M	M	M	L	M		42.	M	M	M	L	M
10.	M	M	M	M	H		43.	M	M	M	M	H
11.	M	M	M	M	L		44.	M	M	M	M	L
12.	H	H	H	H	H		45.	H	H	H	H	H
13.	M	H	H	H	H		46.	M	H	H	H	H
14.	L	H	H	H	H		47.	L	H	H	H	H
15.	H	M	H	H	H		48.	H	M	H	H	H
16.	H	L	H	H	H		49.	H	L	H	H	H
17.	H	H	M	H	H		50.	H	H	M	H	H
18.	H	H	L	H	H		51.	H	H	L	H	H
19.	H	H	H	M	H		52.	H	H	H	M	H
20.	H	H	H	L	H		53.	H	H	H	L	H
21.	H	H	H	H	M		54.	H	H	H	H	M
22.	H	H	H	H	L		55.	H	H	H	H	L
23.	L	L	L	L	L		56.	L	L	L	L	L
24.	M	L	L	L	L		57.	M	L	L	L	L
25.	H	L	L	L	L		58.	H	L	L	L	L
26.	L	H	L	L	L		59.	L	H	L	L	L
27.	L	M	L	L	L		60.	L	M	L	L	L
28.	L	L	H	L	L		61.	L	L	H	L	L
29.	L	L	M	L	L		62.	L	L	M	L	L
30.	L	L	L	H	L							
31.	L	L	L	M	L		63.	L	L	L	M	L
32.	L	L	L	L	H		64.	L	L	L	L	H
33.	L	L	L	L	M		65.	L	L	L	L	M

(S=severity, F=frequency, D=duration, G=generality, %=percent of peers).

As Table 3 shows, high, moderate and low levels of frequency, duration, generality, and percentage of peers were quantifiable and could be easily represented within the problem behavior descriptions. Severity, however, was not quantifiable and required judgments as to what represents a nuisance, disruptiveness, mild emotional upset, and so forth. Because of the subjectivity, interobserver agreement was obtained between the researcher and a second observer regarding the level of severity represented in each problem behavior description.

Table 3

Levels of the Five Factors

	Low	Moderate	High
Severity externalized	nuisance	disruptive	threatens safety of persons or property
Severity internalized	slightly limits social interactions/ mild emotional upset	limits social interactions/ moderate emotional upset	precludes social interaction/ severe emotional upset
Frequency	<= 1/month	several/week	several/day
Duration	< 1/month	1-6 months	> 6 months
Generality	0-33%	34-66%	67-100%
Percentage of Peers	20%+	10-19%	0-9%

Agreement was also obtained as to whether the behavior represented an internalized or externalized problem behavior. The second observer was a graduate student who had been involved in the development of Class.BD and the criteria for each level of the five factors. In addition to the 65 problem behavior descriptions, he was provided specific definitions for externalized and internalized behavior problems and for high, moderate, and low severity behavior problems. These definitions can be found in Appendix D and in Table 3. The second observer was asked to read each problem behavior description and indicate whether he believed the severity of the problem fell into the low, moderate or high range, as defined in Class.BD. He was also asked to indicate whether the problem behavior fit the definition for internalized or externalized as outlined in the USOE guidelines. The observer's responses were compared to those of the researcher who had developed the problem behavior descriptions. When there was disagreement, a new behavior description for that combination of levels of each factor was developed. This continued until 100% agreement was reached.

Design of the Study

The 65 problem behavior descriptions were presented to the six subjects who were asked to rate their seriousness

on an 11-point scale.

Of interest was:

1. The degree of agreement between the subjects and Class.BD regarding the seriousness of the behaviors described.

2. Whether there were differences in the subject's ratings when a behavioral factor (the factor under consideration; i.e., severity, frequency, duration, generality, or percentage of peers) had high versus moderate versus low levels. For example, when severity was the behavioral factor under consideration, was there a difference in ratings when severity was at a high level versus a moderate level versus a low level? Or, when frequency was the behavioral factor under consideration, was there a difference in ratings when frequency was at a high versus a moderate versus a low level? The difference of interest here is shown in Figure 1. Along the right-hand side of each cube in the figure a factor under consideration that can take on a high, moderate, or low value is illustrated.

3. Whether there were differences among the subjects' ratings when the other four behavioral factors (other four factors) were at high, moderate, or low levels. In other words, when severity was the factor under consideration, was there a difference in ratings when the four factors other than severity were at a high versus a

moderate versus a low level? When frequency was the factor under consideration, was there a difference in ratings when the four factors other than frequency were at a high versus a moderate versus a low level? and so forth. The differences of interest here are also shown in Figure 1. Along the bottom of each block in the figure four other factors that can have high, moderate, or low values is illustrated.

4. Whether there were differences in the subjects' ratings of seriousness for problem behavior descriptions of internalized problem behaviors and problem behavior descriptions of externalized problem behaviors. The difference of interest here is also shown in Figure 1 along the top, right-hand side of each block.

5. Whether there were interactions among the levels of the variables (externalized/internalized, level of the factor under consideration, and levels of the other four factors).

Formally stated, the hypotheses tested in the study were:

Hypothesis 1: Given problem behavior descriptions, there are no differences among ratings of the seriousness of the problem behaviors provided by field-based decision makers experienced in making BD/SED classification decisions and certainty factors for those problem behaviors generated using the Class.BD computer program.
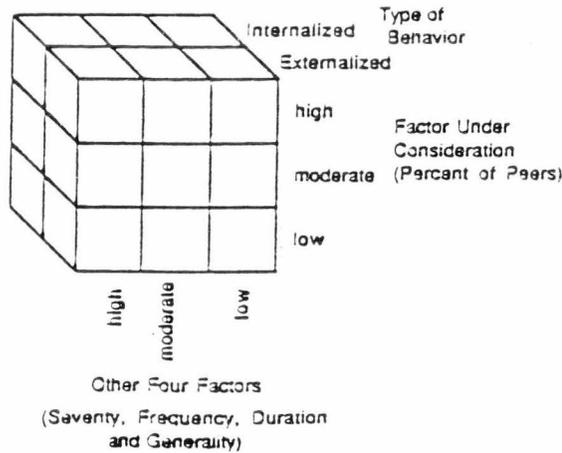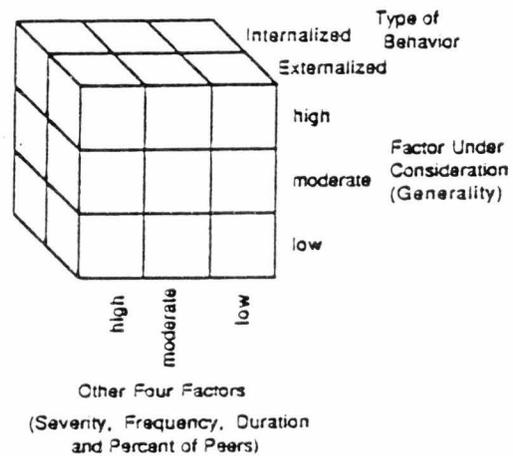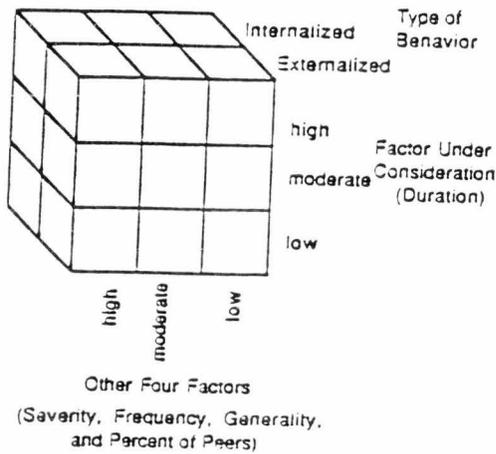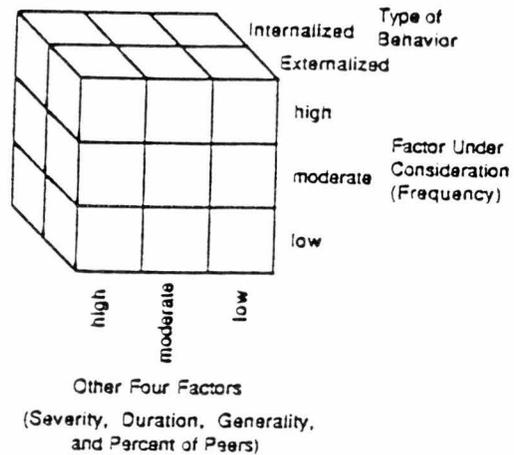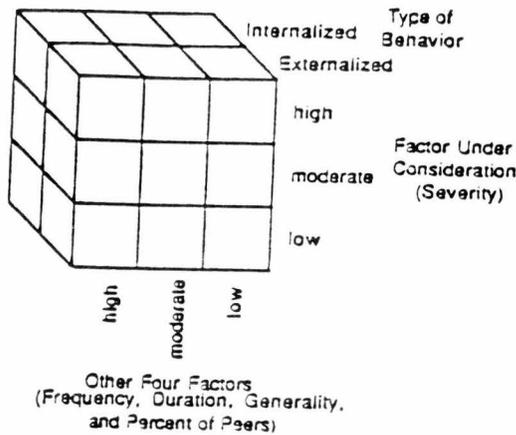
Figure 1. Comparisons between levels of factors

Hypothesis 2: Given problem behavior descriptions of internalized and externalized problem behaviors that are equal in terms of the levels of five behavioral factors described (severity, frequency, duration, generality, and percentage of peers), there are no differences in the ratings of seriousness provided by field-based decision makers on internalized versus externalized problem behavior descriptions.

Hypothesis 3: Given problem behavior descriptions that present one of five behavioral factors at high, moderate, or low levels, there are no differences in ratings of seriousness provided by field-based decision makers when that factor is at high versus moderate versus low levels.

Hypothesis 4: Given problem behavior descriptions that present four behavioral factors at high, moderate, or low levels, there is no difference in ratings of seriousness provided by field-based decision makers when the four factors are at high versus moderate versus low levels.

Hypothesis 5: There is no interaction between ratings of seriousness provided by field-based decision makers on problem behavior descriptions describing externalized and internalized problem behaviors and ratings of seriousness on problem behavior descriptions that present one of five behavioral factors at high, moderate, or low levels.

Hypothesis 6: There is no interaction between ratings of seriousness provided by field-based decision makers on

problem behavior descriptions describing externalized and internalized problem behaviors and ratings of seriousness on problem behavior descriptons that present four behavioral factors at high, moderate, and low levels.

Hypothesis 7: There is no interaction between the ratings of seriousness provided by field-based decision makers on problem behavior descriptions that present one of five behavioral factors at high, moderate, or low levels and the ratings of seriousness on problem behavior descriptions that present four behavioral factors at high, moderate, or low levels.

Hypothesis 8: There is no interaction between the ratings of seriousness provided by field-based decision makers on problem behavior descriptions that present externalized and internalized problem behaviors, one of five behavioral factors at high, moderate, or low levels, and four behavioral factors at high, moderate, or low levels.

Procedures

Presentation of the Problem Behavior Descriptions

The special education personnel and psychologists participating were presented with the 65 descriptions of problem behaviors. The descriptions were placed in random order and presented to each subject in the same order. Subjects were asked to rate the seriousness of

the problem behavior on an 11-point scale, where 1=mild, 6=moderate, and 11=serious. This rating scale is similar to that used in the method of equal-appearing intervals (Edwards, 1957). It was chosen to present a scale that is sufficiently sensitive to pick up differences in ratings across the 65 behaviors and between ratings of the subjects and Class.BD. The subjects were also asked if there were any factors not presented in the descriptions that they felt were important in determining the seriousness of the problem behavior. The instructions given to the subjects can be found in Appendix F.

Reliability

Approximately 2 weeks after the subjects returned their ratings on the 65 problem behavior descriptions to the researcher, each subject was asked to re-rate 10 of the problem behavior descriptions to obtain a measure of reliability. The 10 problem behavior descriptions were chosen at random, with 5 externalized and 5 internalized problem behavior descriptions selected for re-rating.

Follow-up Interviews

When the subjects provided disparate ratings (3 or more points of difference on the scale) on the original 65 problem behavior descriptions, qualitative and

quantitative data were collected in order to determine the reasons for the discrepancy between subjects. The individual giving the disparate rating as well as two subjects whose ratings were similar to the rest of the subjects was interviewed regarding how he/she arrived at his/her decision. The researcher interviewed the subjects regarding their reasons for giving particular ratings on those cases. The subjects were asked which pieces of information provided in the problem behavior description were important in determining the rating they gave it. They were also asked to rank order each factor according to its importance in influencing their decisions. They were asked to describe their decision-making processes and to provide rationales concerning why the factors were important. A copy of this semistructured interview format can be found in Appendix G. In this manner, different perspectives were compared.

Data Analyses

Hypothesis 1: Given problem behavior descriptions, there are no differences among the ratings of seriousness of the problem behaviors provided by field-based decision makers experienced in making BD/SED classification decisions and certainty factors for those problem behaviors generated using the Class.BD computer program.

Agreement among the six subjects and Class.BD was

computed using Kendall's W coefficient of concordance. Kendall's W is a nonparametric measure of agreement among raters that is used when there are ratings on several variables by each rater. Kendall's W was run repeatedly with different combinations of subjects/Class.BD ratings. In this way the researcher was able to determine which source(s) of ratings were providing scores or patterns of scores that were different from the patterns of scores provided by the others. Chi-square tests were applied to determine whether the associations between ratings were statistically significant at the .05 level.

In addition, Pearson's r correlation coefficients were computed to assess the level of association between each pair of subjects' actual ratings of the problem behavior descriptions. Pearson's r was also computed between the certainty factors generated by the first and second versions of Class.BD and the ratings provided by each subject.

Differences between the subjects' ratings were also tested by means of a one-way analysis of variance as follows:

ANOVA

	df
Between subjects	5
Within subjects	384
Total	389

The Fisher's LSD multiple comparison method was applied to determine which pairs of mean ratings were responsible for the statistically significant F ratio.

Hypothesis 2: Given problem behavior descriptions of internalized and externalized problem behaviors that were equal in terms of the levels of five behavioral factors described (severity, frequency, duration, generality, and percentage of peers), there are no differences in the ratings of seriousness provided by field-based decision makers of internalized versus externalized problem behavior descriptions.

Hypothesis 3: Given problem behavior descriptions that present one of five behavioral factors at high, moderate, or low levels, there are no differences in ratings of seriousness provided by field-based decision makers when that factor is at high versus moderate versus low levels.

Hypothesis 4: Given problem behavior descriptions that present four behavioral factors at high, moderate, or low levels, there are no differences in ratings of seriousness when the four factors are at high versus moderate versus low levels.

Hypothesis 5: There is no interaction among ratings of seriousness of problem behavior descriptions of externalized and internalized problem behaviors and ratings of seriousness of descriptions that present one of five behavioral factors at high, moderate, or low

levels.

Hypothesis 6:   There is no interaction between ratings of seriousness of problem behavior descriptions externalized and internalized problem behaviors and ratings of seriousness of descriptons that present four behavioral factors at high, moderate, and low levels.

Hypothesis 7:   There is no interaction between the ratings of seriousness of problem behavior descriptions that present levels of the factor under consideration and the ratings of seriousness of descriptions that present four behavioral factors at high, moderate, and low levels.

Hypothesis 8:   There is no interaction among the ratings of seriousness of problem behavior descriptions that present externalized and internalized problem behaviors; the ratings of the seriousness of problem behavior descriptions that present high, moderate, and low levels of the factor under consideration; and the ratings of seriousness of descriptions that present four behavioral factors at high, moderate, and low levels.

Three-way analyses of variance were used to address Hypotheses 2 through 8.   The following analysis of variance was conducted five times, once with each of the five factors as the factor under consideration.   These analyses are illustrated in Figure 1.

ANOVA

	df
External/Internal (E/I)	1
Factor Under Consideration (FUC)	2
Other Four Factors (O4F)	2
E/I x FUC	2
E/I x O4F	2
FUC x O4F	4
E/I x FUC x O4F	4
Error	85

The analyses of variance were blocked on the subjects variable to control for variability associated with subject differences, enabling more precise comparisons among the means for the other variables in the ANOVA (Ott, 1988). The externalized/internalized main effect was tested to determine whether statistically significant differences existed between ratings of those problem behavior descriptions in which the behaviors were designated as internalized versus those in which they were designated as externalized (Hypothesis 2, E/I). The F ratio for the factor under consideration main effect (severity, frequency, duration, generality, or percentage of peers) was tested to determine whether significant differences existed among the ratings of problem behavior descriptions in which that factor had high, moderate, or low values (Hypothesis 3, FUC). The main effect for the other four factors was tested to determine if statistically significant differences existed between the ratings of the problem behavior descriptions where the

other four factors were at high versus moderate versus low values (Hypothesis 4, O4F). The interaction between the factor under consideration and externalized/internalized variables was tested to determine whether the ratings of those problem behavior descriptions in which the factor under consideration was at high, moderate, or low values varied in significantly different patterns when the behaviors were designated as internalized versus externalized (Hypothesis 5, E/I x FUC). The interaction between the other four factors by the externalized/internalized variable was tested to determine whether the ratings of those descriptions in which the other four factors were at high, moderate, or low values varied in the same patterns when the behaviors were designated as externalized versus internalized (Hypothesis 6, E/I x O4F). The interaction between the factor under consideration and the other four factors was tested to determine whether the ratings of descriptions where the factor under consideration was at high, moderate, or low levels differed by the same amount under conditions where the other four factors were at high, moderate, and low values (Hypothesis 7, FUC x O4F). The interaction among externalized/internalized, level of the factor under consideration, and level of the other four factors was tested to determine whether the ratings varied in similar patterns when all three variables were

taken into account (Hypothesis 8, E/I x FUC x O4F).

Fisher's protected Least Significant Difference (LSD) test (Ott, 1988) was used to determine which means were statistically significantly different from one another when F ratios were statistically significant (p<.05). Fisher's test, which is a less conservative test than other multiple comparison methods, was chosen because of the high variability typically found in studies using human subjects. Because it is less conservative than most other multiple comparison methods, a researcher is more likely to find differences that actually exist between means. It is "protected" by the F test in that it is only applied when the F ratios for main effects or interactions are statistically significant. The error rate for the protected LSD is believed to be controlled on an experimentwise basis, with the alpha level approximately equal to that of the F test (Ott, 1988).

Interview Data

For each problem behavior description in which one or more subjects provided ratings three or more points different from the median rating of the majority of the subjects, that subject as well as two of the subjects giving average ratings was interviewed in an attempt to determine the reasons for the disparate ratings. The data obtained during the follow-up interviews were analyzed in

the following manner. Data obtained on the first question, which required subjects to rank order each factor in terms of its importance in influencing their rating, were analyzed using Kendall's W coefficient of concordance. Kendall's W is a nonparametric measure of agreement among raters or judges. Kendall's W was used to determine on which problem behavior descriptions there was significant disagreement among subjects regarding the importance of each factor and to determine whether the subjects ranked the factors in a similar manner, even though they provided disparate ratings of the problem behavior descriptions. Chi-square tests were applied to determine whether the associations between the rankings were significant at the .05 level.

In addition, Spearman's rho was computed between each pair of subject's rankings of each problem behavior description on which they ranked the five factors to determine whether the two persons giving average ratings ranked the five factors in a manner more similar to each other than to the rankings of the person giving a disparate rating. A frequency count was taken of the responses to why they ranked the factor under consideration where they did and how they approached the problem of determining the seriousness of the behavior

problem in each problem behavior description to look for consistencies across subjects.

CHAPTER IV

RESULTS

The objectives of this study were to 1) evaluate how well the Class.BD expert system emulates the decision-making process used by special educators and psychologists in evaluating the seriousness of student problem behaviors; 2) determine whether field-based decision makers (FBDMs) agree regarding the seriousness of problem behavior descriptions; and 3) if the FBDMs were in agreement, to attempt to clarify the rules they use in rating the seriousness of problem behavior descriptions. In particular, do field-based decision makers differentially rate internalized and externalized problem behaviors in determining seriousness, and do different levels of five behavioral factors set the occasion for significant differences in ratings on the problem behavior descriptions? In addition, the reliability of the field-based decision makers' ratings was assessed to see if their ratings would be consistent over time.

The specific research hypotheses appear in the previous section. The reliabilty for the ratings and the results related to the research hypotheses are as follows.

Reliability of Ratings

Approximately 2 weeks after the subjects completed their initial ratings, 10 of the 65 problem behavior descriptions were randomly selected (5 externalized and 5 internalized). These 10 descriptions were sent to the subjects to rate again. This was done to get a measure of reliability of the subjects' ratings. A pair of ratings for a problem behavior description was considered to be in agreement if there were 2 or fewer point differences. The agreement between the first and second ratings was calculated by dividing the number of agreements by 10. The results are presented in Table 4.

Table 4.

Reliability of Ratings on Problem Behavior Descriptions

Subject	Agreement
1	.90
2	.80
3	.60
4	.60
5	.90
6	.90
Overall Reliability	.78

Overall, the subjects provided similar ratings on the ten problem behavior descriptions 78% of the time.

Hypothesis 1: Agreement Between Class.BD and Subjects

Agreement between the six subjects and Class.BD was computed by changing the rating data and the Class.BD certainty factors into ranks (based on 65 ranks) and computing Kendall's W coefficient of concordance. A chi-square test was applied to determine whether the association between ranks was statistically significant at the .05 level. These results are shown in Table 5.

In Table 5, BD1 refers to the original version of Class.BD. BD2 refers to a revision, done in March 1989, based on preliminary findings from this research. The X's denote which computer program/subjects were used in each analysis. As illustrated in the table, all the combinations of computer programs/subjects resulted in high Ws (range .7978 to .8318). The chi-squares associated with each combination were highly statistically significant. A slight increase in Kendall's W was observed with the March revision of the behavior problem section over the original. Computing Kendall's W with any five of the six subjects' ranks did not change Kendall's W appreciably (range .8254 to .8546).

In addition, Pearson's r correlation coefficients were computed between the actual ratings provided on the problem behavior descriptions for all combinations of pairs of subjects and versions of Class.BD. These are

shown in Table 6. Scatterplots for each of these correlation coefficients can be found in Appendix H. As shown in Table 6, the correlations among the subjects ratings on the 65 problem behavior descriptions were quite high, ranging from .6461 to .8717. The correlations between the subjects' ratings and the certainty factors in the two versions of Class.BD were also high, with an increase in the correlation coefficients observed with the second version of Class.BD. All the Pearson's r's were statistically significant, with p<.001.

Table 5.

Kendall's W Coefficient of Concordance.

BD1	BD2	S1	S2	S3	S4	S5	S6	W	Chi Square
X		X	X	X	X	X	X	.7978	357.4172
	X	X	X	X	X	X	X	.8198	367.2486
		X	X	X	X	X	X	.8318	319.4000

For all Chi Square values p<.0001

A one-way analysis of variance was also conducted to test whether there were significant differences among the mean ratings given by the FBDMs. The results are shown in Table 7.

Table 6

Pearson's r Correlations Coefficients Among the Ratings
Provided by the Six Subjects and Two Versions of Class.BD

	S1	S2	S3	S4	S5	S6
BD1	.6921	.6978	.6534	.6659	.6733	.6761
BD2	.7744	.7404	.6591	.7014	.8113	.7643
S1		.7737	.7386	.7915	.7683	.8122
S2			.7631	.8717	.7548	.8430
S3				.7779	.6461	.8168
S4					.7470	.8398
S5						.8026

For all r's p<.001

Table 7.

Analysis of Variance for Differences Among Subjects.

	df	MS	F
Between judges	5	48.62	
Within	384	10.51	
Total	389		4.63

The critical value of F with alpha set at .01 is 3.06. Thus, F=4.63 was statistically significant at the .01 level, indicating statistically significant differences among the mean ratings provided by the subjects.

To determine which pairs of means were different, Fisher's LSD multiple comparison procedure was used. The

means for each subject are listed in Table 8.

Table 8.

Mean Ratings for Subjects

Subject	1	2	3	4	5	6
Mean	7.00	5.75	4.66	4.92	5.74	4.92

For the Fisher's LSD test, for alpha=.05 a difference of 4.019 between means is required for a difference to be considered statistically significantly different. None of the differences between the means listed in Table 7 were significantly different according to this multiple comparison procedure.

These results indicate a high level of agreement among subjects on their ratings of the problem behavior descriptions. There is also a high level of agreement between subjects and the two versions of Class.BD. Thus, Hypothesis 1, which states that there are no significant differences between the ratings provided by the FBDMs and Class.BD, is supported.

Hypotheses 2 through 8

Hypotheses 2 through 8 were tested by means of 5 three-way analyses of variance. As described in the analysis section, the effects tested were (a) externalized versus internalized (Ex/In), (b) level of

factor under consideration (FUC), (c) level of other four factors (O4F), (d) Ex/In by FUC, (e) Ex/In by O4F, (f) FUC by O4F, and (g) Ex/In by FUC by O4F. An overall summary of the results is shown in Table 9. The various ANOVAs are listed down the right-hand column and are designated by the main factor under consideration. Across the top, the effects tested and the hypotheses they relate to are specified. Within the table the significance of each effect for each analysis is presented. The ANOVA tables can be found in Appendix I.

Table 9.

ANOVA Results.

Factor	Ex/In	FUC	O4F	E/I x FUC	E/I x O4F	FUC x O4F	3-way
Hypotheses	2	3	4	5	6	7	8
Severity	NS	<.001	<.001	.026	<.001	<.001	<.001
Frequency	.002	<.001	<.001	.002	.03	.007	.004
Duration	NS	.008	<.001	NS	.011	<.001	.005
Generality	<.001	.009	<.001	NS	<.001	NS	.026
% of Peers	NS	<.001	<.001	<.001	<.001	NS	.005

Hypothesis 2: Externalized Versus Internalized Problem Behaviors.

The F ratio for the externalized/internalized main effect was not statistically significant in the severity, duration, and percentage of peers ANOVAs. This main

effect was statistically significant in the frequency and duration ANOVAs. Comparison of the means for externalized and internalized problem behaviors revealed a difference of .8 points in both the frequency and duration ANOVAs.

Thus, three of the five analyses provided support for Hypothesis 2, that there is no difference between the ratings for externalized and internalized problem behaviors.

Hypothesis 3: Levels of the Factor Under Consideration.

In all five ANOVAs, the main effect for the factor under consideration was statistically significant. The results of the Fisher's LSD comparisons are listed in Table 10, which designates which pairs of means had statistically significant (S) differences and which had nonsignificant (NS) differences.

Hypothesis 3, that there are no differences in the ratings on problem behavior descriptions when the factor under consideration is at high, moderate, or low levels, is not supported. Significant differences were observed between the ratings when the factor under consideration was at moderate versus low levels. Significant differences were observed between the ratings for high

Table 10.

Pairwise Comparisons for Levels of the

Factor Under Consideration.

| | Comparison | | |
	High/Mod	Mod/Low	High/Low
Severity	S	S	S
Frequency	NS	S	S
Duration	NS	S	S
Generality	NS	S	S
% of Peers	NS	S	S

versus moderate levels of severity; there was no difference between the mean ratings on the problem behavior descriptions with high versus moderate levels of frequency, duration, generality, and percentage of peers.

Hypothesis 4: Levels of the Other
Four Factors.

In all five ANOVAs, the main effect for the other four factors was statistically significant. The results of the Fisher's LSD comparisons are listed in Table 11, which designates which pairs of means had statistically significant (S) differences and which had nonsignificant (NS) differences.

Hypothesis 4, that there are no differences among the ratings when the other four factors are at high, moderate, and low levels, is rejected. Significant differences were found among the mean ratings for high,

moderate, and low levels of the other four factors across all five ANOVAs.

Table 11.

Pairwise Comparisons for Levels of the

Other Four Factors.

| | Comparison | | |
	High/Mod	Mod/Low	High/Low
Severity	S	S	S
Frequency	S	S	S
Duration	S	S	S
Generality	S	S	S
% of Peers	S	S	S

Hypothesis 5: Interaction Between
Externalized/Internalized and
Level of the Factor Under Consideration.

The interaction between externalized/internalized problem behaviors and the level of the factor under consideration was statistically significant in three of the five ANOVAs: severity, frequency, and percentage of peers.

The interaction between externalized/internalized and level of severity is shown in Figure 2.

There were statistically significant differences among the mean ratings for high, moderate, and low levels of severity for externalized problem behaviors.
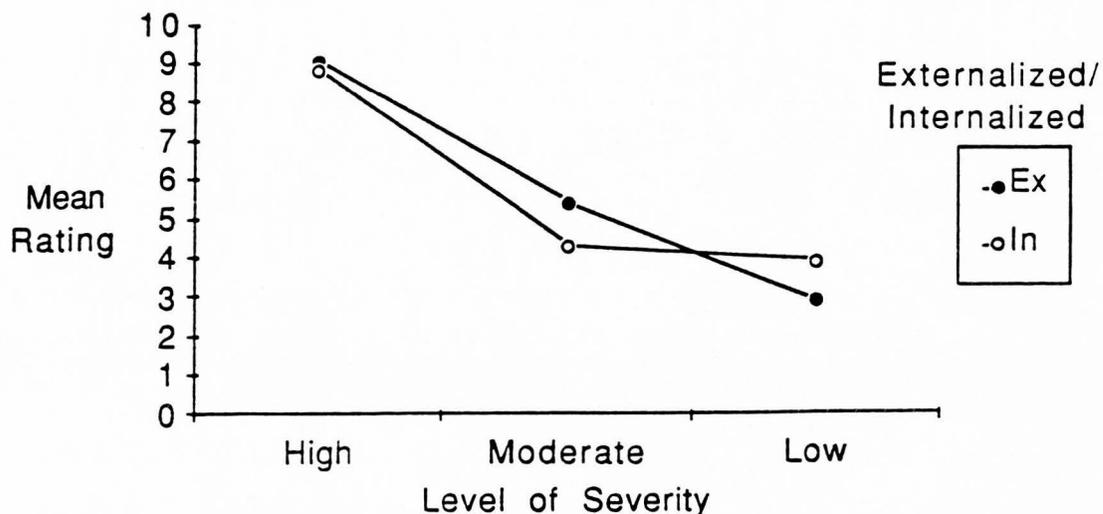
Figure 2 Externalized/internalized by level of severity.

There were statistically significant differences between high and moderate levels of severity for internalized problem behaviors. There was no difference between the mean ratings for moderate and low levels of severity for internalized problem behaviors. There was a significant difference between externalized and internalized problem behaviors when the level of severity was moderate. There was no difference between the mean ratings for externalized and internalized problem behaviors when the level of severity was high or low. The means and Fisher's LSD's for this interaction can be found in Appendix J.

The interaction between externalized/internalized and level of frequency is shown in Figure 3.
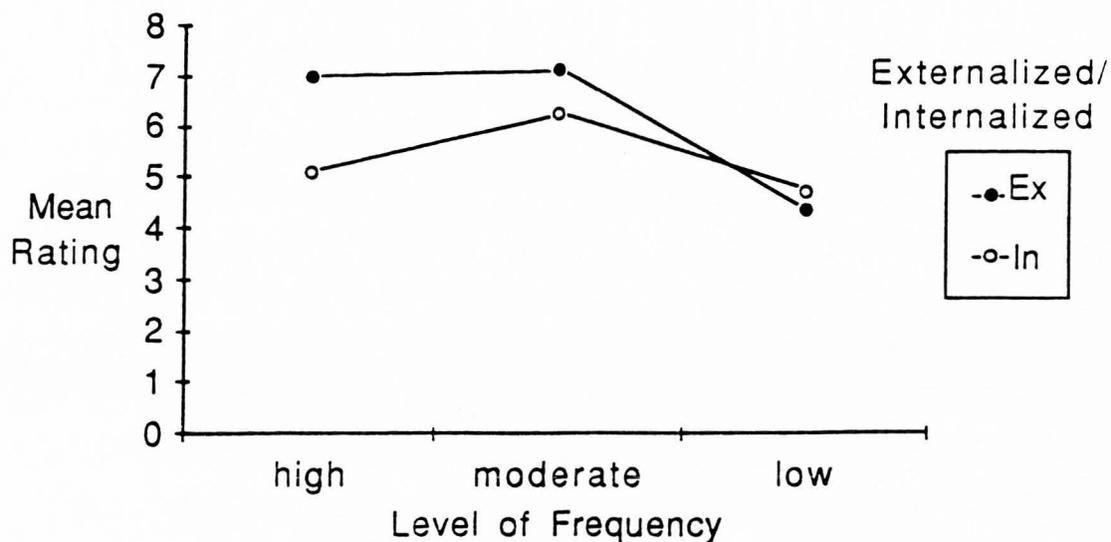
Figure 3 Externalized/internalized by level of frequency.

There were statistically significant differences between the mean ratings for high and moderate and between moderate and low levels of frequency when the behavior was internalized. However, there was no difference between the mean ratings for high and low levels of frequency for internalized problem behaviors. As is shown in Figure 3, the mean rating for moderate frequency for internalized problem behaviors was higher than the mean rating for high-frequency level internalized problem behaviors. There was a significant difference between the mean ratings for moderate- and low- level frequency externalized problem behaviors, but the mean ratings for high- and moderate-frequency externalized problem behaviors were not different. There

were significant differences between mean ratings for externalized and internalized problem behaviors when the level of frequency was high and when it was moderate. There was no difference between mean ratings for externalized and internalized problem behaviors when the level of frequency was low. The means and Fisher's LSDs for the frequency ANOVA can be found in Appendix K.

The interaction between externalized/internalized and level of percentage of peers is shown in Figure 4.
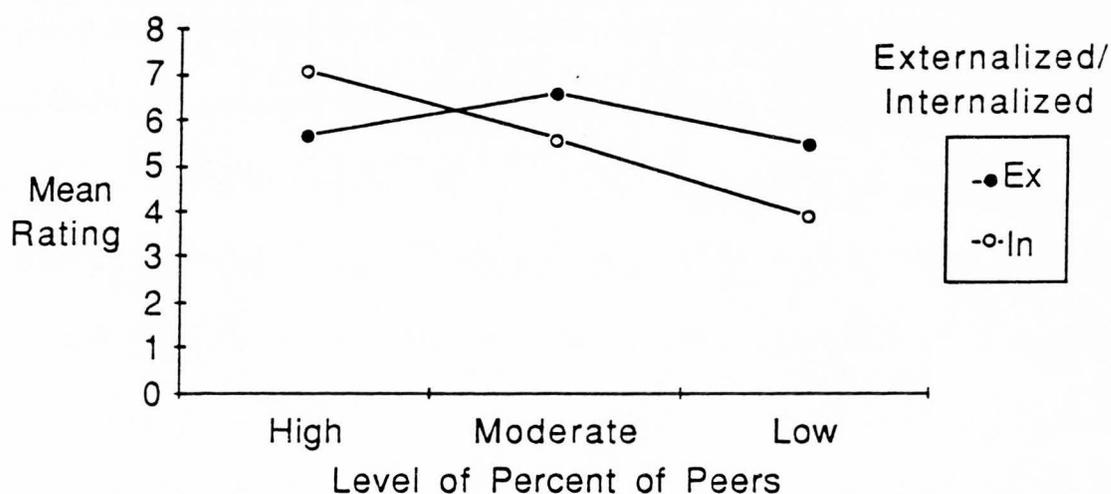
Figure 4 Externalized/internalized by level of percentage of peers.

There was no difference between the mean ratings for high and moderate, and between high and low levels of percentage of peers when the problem behaviors were externalized. There was a significant difference between moderate and low levels of percentage of peers;

as can be seen in Figure 4, the mean rating for moderate percentage of peers when the behavior was externalized was higher than the mean rating for high level of percentage of peers when the behavior was externalized. There were significant differences between the mean ratings for high, moderate, and low percentage of peers when the problem behaviors were internalized. There were significant differences between the means for externalized and internalized problem behaviors across the three levels of percentage of peers, with the mean for internalized higher when the level of percentage of peers was high and the means for externalized higher when the percentage of peers was at moderate and low levels. The means and Fisher's LSD comparisons for the percentage of peers ANOVA can be found in Appendix O.

The interaction between externalized/internalized and level of duration was not statistically significant. The mean ratings for the three levels of duration were similar across both externalized and internalized problem behaviors. These means and the Fisher's LSD comparisons can be found in Appendix M.

The interaction between externalized/internalized and level of generality was also not statistically significant. As in the duration by externalized/internalized interaction, the mean ratings for the three levels of generality were similar across

both externalized and internalized problem behaviors. These means and the Fisher's LSD comparisons can be found in Appendix N.

Hypothesis 5, that there is no interaction between the ratings on externalized and internalized problem behaviors and the ratings of seriousness on high, moderate, and low levels of the factor under consideration, was rejected in three cases and accepted in two cases.

Hypothesis 6: Interaction Between Externalized/ Internalized and Level of the Other Four Factors.

The interaction between externalized/internalized and level of the other four factors was statistically significant in all five ANOVAs.

The interaction between externalized/internalized and level of the other four factors in the severity ANOVA is shown in Figure 5.

There was no difference between the mean ratings for high and moderate levels of the other four factors when the problem behavior was externalized. There was a significant difference between the mean ratings when the other four factors were at moderate versus low levels and the behavior was externalized. There was a significant difference between the mean ratings for high and moderate levels of the other four factors when the problem

behavior was internalized. There was no difference between the mean ratings for moderate versus low levels of the other four factors when the problem behavior was internalized.
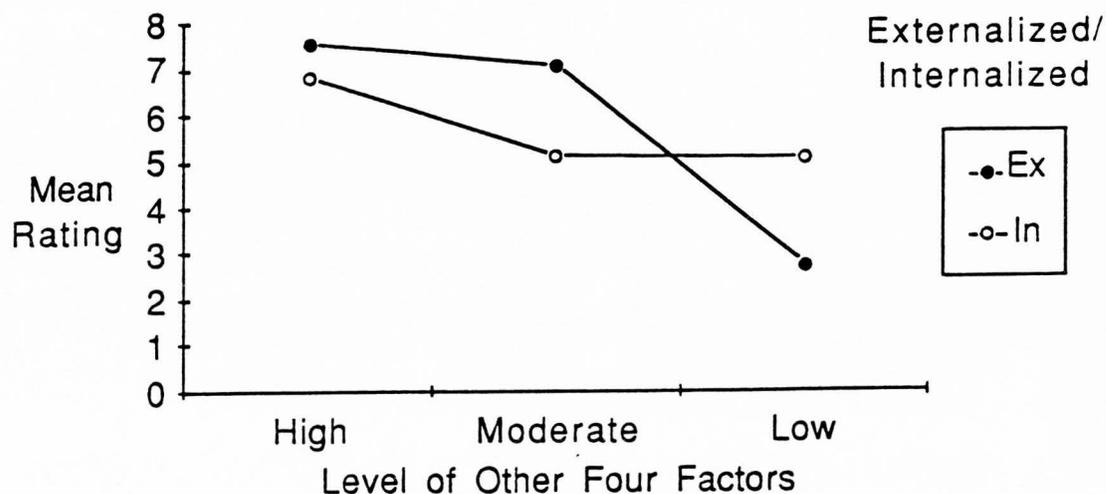
Figure 5 Externalized/internalized by level of other four factors - severity ANOVA.

There was no difference between the mean ratings for externalized and internalized problem behaviors when the level of the other four factors was high. There were significant differences between externalized and internalized when the level of the other four factors was moderate and low. When the level of the other four factors was moderate, the mean for externalized problem behaviors was higher than the mean for internalized problem behaviors. When the level of the other four factors was low, the mean for internalized problem behaviors was higher. The means and Fisher's LSDs can be

found in Appendix J.   The interaction between externalized/internalized and level of the other four factors in the frequency ANOVA is shown in Figure 6.

There were significant differences among the means for high, moderate, and low levels of the other four factors when the behavior was internalized and when it was externalized.   There were no differences between the means for externalized and internalized problem behaviors when the level of the other four factors was high and when it was low.   There was a significant difference
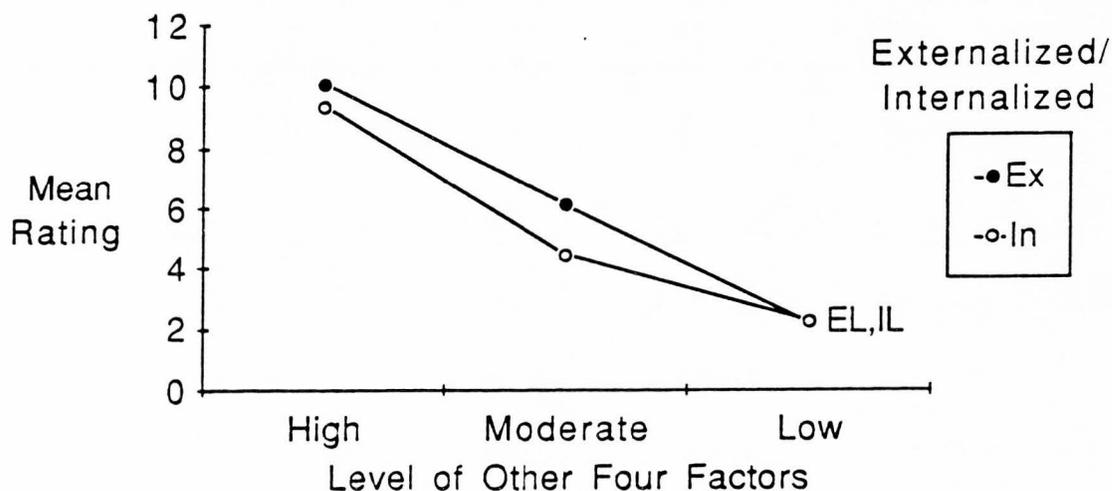
Figure 6 Externalized/internalized by level of the other four factors - frequency ANOVA.

between the means for externalized and internalized problem behaviors when the level of the other four factors was moderate.   The means and Fisher's LSD comparisons can be found in Appendix K.

The interaction between externalized/internalized and level of the other four factors in the duration ANOVA is shown in Figure 7.

There were significant differences among the means for high, moderate, and low levels of the other four factors when the behavior was internalized and when it was externalized. There were significant differences between the means for externalized and internalized problem
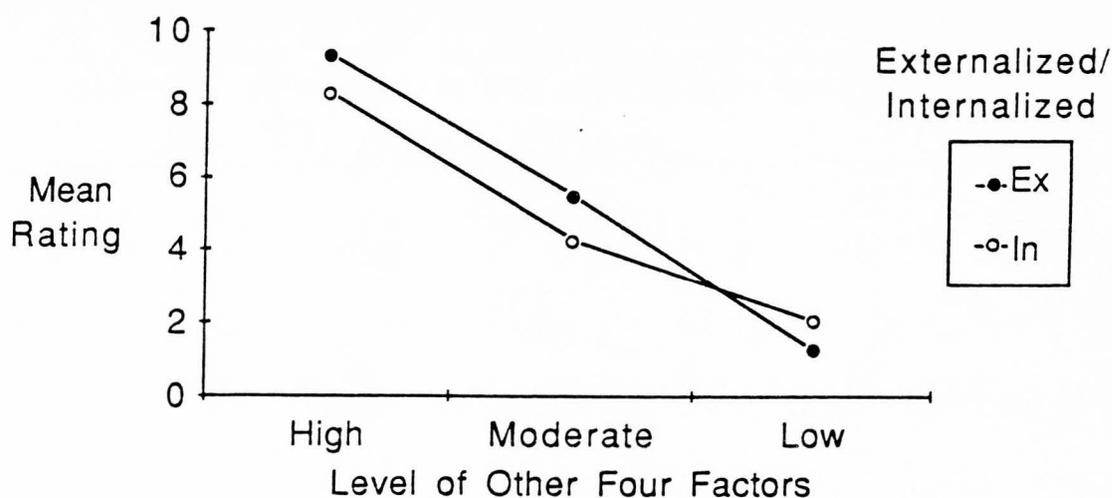
Figure 7 Externalized/internalized by level of other four factors - duration ANOVA.

behaviors when the other four factors were at high and moderate levels, with the mean ratings for externalized problem behaviors higher. There was no difference between the means for externalized and internalized problem behaviors when the other four factors were at a

low level. The means and Fisher's LSD comparisons can be found in Appendix L.

The interaction between externalized/internalized and level of the other four factors in the generality ANOVA is shown in Figure 8.

There were significant differences among the means for high, moderate, and low levels of the other four factors when the behavior was internalized and when it was
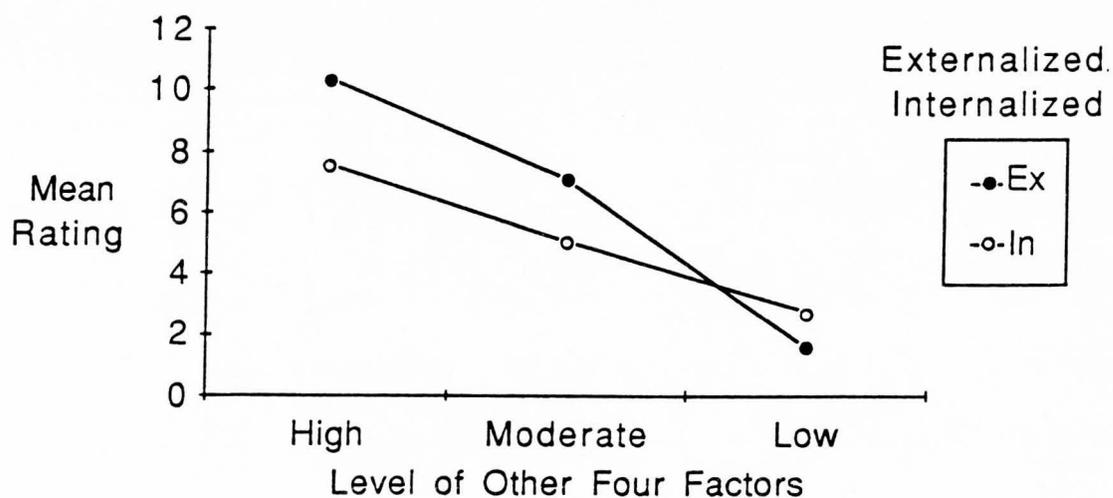
Figure 8 Externalized/internalized by level of other four factors - generality ANOVA.

externalized, with externalized problem behaviors rated higher at high and moderate levels and internalized problems rated higher at the low level. There were significant differences between the means for externalized and internalized problem behaviors when the other four factors were at high and moderate levels.

There was no difference between the mean ratings for externalized and internalized problem behaviors when the other four factors were at a low level. The means and Fisher's LSD comparisons can be found in Appendix M.

The interaction between externalized/internalized and level of the other four factors in the percentage of peers ANOVA is shown in Figure 9.
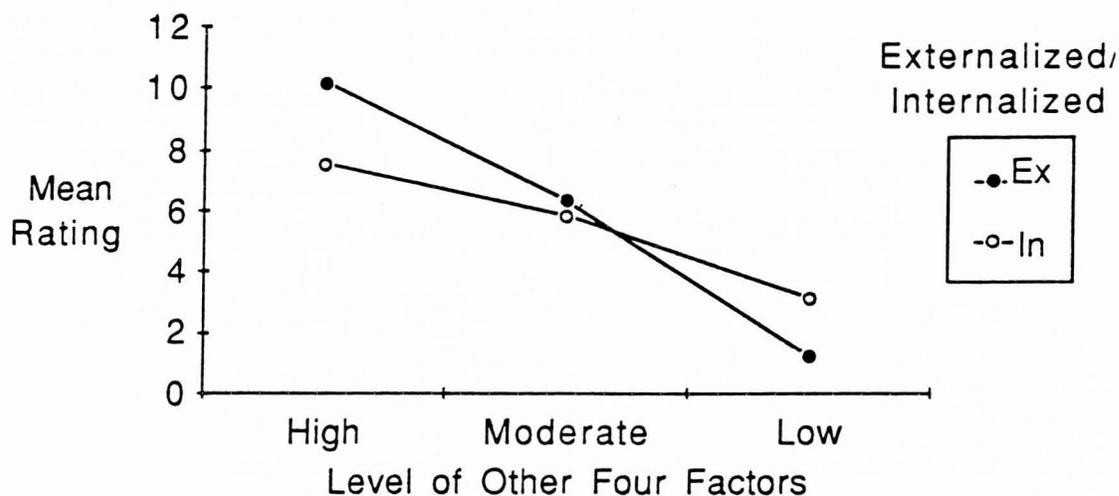
Figure 9 Externalized/internalized by level of other four factors - percentage of peers ANOVA.

There were significant differences among the means for high, moderate, and low levels of the other four factors when the behavior was internalized and when it was externalized. There were significant differences between the means for externalized and internalized problem

behavior when the levels of the other four factors were high and low. The mean rating for externalized problem behaviors was higher when the other four factors were high; the mean rating for internalized problem behaviors was higher when the other four factors were at a low level. There was no difference between the means for externalized and internalized problem behaviors when the level of the other four factors was moderate. The means and Fisher's LSDs can be found in Appendix O.

Hypothesis 6 states that there is no interaction between ratings of externalized and internalized problem behaviors and high, moderate, and low levels of the other four factors. This was not supported.

Hypothesis 7:
Interaction Between Level of the Factor Under Consideration and Level of the Other Four Factors.

The interaction between the levels of the factor under consideration and the levels of the other four factors was statistically significant in the severity, frequency, and duration ANOVAs. This interaction was not significant in the generality and percentage of peers ANOVAs.

The interaction between level of severity and level of the other four factors is illustrated in Figure 10.

The difference among the mean ratings when severity

was high versus moderate was significant when the level
of the other four factors were high, moderate, and low.
The difference between moderate and low severity was
significant when the level of the other four factors was
moderate but not when they were at high or low levels.
When severity was high, the mean rating for the other
four factors at a high level was significantly greater
than the ratings for the other four factors at moderate
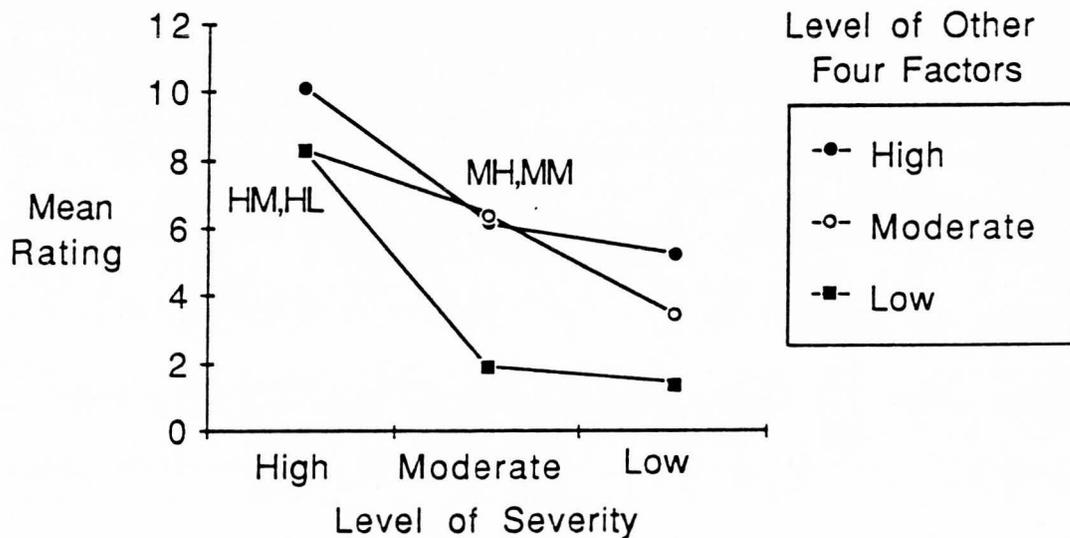
Figure 10. Interaction between level of severity and
level of the other four factors.

or low levels. When severity was high, there was no
difference in ratings between the other four factors at
moderate and low levels. There was no difference between
the mean rating for moderate severity when the other four
factors were at high and moderate levels, but it was
significantly lower when the other four factors were at a

low level. When the level of severity was low, there were significant differences between the mean ratings for problem behavior descriptions where the level of the other four factors was high versus moderate versus low. These means and the Fisher's LSD comparisons can be found in Appendix J.

The interaction between the level of frequency and the level of the other four factors is illustrated in Figure 11.
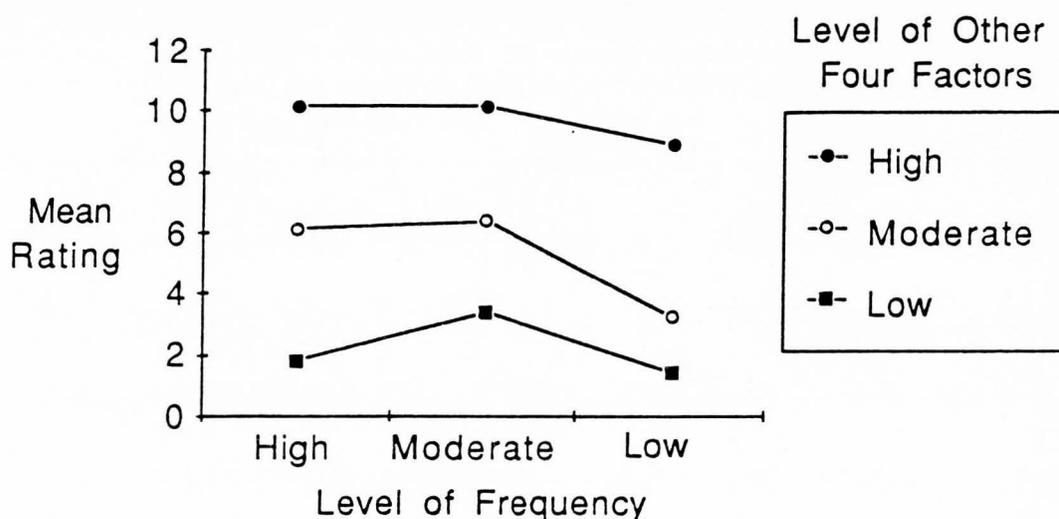
Figure 11. Interaction between level of frequency and level of the other four factors.

There were no significant differences between the mean ratings for high and moderate levels of frequency when the level of the other four factors was high and when it was moderate. When the level of the other four

factors was high, the mean ratings for problem behavior descriptions with moderate-level frequency was significantly higher than for those with low-level frequency. There was no difference between the mean rating for high frequency and low frequency when the level of the other four factors was low. There were significant differences between the mean ratings for problem behavior descriptions in which level of frequency was moderate versus low when the level of the other four factors was moderate and low. There were significant differences among mean ratings for high, moderate, and low levels of the other four factors across all three levels of frequency. The means and Fisher's LSD comparisons for this interaction can be found in Appendix K.

The interaction between level of duration and level of the other four factors is illustrated in Figure 12.

When the level of the other four factors was high, there was a significant difference between the mean ratings for high versus moderate duration but no difference between the mean ratings for moderate versus low duration. When the level of the other four factors was moderate, the mean rating was significantly higher when duration was at a moderate level than when it was at high or low levels. There was no difference between the mean ratings for high and low duration when the other

four factors were at a moderate level. There was no
difference among the mean ratings for high, moderate and
low duration when the level of the other four factors was
low. There were significant differences among the means
for high, moderate, and low levels of the other four
factors across all three levels of duration. These means
and Fisher's LSD comparisons can be found in Appendix L.
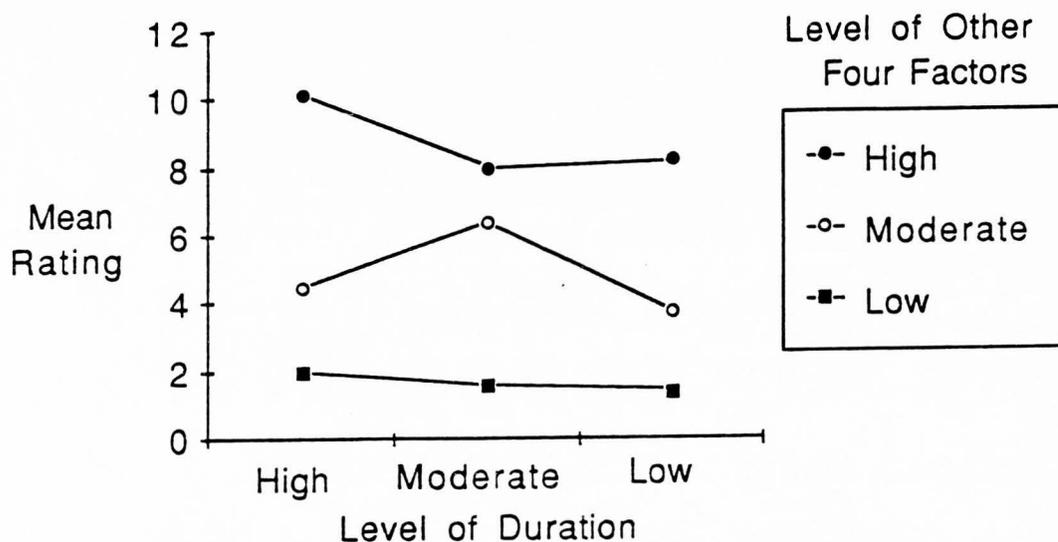
Figure 12. Interaction between level of duration and
level of the other four factors.

The interaction between level of generality and level
of the other four factors was not statistically
significant. The mean ratings for high, moderate, and
low generality varied in the same pattern across high,
moderate, and low levels of the other four factors.
These means and Fisher's LSD comparison can be found in
Appendix M.

The interaction between level of percentage of peers and level of the other four factors was not statistically significant. The mean ratings for high, moderate, and low percentage of peers varied in the same pattern across high, moderate, and low levels of the other four factors. These means and Fisher's LSD comparison can be found in Appendix N.

Hypothesis 7, which states that there is no interaction among the ratings of seriousness on problem behavior descriptions that present levels of the factor under consideration and the ratings of seriousness of problem behavior descriptions that present four behavioral factors at high, moderate, and low levels, is accepted in two cases and rejected in three.

Hypothesis 8: 3-Way Interaction Among
Externalized/Internalized, Level of the Factor
Under Consideration and Level of the Other Four Factors.

The three-way interactions were statistically significant in all five ANOVAs.

The three-way interaction among externalized/internalized, level of severity, and level of the other four factors is illustrated in Figure 13.

As can be seen on the top graph in Figure 13, when the problem behavior was externalized, there were significant differences in the mean ratings between high versus

Externalized

Internalized
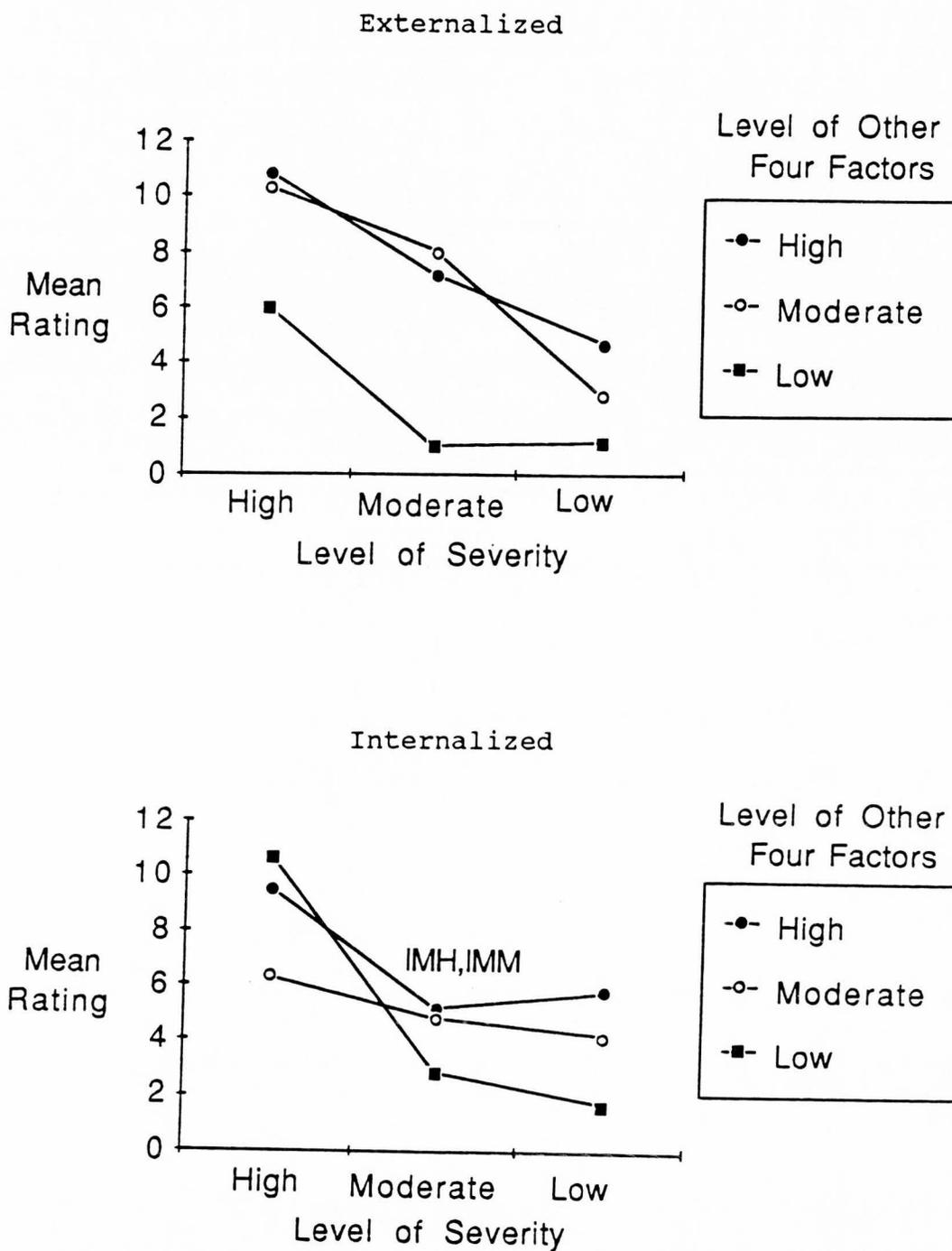
Figure 13. 3-way interaction - severity ANOVA.

moderate severity when the level of the other four factors was high, moderate, and low. Significant differences were also found between the mean ratings for moderate and low severity when the level of the other four factors was high and moderate but not when it was low. There was no difference among the mean ratings for high severity when the other four factors were at high versus moderate levels. The mean rating for high severity when the other four factors were low was significantly lower than when the other four factors were at moderate or high levels. Similarly, there was no difference among the mean ratings for moderate severity when the level of the other four factors was high versus moderate, but the mean rating for moderate severity was significantly lower when the other four factors were at a low level.

As is illustrated in the lower graph in Figure 13, when the problem behavior was internalized, there were significant differences between the mean ratings for high versus moderate severity when the level of the other four factors was high and low but not when it was at a moderate level. There were no differences among the mean ratings when the severity was moderate versus low across all three levels of the other four factors. There was no difference between the mean ratings for high and low levels of the other four factors when the level of

severity was high; the mean rating was significantly lower when the other four factors were at a moderate level. There was no difference between the mean ratings for high and moderate levels of the other four factors when severity was at a moderate and at a low level; the mean ratings were significantly lower when the level of the other four factors was low for both moderate and low severity.

Comparison of the mean ratings for internalized versus externalized problem behaviors is outlined in Table 12, which lists whether the Fisher's LSD comparison was statistically significant (S) or nonsignificant (NS).

Table 12.

Pairwise Comparisons for Externalized and Internalized Problem Pehaviors - Severity ANOVA.

	External/Internal
severity high, other 4 high	NS
severity high, other 4 moderate	S
severity high, other 4 low	S
severity moderate, other 4 high	S
severity moderate, other 4 moderate	S
severity moderate, other 4 low	NS
severity low, other 4 high	NS
severity low, other 4 moderate	NS
severity low, other 4 low	NS

There was no difference between the ratings for externalized and internalized problem behaviors when the level of severity was low across all three levels of the

other four factors. There was also no difference when the severity was high and the other four factors were high, nor when the level of severity was moderate and the other four factors were low. The mean for externalized problem behaviors was significantly higher when 1) severity was high and the other four factors were moderate, 2) severity was moderate and the other four factors were high, and 3) severity was moderate and the other four factors were moderate. The mean rating for internalized problem behavior was higher when severity was at a high level and the other four factors were at a low level.

The means and Fisher's LSD's for this interaction can be found in Appendix J.

The interaction among level of frequency, externalized/internalized, and level of the other four factors is illustrated in Figure 14.

As shown in the top graph of Figure 14, the mean ratings for externalized problem behavior, there was no difference in the mean ratings among high, moderate, and low frequency when the level of the other four factors was high. There was no difference between high and moderate levels of frequency when the level of the other four factors was moderate; when the level of frequency was low, however, the mean rating was significantly lower. Similarly, there was no difference between high

Externalized

Internalized
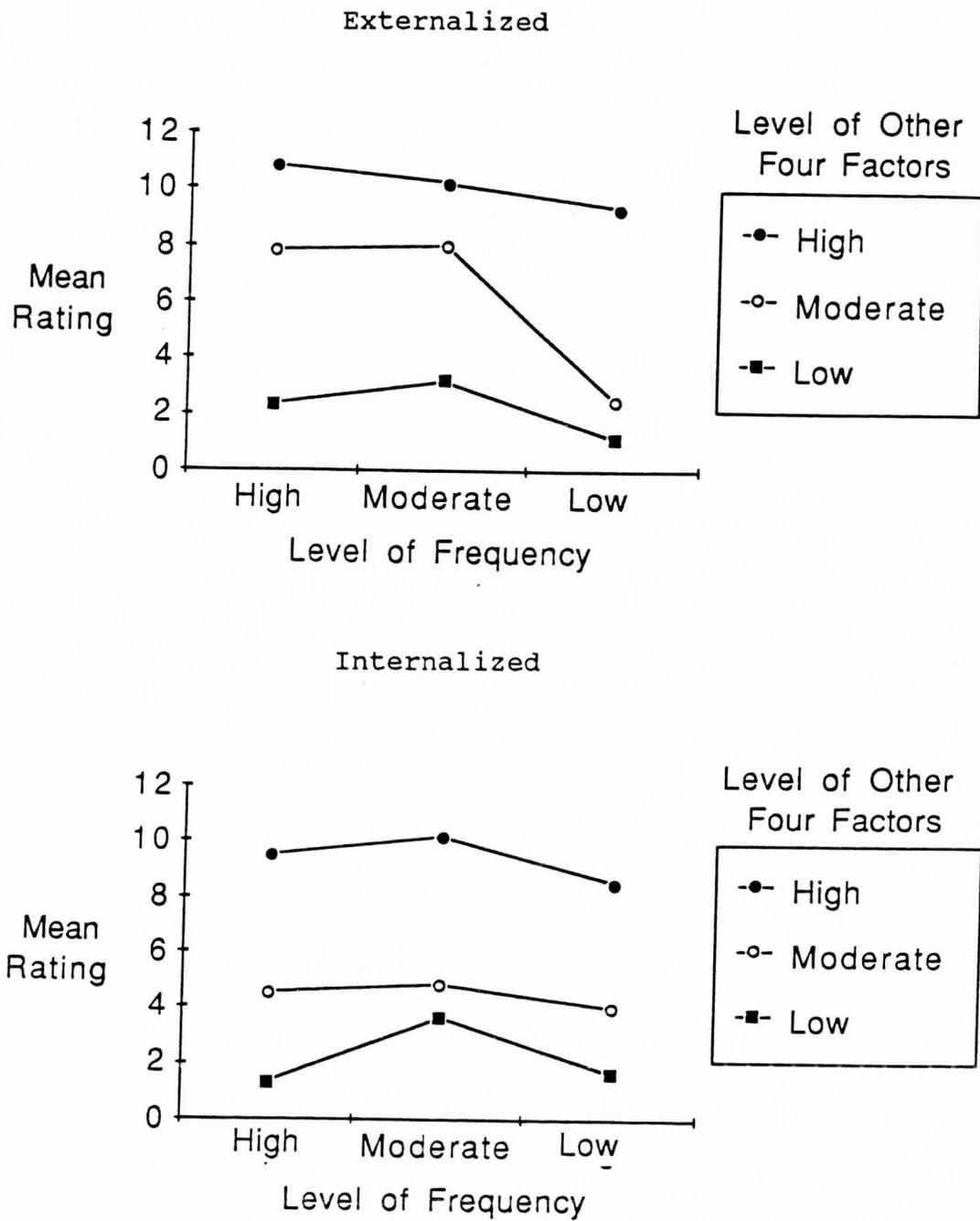
Figure 14.   3-way interaction - frequency ANOVA.

and moderate levels of frequency when the level of the other four factors was low, but the mean rating for low frequency problem behaviors was significantly lower. The difference between the mean ratings for high versus moderate versus low level of the other four factors was significant when the level of the other four factors was high and when it was moderate. There was a significant difference between the mean rating for high and moderate levels of the other four factors when the level of frequency was low; there was no difference between the mean ratings for moderate and low levels of the other four factors when the level of frequency was low.

As illustrated in the bottom graph of Figure 14, when the problem behavior was internalized, there were no differences among mean ratings for high, moderate, and low frequency problem behaviors when the level of the other four factors was high and when it was at a moderate level. There was no difference between the mean rating for high and low level frequency when the level of the other four factors was low, but the mean rating for moderate frequency was significantly higher. There were significant differences among mean ratings for high, moderate, and low levels of the other four factors when the level of frequency was high and when it was low. When the level of frequency was moderate, there was a significant difference between the mean rating for high

and moderate levels of the other four factors, but the mean ratings for moderate and low frequency were not significantly different.

Comparison of the mean ratings for internalized versus externalized problem behaviors is outlined in Table 13, which lists whether the Fisher's LSD comparison was statistically significant (S) or nonsignificant (NS).

Table 13.

Pairwise Comparisons for Externalized and Internalized Problem Behaviors - Frequency ANOVA.

	External/Internal
frequency high, other 4 high	NS
frequency high, other 4 moderate	S
frequency high, other 4 low	NS
frequency moderate, other 4 high	NS
frequency moderate, other 4 moderate	S
frequency moderate, other 4 low	NS
frequency low, other 4 high	NS
frequency low, other 4 moderate	NS
frequency low, other 4 low	NS

There was no difference between the mean ratings for externalized and internalized problem behaviors except that 1) when frequency was high and the other four factors were moderate, the mean for externalized was higher; and 2) when frequency was at a moderate level and the other four factors were at a moderate level, the mean for externalized was significantly higher.

These means and the Fisher's LSD comparison can be

found in Appendix K.

The three-way interaction among level of duration, level of the other four factors, and externalized/internalized is illustrated in Figure 15.

As is illustrated in the top graph in Figure 15, when the problem behavior was externalized there was a significant difference between the mean ratings for high and moderate duration when the other four factors were at a high level but no difference between the means for moderate- and low-level durations. There was no difference between the mean ratings for high and low duration when the other four factors were at a moderate level, but the mean for moderate duration was significantly higher than both of them. There was no difference between the mean ratings for high versus moderate versus low duration when the other four factors were at a low level. When the level of duration was high, there were significant differences among the mean ratings for high, moderate, and low levels of the other four factors. When the level of duration was moderate, there was no difference between the mean ratings for high and moderate levels of the other four factors, but the mean rating for low level of the other four factors was significantly lower. When the level of duration was low, there were significant differences between the mean ratings for high versus moderate versus low duration.
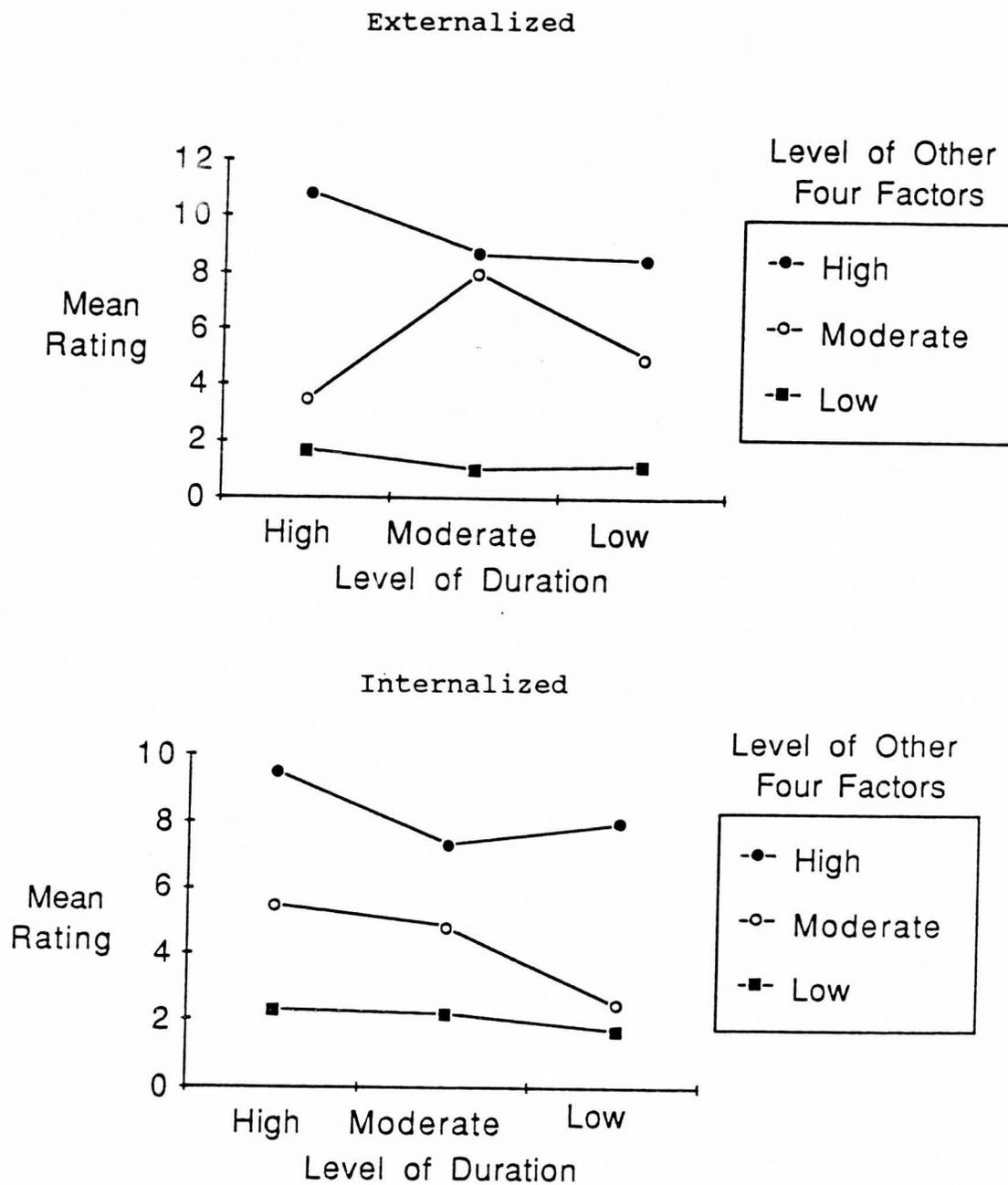
Figure 15. 3-way interaction - duration ANOVA.

As illustrated in the lower graph in Figure 15, when the problem behaviors were internalized there was no difference between the ratings for high and low levels of duration and between moderate and low levels of duration when the other four factors were at a high level, but the mean for moderate duration was significantly lower than the mean for high duration. When the level of the other four factors was moderate, there was no difference between the mean ratings for high and moderate duration, but the mean for low duration was significantly lower. When the level of the other four factors was low, there was no difference between the mean ratings for high versus moderate versus low levels of duration. When duration was high and moderate, there were significant differences between the mean ratings for high versus moderate versus low levels of the other four factors. When duration was at a low level, there was a significant difference between the mean ratings for high and moderate levels of the other four factors but no difference between the mean ratings for moderate and low levels of the other four factors.

Comparison of the mean ratings for internalized versus externalized problem behaviors is outlined in Table 14, which lists whether the Fisher's LSD comparison was statistically significant (S) or nonsignificant (NS).

There was no difference between the mean ratings for

externalized and internalized problem behaviors except 1) when duration was high and the other four factors were at a moderate level, the mean rating for the internalized problem behaviors was significantly higher; 2) when duration was at a moderate level and the other four factors were at a moderate level, the mean rating for the externalized problem behaviors was higher; and 3) when the level of duration was low and the other four factors

Table 14.

Pairwise Comparisons for Externalized and Internalized Problem Behaviors - Duration ANOVA.

	External/Internal
duration high, other 4 high	NS
duration high, other 4 moderate	S
duration high, other 4 low	NS
duration moderate, other 4 high	NS
duration moderate, other 4 moderate	S
duration moderate, other 4 low	NS
duration low, other 4 high	NS
duration low, other 4 moderate	S
duration low, other 4 low	NS

were at a moderate level, the mean rating for the externalized problem behaviors was significantly higher.

These means and Fisher's LSD comparison can be found in Appendix L.

The three-way interaction among level of generality, level of the other four factors and externalized/internalized is illustrated in Figure 16.

As is illustrated in the top graph in Figure 16, when the problem behavior was externalized there was no difference between the mean ratings for high versus moderate versus low generality when the level of the other four factors was at a high level and when it was at a low level. There was no significant difference between the mean ratings for high and moderate generality when the other four factors were at a moderate level but the mean for low generality was significantly lower. There were significant differences between the mean ratings for high versus moderate versus low levels of the other four factors across all three levels of the other four factors.

As illustrated in the bottom graph in Figure 16, when the problem behavior was internalized and when the other four factors were at a high level, there was a significant difference between the mean ratings for high and moderate generality and no difference between the mean ratings for moderate and low generality. There was no difference among the mean ratings for high, moderate, and low generality when the level of the other four factors was moderate. There was no difference between the mean ratings for moderate and low generality when the level of the other four factors was low. There was a significant difference between the mean ratings for high and moderate levels of the other four factors when
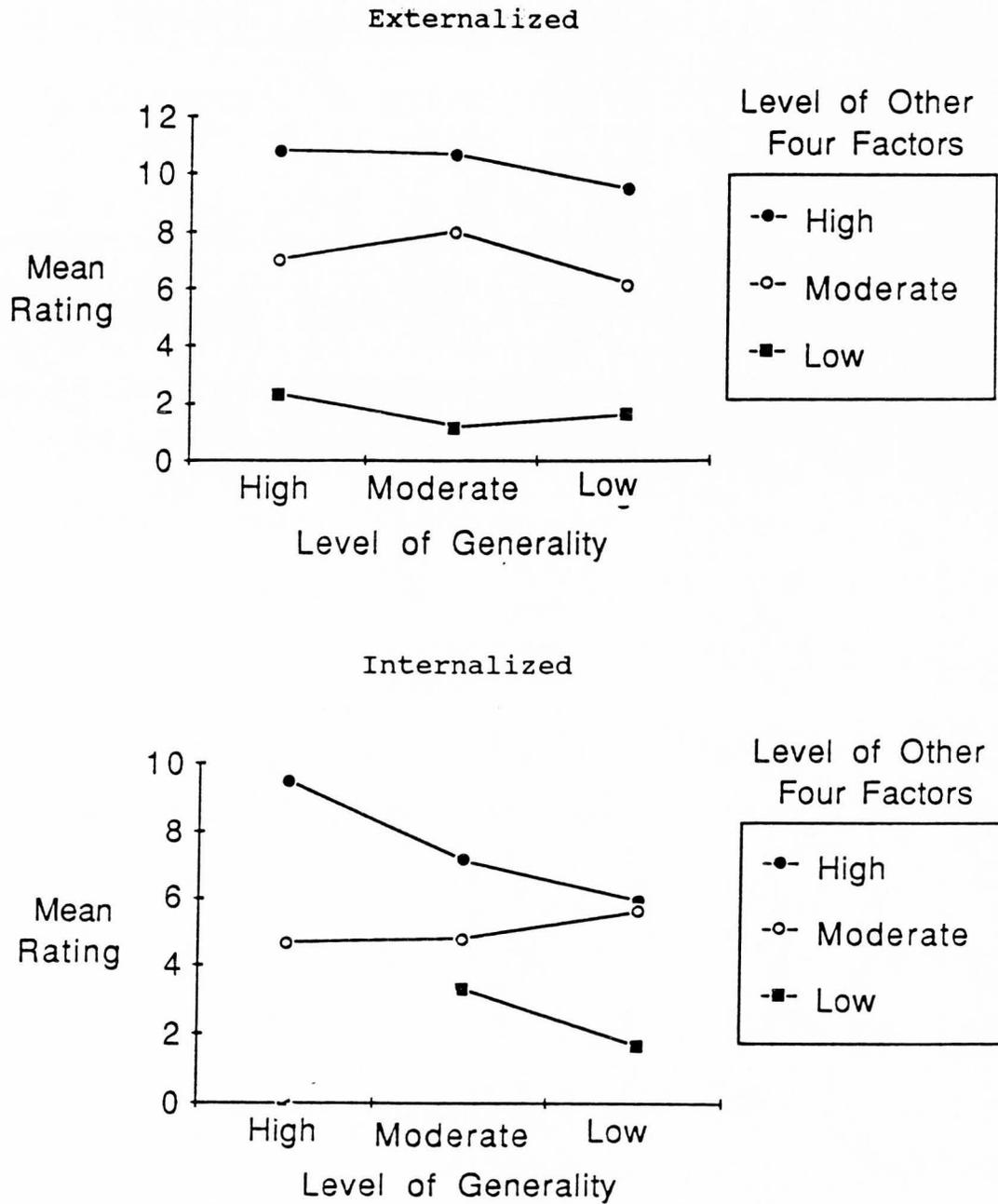
Figure 16. 3-way interaction - generality ANOVA.

generality was at high and moderate levels but not when generality was at a low level. There was no difference between the means for moderate versus low level of the other four factors when generality was at a moderate level but a significant difference when generality was at a low level. As mentioned earlier, there was no problem behavior description in which generality was high and the other four factors were at a low level.

Comparison of the mean ratings for internalized versus externalized problem behaviors is outlined in Table 15, which lists whether the Fisher's LSD comparison was statistically significant (S) or nonsignificant (NS).

Table 15.

Pairwise Comparisons for Externalized and Internalized Problem Behaviors - Generality ANOVA.

	External/Internal
generality high, other 4 high	NS
generality high, other 4 moderate	S
generality moderate, other 4 high	S
generality moderate, other 4 moderate	S
generality moderate, other 4 low	S
generality low, other 4 high	S
generality low, other 4 moderate	NS
generality low, other 4 low	NS

The means for externalized and internalized problem behaviors were significantly different 1) when generality was high and the other four factors were moderate, 2) when generality was moderate and the other four factors

were high, 3) when generality was moderate and the other
four factors were moderate, 4) when generality was
moderate and the other four factors were low, and 5) when
generality was low and the other four factors were high.

These means and Fisher's LSD comparisons can be found
in Appendix N.

The three-way interaction among level of percentage of
peers, level of the other four factors, and
externalized/internalized is illustrated in Figure 17.

As illustrated in the top graph in Figure 17, when the
problem behavior was externalized there was no difference
between the mean ratings for high versus moderate versus
low percentage of peers when the other four factors were
at a high level. There was no difference between the
mean ratings for high and low levels of percentage of
peers but a significantly higher mean rating for moderate
percentage of peers when the level of the other four
factors was moderate. There was no difference between
the mean ratings for high versus moderate versus low
percentage of peers when the other four factors were at a
low level. There were significant differences between
the mean ratings for high versus moderate versus low
levels of the other four factors across all three levels
of percent of peers.

Externalized
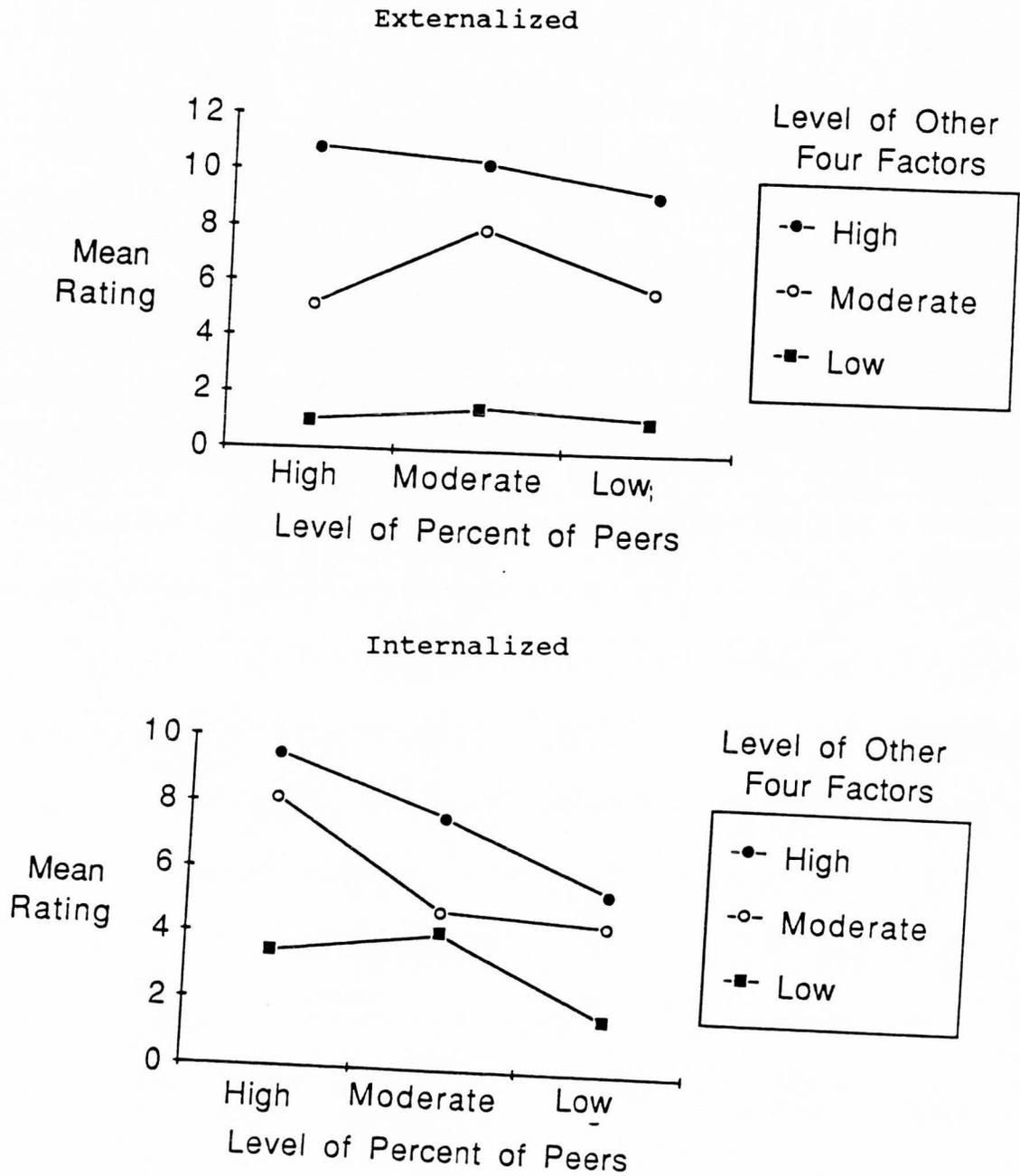
Internalized

<u>Figure 17.</u>   3-way interaction - percent of peers ANOVA.

As illustrated in the bottom graph in Figure 17, when the problem behavior was internalized there were significant differences between the mean ratings for high versus moderate versus low levels percentage of peers when the other four factors were at a high level. There was a significant difference between the mean ratings for high and moderate levels of percentage of peers but no difference between the mean ratings for moderate and low percentage of peers when the other four factors were at a moderate level. There was no difference between the mean rating for high versus moderate percentage of peers, but a significant difference between moderate and low percentage of peers when the other four factors were at a low level. When percentage of peers was high, there was no significant difference between the mean ratings for high and moderate levels of the other four factors but a significant difference between the mean ratings for moderate and low levels of the other four factors. When percentage of peers was at a moderate level, there was a significant difference between the mean ratings for high versus moderate levels of the other four factors but no difference between the mean ratings for moderate and low levels of the other four factors. When percentage of peers was at a low level, there was no difference between the mean ratings for high and moderate levels of the other four factors, but the mean rating for low level of

the other four factors was significantly lower. Comparison of the mean ratings for internalized versus externalized problem behaviors is outlined in Table 16, which lists whether the Fisher's LSD comparison was statistically significant (S) or nonsignificant (NS). There were significant differences between the mean ratings for externalized and internalized problem behaviors when 1) percentage of peers was high and the

Table 16.

Pairwise Comparisons for Externalized and Internalized Problem Behaviors - Percent of Peers ANOVA.

	External/Internal
% of peers high, other 4 high	NS
% of peers high, other 4 moderate	S
% of peers high, other 4 low	S
% of peers moderate, other 4 high	S
% of peers moderate, other 4 moderate	S
% of peers moderate, other 4 low	S
% of peers low, other 4 high	S
% of peers low, other 4 moderate	NS
% of peers low, other 4 low	NS

level of the other four factors was both moderate and when it was low, the mean rating for internalized problem behaviors was higher; 2) percentage of peers was moderate and the other four factors were at a high or at a moderate level, the mean rating for the externalized problem behavior was higher; 3) when percentage of peers was at a moderate level and the other four factors were

at a low level, the mean rating for the internalized problem behavior was higher; and 4) when percentage of peers was at a low level and the other four factors were at a high level, the mean rating for externalized problem behaviors was significantly higher.

These means and the Fisher's LSD comparisons can be found in Appendix N.

Hypothesis 8, which states that there is no interaction among ratings on externalized and internalized problem behaviors, ratings on level of the factor under consideration, and the level of the other four factors, was not supported.

Association Among the Ratings, the Main Effects
and the Interactions in the ANOVAs

R squared was computed for each of the ANOVAs to obtain an index of the amount of variability accounted for in the analyses. R squared is the proportion of the total variance in the ratings accounted for by the variables and interactions in the ANOVAs. The R squares are listed in Table 17 below.

As the table shows, a high percentage of the variability in the ratings on the problem behavior descriptions was accounted for in these analyses.

Table 17.

R Squared for Each ANOVA

ANOVA	R-Squared
Severity	.833
Frequency	.897
Duration	.855
Generality	.858
Percent of Peers	.862

The correlation ratio, Eta squared, was calculated for each main effect and interaction. This was done in order to obtain an index describing the relationship between the ratings on the problem behavior descriptions and each of the variables and interactions in the five ANOVAs. These Eta squares are presented in Table 18 below.

Table 18

Eta Square for Each Variable and Interaction in the Five Three-Way ANOVAs.

	Severity	Frequency	Duration	Generality	Percent of Peers
Subjects	.057	.0365	.064	.061	.047
E/I	.00025	.012	.064	.019	.0034
FUC	.4432	.0613	.017	.0403	.0456
O4F	.151	.724	.674	.647	.625
E/I x FUC	.015	.0165	.009	.0005	.034
E/I x O4F	.065	.0007	.016	.056	.070
FUC x O4F	.048	.018	.042	.018	.011
E/I x FUC X O4F	.054	.020	.027	.017	.026

As shown in Table 18, the greatest proportion of the variability in the ratings in the severity ANOVA was attributed to the three levels of severity (44.32%). In the other ANOVAs, the greater proportion of the variability in the scores was associated with the level of the other four factors. In the frequency ANOVA, the proportion was 72.4%; in the duration ANOVA, the proportion was 67.4%; in the generality ANOVA, the proportion was 64.7%; and in the percentage of peers ANOVA, the proportion was 62.5%. Thus it appears that the greater proportion of the variability in the ratings is associated with the level of severity in the problem behavior descriptions.

Comments on Problem Behavior Descriptions

In the instructions, the subjects were asked to note whether there were any information not included in the behavior problem descriptions they felt was important in determining the seriousness of the problem. These comments were tallied and are presented in Table 19. The subjects made relatively few comments regarding important missing information in the problem behavior descriptions. Subject 6 made the most with 22 comments and Subject 4 the least with none.

Table 19

Information Not Included in

Problem Behavior Descriptions

	S1	S2	S3	S4	S5	S6
Adequacy of classroom management procedures	2	10	4			9
Family/home conditions	2	3				
Prior interventions		1	2		4	
Appropriateness of educational placement/ curriculum		3				6
Drug use		1				
Vision/Hearing/Motor Problems		1				3
Analysis of environment/ situational variables			3		2	
Age/grade of student			5			1
Mental Illness vs Malingering?			2			
Language problems?						2
Recent head injury?						1

Interview Data

In rating the problem behavior descriptions, one or more of the subjects provided ratings that were three or more points different from the ratings of the other subjects on 42 of the 65 problem behavior descriptions.

For each of the 42 descriptions, the person giving the disparate rating as well as two of the subjects giving typical ratings was interviewed regarding that description.

The first question in the semi-structured interview was:

Of the information provided in this case, which pieces of information were most important in helping you determine your rating on this case?

Rank order: _____ frequency
_____ severity
_____ duration
_____ generality
_____ percent of peers

Kendall's W coefficient of concordance was used to determine, for each problem behavior description, whether the subjects ranked the factors in a similar manner. The Kendall's W's were statistically significant in only 4 of the 42 cases. Thus, there was little association among the subjects with respect to the way they ranked the importance of the five behavioral factors. The results of these analyses can be found in Appendix O.

In addition, Kendall's W coefficient of concordance was run on the rankings of all of the cases ranked by each subject. This was done to see if there were within-subject consistencies in the rank ordering of the five factors. The results of these analyses can be found in Table 20. For five of the six subjects, the correlation

among rankings of the five factors was small (range .08 to .28) yet statistically significant p<.05. Thus, there was a small tendency for all but one subject to rank the five behavioral factors in the same way from one problem behavior to another. The mean ranks for each factor for each subject can be found in Appendix P.

Table 20

Kendall's W for Subject's Rankings

on Interview Question 1.

	#PBDs	W	Chi-Square	Signif.
Subject 1	28	.0890	9.9714	.0409
Subject 2	23	.0465	4.2783	.3697
Subject 3	27	.2869	30.9885	<.0001
Subject 4	26	.1139	11.8450	.0185
Subject 5	25	.1123	11.2349	.0240
Subject 6	24	.1205	11.5667	.0209

To determine whether there were a higher degree of association among the rankings provided by subjects who gave typical ratings on the problem behavior descriptions than between the typical subjects and subjects giving disparate ratings, Spearman's rho was calculated between the rankings provided by each of the subjects interviewed for each problem behavior description. Spearman's rho is a measure of the association between two sets of ranks. A total of 201 Spearman's rhos was computed. Of these, 20 were statistically significant at the .05 level. Thus,

there appears to be little or no tendency for subjects to rank the behavioral factors similarly.

Of the 20 statistically significant rho's 1) 8 were among the rankings provided by subjects who gave typical ratings on the problem behavior description, 2) 8 were between a subject who had provided a typical rating and a subject who had provided a disparate rating on the problem behavior description, and 3) 4 were among subjects who had provided disparate ratings on the problem behavior descriptions. Thus, among the few cases where significant rho's were found, there appears to be no greater tendency for typical subjects to rank the five behavioral factors in the same order than for typical and disparate subjects to rank them in the same order.

A frequency count was taken on the number of times a "typical by typical" rho was highest, a "typical by disparate" rho was highest, and a "disparate by disparate" rho was highest for each problem behavior description. These frequencies are shown in Table 21. There were two-way ties on two problem behavior descriptions and a three-way tie on one problem behavior description. As the data show, there was no higher association among the rankings provided by persons giving similar ratings. Their approaches to ranking the five behavioral factors during the interview were not similar. The Spearman's rhos for each problem behavior description

can be found in Appendix Q.

Table 21.

Frequency of Highest Spearman's Rhos for Each Problem Behavior Description.

typical x typical	typical x disparate	disparate x disparate
17	24	5

Frequency counts were also taken of the comments made during the interviews regarding each of the problem behavior descriptions to determine if typical and disparate subjects reported different approaches to evaluating the seriousness of the problem behaviors. There were no consistencies among typical raters or any consistent disparities between typical and disparate raters. Comments and frequency counts can be found in Appendix Q.

A frequency count was taken on pieces of information not presented in the problem behavior descriptions that the subjects asked about during the interviews. These were tallied to look for information the subjects found missing and would have liked to have known more about when rating the problem behavior descriptions. The data are presented in Table 22.

As the table shows, few comments were made, and even fewer were made more than three times. Subjects most often requested information regarding classroom management procedures, the students' ages and the environment the behavior was occurring in. The rest of the comments can be found in Appendix Q.

Table 22

Questions Requesting Additional Information

Asked During the Follow-Up Interviews

Comment	Frequency
Adequacy of classroom management procedures	10
Family/home conditions	5
Prior interventions	7
Appropriateness of educational placement/ curriculum	7
Drug use	3
Vision/Hearing/Motor Problems	2
Analysis of environment/ situational variables	9
Age/grade of student	17
Mental Illness vs Malingering/manipulative	6
Language problems?	2
Recent head injury?	2
Social/emotional problems?	3
Duration of episodes?	1
Pattern of occurrance	1
Learning disability?	3
Attention Deficit Disorder?	1
Police/legal involvement?	2

CHAPTER V

DISCUSSION

Study Overview

The present study sought to determine whether or not field-based decision makers would be in agreement on their ratings of seriousness of a set of descriptions of problem behaviors. The study also sought to validate the rules for determining the seriousness of problem behaviors contained in the computer expert system Class.BD. To these ends, problem behavior descriptions were developed, each of which contained information regarding five factors of problem behavior. Each of these factors was represented within the problem behavior descriptions at high, moderate, and low values. Correlational techniques and analysis of variance were used to study agreement among subjects and between subjects and the problem behavior section of Class.BD. Analysis of variance was used to study the impact of the presence of the five behavioral factors at three levels on the ratings the subjects provided. The results of the analyses of variance were used to describe how the subjects used the levels of the factors in determining their ratings of seriousness of the problem behaviors.

Limitations

In the present study the subjects were presented with written descriptions of problem behaviors. Each of the written descriptions contained a specified set of information, which included the nature of the behavior (severity), its frequency, duration, and generality, and the percentage of peers who engaged in the same behavior. Whether or not the subjects would actually attend to all five factors when evaluating real students in a school setting was not determined. All that can be said is when these factors were present and the subjects' attentions were brought to them, they had an impact on how they rated the seriousness of the problem behaviors.

The subjects in the present study were not randomly selected, and the generalizability of the results to other field-based decision makers is therefore open to question. The purpose of the present research was to study the responses of the best field-based decision makers possible in an attempt to validate the Class.BD expert system. The responses of the subjects may not be representative of those of other classification-decision makers within the state of Utah or other states.

Analysis of Variance

Five randomized block design three-way ANOVAs were conducted, one for each of the five factors under

consideration. These analyses were blocked on the subject variable to control for variability associated with subject differences. This allowed for greater precision in comparing the means for the other variables in the ANOVAs.

To determine whether the assumption of homoscedasticity was met, residual plots were generated for each ANOVA. There were no patterns in the residual plots that would indicate problems with the assumption of homoscedasticity.

To determine whether the assumption of normality had been met, normal probability plots were generated for each of the five ANOVAs. The normal probability plots produced straight lines, indicating no serious deviations from normality.

Findings

Agreement Between Subjects/Validation of Class.BD

The findings of the present study indicate that there is agreement among ratings of seriousness on problem behavior descriptions provided by the subjects, as well as agreement between subjects and the problem behavior evaluation section of the Class.BD computer program. All Kendall's W's among the ratings of the subjects and between their ratings and Class.BD were high and

statistically significant. The Pearson's r correlation coefficients between pairs of subjects and the two versions of Class.BD were also quite high (range .6461 to .8717) and statistically significant.

Although the analysis of variance for differences between subjects yielded a statistically significant F, the Fisher's LSD multiple comparisons showed no significant differences between individual subjects' mean ratings. The mean rating for Subject 1 was more than one point higher than the mean ratings for the other five subjects. The reasons Subject 1 gave higher ratings than the other subjects were not empirically verified, but there are some possible explanations. First, Subject 1 was the only participant from a rural school district, where perhaps the staff members are less tolerant of deviant behavior. Second, Subject 1 was the only subject who worked exclusively in a high school setting. As was often noted by the subjects during the follow-up interviews, the seriousness of a problem behavior often depends on the age of the student. Behaviors that might seem quite normal for an early-elementary-school-aged child might be considered quite deviant in a teenage student. Also, she was the only subject who was a classroom teacher. These speculations, that school personnel in rural areas consider the degree of seriousness of a problem behavior differently than urban

school personnel and that high school and elementary school personnel consider the seriousness of a problem behavior differently, remain to be empirically tested.

In summary, both correlational analyses and the ANOVA support the same conclusion. Subjects presented with problem behavior descriptions show high agreement regarding their ratings of seriousness. Further, the Kendall's W analyses indicate that Class.BD's decisions regarding seriousness of problem behaviors are highly similar to those of the subjects. Thus, the problem behavior section of the Class.BD expert system was validated.

ANOVA Results/Implications for Knowledge Engineering

Differences in ratings of externalized versus internalized problem behavior. The results of the present study generally support Hypothesis 2, that there are no differences in the subjects' ratings of externalized versus internalized problem behaviors. The main effects for the externalized/internalized variable in the severity, duration, and percentage of peers ANOVAs were not statistically significant. While those effects were statistically significant in the frequency and generality ANOVAs, comparison of the means shows that the mean rating for externalized problem behaviors was only .8 points higher in the frequency ANOVA and 1.2 points

higher in the generality ANOVA. Thus, although the overall trend was for subjects not to discriminate between externalized and internalized problem behaviors, there was a slight tendency for them to respond differently when frequency and duration varied across different levels.

Walker, Reavis, Rhode, and Jenson (1985) suggested that students who internalize their problems are less likely to be referred for evaluation by classroom teachers. Presumably this is because their problems are less likely to result in classroom disruption, and teachers are less likely to notice that the student has problems. The findings of the present study suggest that, at least for the six subjects who participated in the study, there are few differences in perception of the seriousness of internalized and externalized problem behaviors. Once students with internalized problem behaviors are referred to these individuals for evaluation, their problems are, for the most part, considered equally as serious as those of their peers who exhibit externalized problem behaviors.

Differences among levels of the factors under consideration. The main effect for the factor under consideration was statistically significant in all five ANOVAs.

For the severity factor, subjects discriminated

among high, moderate, and low levels of severity. The mean ratings for each level were significantly different from each other.

For the frequency factor, the subjects' mean ratings for high versus moderate levels of frequency were not significantly different; the subjects did not discriminate between high (several times per day) and moderate (several times per week) levels of frequency in making their ratings. The subjects did discriminate between moderate and low (once per month or less) levels of frequency, providing a significantly lower mean rating for low-frequency problem behaviors. Thus, the subjects in this study considered problem behaviors occurring several times per week to be as serious as those occurring several times per day. When the problem behavior occurred once a month or less, they considered the problem behavior to be significantly less serious.

For the duration factor, the mean ratings for high and moderate duration were not significantly different. In providing their ratings, the subjects did not discriminate between high (more than six months) and moderate (one to six months) levels of duration. The mean rating for low duration (less than one month) was significantly lower than the means for high and moderate duration. Thus, the subjects considered problem behaviors with high and moderate duration to be equally

serious.   Problem behaviors having a low duration were rated as significantly less serious than those with high and moderate levels of duration.

For the generality factor, the subjects' mean ratings for behavior problems with high (67-100% of school environments) and moderate (34-66%) generality were not significantly different.   The mean rating for problem behavior descriptions with low generality (1-33% of school environments) was significantly lower.   Thus, the subjects gave equally high ratings for problem behaviors that occurred in 34% to 100% of the students' school environments.   Their ratings were significantly lower when the problem behaviors occurred in only one or a few settings.   According to the subjects' responses in the follow-up interviews, when problem behaviors occurred in only a small proportion of school settings, they became suspicious about the adequacy of the behavior management plans in those settings or wondered what situational variables were setting the occasion for the problem behavior to occur.

For the percentage of peers factor, the mean ratings for high (0-9%) and moderate (10-19%) levels of percentage of peers were not significantly different. When the percentage of peers engaging in the same behavior was at a low level (in this instance a higher actual value, 20% or more of the student's peers), the

subjects' ratings on the problem behavior descriptions were significantly lower. Thus, the subjects considered problem behaviors where 0 to 19% of the student's peers also engaged in the same behavior to be equally as serious. If 20% or more of the student's peers engaged in the same behavior, they rated the behavior problem as significantly less serious. According to their responses in the follow-up interviews, if a high percentage of the students engaged in a particular problem behavior, they became suspicious regarding the adequacy of the behavior management system or wondered what environmental factors might be setting the occasion for students to behave in such a manner.

In summary, the results indicate that subjects discriminated among three levels of severity in rating the seriousness of problem behavior descriptions. In contrast, they discriminated only two levels when focused on frequency, duration, generality, and percentage of peers. This suggests the possibility of a threshold that once passed, adds to concern regarding the seriousness of problem behaviors.

Differences between levels of the other four factors. The main effect for level of the other four factors was statistically significant in all five ANOVAs. The Fisher's LSD comparisons revealed significant differences

between ratings for problem behavior descriptions when the other four factors were at high versus moderate versus low levels.

Because all possible combinations of the levels of the five factors were not presented to the raters (it was not feasible to do so), it was impossible to determine which, if any, of the particular factors was responsible for the difference in the mean ratings. What can be concluded is that the subjects discriminated between high, moderate and low levels of the other four factors when making their ratings. Regardless of the level of the factor under consideration or which of the factors was under consideration, the subjects' ratings were significantly higher when the level of the other four factors was high, significantly lower when the other four factors were at a low level; and when the other four factors were at a moderate level the mean rating fell in between. It can be concluded that the subjects discriminated differences in levels of the other four factors in making decisions about the seriousness of the problem behaviors.

In comparing these results to those of the main effects for the factor under consideration, it appears that severity (one of the other four factors in the frequency, duration, generality, and percentage of peers ANOVAs) might be responsible for the significantly different ratings between high and moderate levels of the

other four factors. Further evidence for this conclusion was provided by the Eta-squared analysis, which indicates that the largest proportion of the variability in the ratings is associated with the levels of severity in the severity ANOVA, and with the level of the other four factors in the other four ANOVAs.

Interactions. Twenty-one of the 25 interaction effects tested were found to be statistically significant. These interactions were described in the preceding Results section. Taken together, they indicate that the subjects responded to the problem behavior descriptions in complex ways when making decisions regarding the seriousness of the problem behavior descriptions. However, the Eta squared data, which are presented in Table 18, indicate that these interactions accounted for a very small proportion of the variability in the ratings. In the severity ANOVA, over half of the variability in the ratings accounted for in the model was associated with the level of severity. In the frequency, duration, generality and percentage of peers ANOVAs, 72% to 81% of the variance accounted for in the models was associated with the level of the other four factors. Thus, it appears that the level of severity presented in the problem behavior descriptions accounted for the majority of differences among the ratings on the problem behavior descriptions.

As the subjects frequently mentioned during the follow-up interviews, low generality and 20% or more peers engaging in the same behavior were important cues to the subjects indicating a possible problem in the setting, such as poor behavior management in the classroom. When generality and percentage of peers were at low levels (a high percentage of peers=low-level percentage of peers in terms of Class.BD weightings), they gave low ratings on the problem behavior descriptions fairly consistently. The levels of these two factors, then, may be more important in determining the rating than the levels of frequency and duration.

Implications for knowledge engineering. The present study extends the application of the ANOVA method of knowledge elicitation (Triggs, 1988) to knowledge engineering in clinical decision making. This method has been used in previous research (Triggs, 1988) on materiality judgments of auditors, rain forecasting, and nurses judgments of when to call for the doctor. In these studies, specific cues and combinations of specific cues were used to form specific judgments.

In the present study, levels of factors rather than specific factors (pieces of information or cues) were used. In addition, the severity factor, which was not easily quantifiable, was used, representing a departure

from typical knowledge engineering situations. The results indicate that the ANOVA method can be useful in determining the importance of factors that are not mutually exclusive (do not occur independently of one another), enabling the researcher to determine the relative importance of the level of each factor in the decision-making process.

While the correlations between the subjects' ratings and the problem behavior section of Class.BD were quite high and statistically significant, it appears that the rules programmed into Class.BD are quite different from those used by the subjects. While there was support for three levels of severity, there was not support for three levels of frequency, duration, generality, and percentage of peers. Rather, the results suggest that there are two levels of these four factors that are useful to decision makers in determining the seriousness of a problem behavior.

As mentioned earlier, the variance model used in developing Class.BD involves assigning weights to each level of each factor, then the weightings are combined in an additive model to get an index of certainty regarding whether the student qualifies for special education. In the first version of Class.BD all of the weightings were positive. With the addition of a factor into the knowledge base, more of the remaining variance was

subsumed. The results of this study suggest that positive weights might not always be appropriate for determining the seriousness of a problem behavior. Particularly in situations where the percentage of peers engaging in the same behavior is high or the generality of the problem behavior is low, it appears to be more appropriate to use negative weightings. This may also be true for situations in which severity, frequency, and duration are extremely low. When negative weightings are used, the certainty factor is reduced by a proportion of the variance already subsumed.

Interview Data/Implications for Knowledge Engineering.

The Kendall's W's and Spearman's rhos among the rankings provided by the subjects on the problem behavior descriptions showed little consistency across subjects in the way they ranked the five factors. Each of the subjects tended to rank order the factors differently. However, each of the subjects exhibited some internal consistency in ranking the factors, based on statistically significant correlations among the subject's rankings on the problem behavior descriptions about which they were interviewed.

One possible reason for the lack of consistency across subjects is that they appeared to approach the task in different ways. For example, some subjects rank-ordered

the factors in terms of which factor made the problem most serious. The second factor was the one that had the second most impact on making the problem serious, and so forth. Others rank ordered the factors in terms of how much impact each piece of information had on their rating, without necessarily making it most serious. For example, percentage of peers might be ranked number one because it made the subject rate the problem behavior as much less serious and thus had more impact on their decision than the other four factors. Future researchers using such an interview format might find more consistency among subject responses if the directions to the subjects on how to approach the ranking task were more explicit.

When noting information they found lacking in the problem behavior descriptions, the subjects did not ask for additonal specifics regarding the behaviors. Rather, they asked for contextual information such as the adequacy of the classroom behavior management program, effectiveness and nature of prior interventions, other child characteristics, and so forth. On one occassion, one of the subjects wondered about the duration of each episode of a problem behavior. This was the only time any of the subjects asked about a behavioral factor not included in the Class.BD model.

These findings provide support for the concerns

expressed earlier in this paper on the adequacy of the interview method for knowledge engineering. They also support concerns over the practice of interviewing only one expert. The subjects in this study approached the task of ranking the five behavioral factors differently and gave very different responses to the open-ended interview questions.

In summary, it appears that the ANOVA method provides a much more useful and efficient means for clarifying the subjects' knowledge of how to evaluate the seriousness of a problem behavior than do the interview techniques employed.

CONCLUSIONS

The results of this study support the following conclusions:

1.   The results support the hypothesis that there is agreement among the subjects and between the subjects and Class.BD regarding the ratings of seriousness on the problem behavior descriptions.   Both the correlational analyses and the analysis of variance support the hypothesis.

2.   The results do not support the notion that externalized problem behaviors are perceived as more serious than internalized problem behaviors.   For the most part, the subjects gave highly similar ratings when the factor under consideration and the level of the other four factors were comparable.

3. The results support the notion that severity can be separated into three distinct levels, with the subjects providing significantly different ratings across the three levels of severity.   The results do not support the notion of three levels for frequency, duration, generality, and percentage of peers.   The subjects gave highly similar ratings when these factors were present at high and moderate levels and gave much lower ratings when they were at low levels.

4. For the other four factors (frequency, duration,

generality, and percentage of peers), there does not appear to be support for the notion of three levels. The subjects did not provide significantly different ratings between high and moderate levels of these factors. Significant differences were found between moderate and low levels of these factors, providing support for the notion of two levels of these factors or indicating some threshold level of these factors below which the problem behaviors were given significantly lower ratings.

5. The combined levels of the other four factors had an impact on how the subjects rated the problem behavior descriptions. Consistently, there were significant differences between the ratings when the level of the other four factors was high versus moderate versus low. Based on the results of the Eta squared tests, it appears that these significant differences are attributable to the level of severity in each problem behavior description.

Although a number of the interactions in the ANOVAs were statistically significant, a relatively small proportion of the variability in the ratings was accounted for in each of them, according to the Eta-squared tests.

6. The results also call into question the use of only positive weightings for determining the certainty factors for seriousness of problem behaviors in Class.BD. When

the levels of frequency, duration, generality, and percentage peers are low, it may be more appropriate to subtract from the certainty factor, rather than add a smaller number than programmed for moderate and high levels.

7. The results support the notion that an ANOVA approach is a viable method for knowledge engineering, particularly in clinical decision-making situations where cues and decision points are not as discrete as in more concrete areas (such as chemistry and medicine).

8. The results from the interviews showed little consistency among subjects on either ranking the factors or in describing their approach to determining the seriousness of the problem behaviors. Little support was shown for the interview methods employed as viable knowledge engineering methodologies. The ANOVA method proved to be much less time consuming and provided results that were much more useful in the knowledge engineering process.

Future Research

This study was a preliminary attempt to describe the impact of the level of the five behavioral factors as well as whether the behavior is internalized or externalized on the field-based decision maker's perception of the seriousness of a problem behavior. As

such, it has heuristic value, and the results generated a number of further areas to be explored.

Further research is needed to address the following questions:

1.   The existence of some sort of "cutoff point" for determining whether a child is eligible for special education because of a behavioral disorder has not been established. A special education classification decision differs from a diagnosis in that the question is not whether a child has a problem but whether the problem is serious enough to warrant placement and expenditure of special education funds.   Establishing such a cutoff empirically might be useful.

2.   One issue not currently addressed in Class.BD or in this research project is an index of the age appropriateness for a problem behavior.   What might be considered typical behavior for a second grader might be considered extremely deviant in a high school student. Perhaps in some cases the reverse may be true.   Some portion of the variability not accounted for in the three-way ANOVAs in this study is likely attributable to the assumptions each subject had about the age of the student described in each problem behavior description. A study looking at the impact of this variable and whether it is important enough to be included in a future revision of Class.BD is needed.

3. Whether or not field-based decision makers from rural versus urban, northern versus southern Utah, and so forth would rate certain problem behaviors differently remains to be empirically tested. Five of the six subjects in the present study work in Salt Lake City, a large metropolitan area. All six subjects work in northern Utah. Whether their colleagues in other parts of the state or in other states would give highly similar ratings on the problem behavior descriptions remains to be demonstrated.

4. For individuals who prefer to use an interview procedure for knowledge engineering, a more structured approach needs to be developed. Based on the findings regarding the subjects' rankings on the five behavioral factors during the follow-up interviews, it appears that they each approached the task with a different strategy or assumed expectations. A more structured format with specific instructions regarding how to rank the factors might yield more consistent and informative results. Given the more useful results obtained through the ANOVA method in the present study, perhaps the interview procedures are best used only in preliminary stages of knowledge engineering for initial identification of potentially important factors in the decision-making process.

REFERENCES

Balow, B. (1979). Definitional and prevalence problems in behavioral disorders in children. School Psychology Digest, 8(4), 348-354.

Berry, D.C. (1987). The problem of implicit knowledge. Expert Systems, 4(3), 144-151.

Cullinan, D., Epstein, M.H., & McLinden, D. (1986). Status and change in state administrative definitions of behavior disorder. School Psychology Review, 15(3), 383-392.

Duda, R.O., Gaschnig, J.G., & Hart, P.E. (1979). Model design in the PROSPECTOR consultant system for mineral exploration. In D. Michie (Ed.), Expert systems in the microelectronic age. Edinburgh: Edinburgh University Press. 152-167.

Edwards, A.L. (1957). Techniques of attitude scale construction. New York: Appleton-Century-Crofts, Inc.

Epstein, M.H., Cullinan, D., & Sabatino, D.A. (1977). State definitions of behavior disorders. The Journal of Special Education, 11(4), 417-425.

Executive Committee of the Council for Children with Behavioral Disorders (1987). Position paper on the definition and identification of students with behavioral disorders. Behavioral Disorders, 13(1), 9-19.

Federal Register (1977, August 23). Part II (Rules and regulations for amendments to Part B, Public Law 94-142, Education for All Handicapped Children Act of 1975).

Ferrara, J.M., & Baer, R., Althouse, B., & Reavis, K. (1988). Class.BD [computer program]. Logan, UT: Utah State University.

Ferrara, J.M., & Hofmeister, A.M. (1984). Class.LD2: An expert system for classifying learning disabilities. [computer program]. Logan, UT: Utah State University.

Ferrara, J.M., Williams, D., & Giere, S. (1987). Class.IH [computer program]. Logan, UT: Utah State University.

Geissman, J.R., & Schultz, R.D. (1988). Verification and validation of expert systems. AI Expert, 3(2), 26-33.

Gold, S., & Peterson, A. (1988). ClassPH [computer program]. Logan, UT: Utah State University.

Greenburg, D. (1983). A survey of definitions and identification of seriously emotionally disturbed youngsters: Local special education administrator perspectives and processes. A report of survey information. Council of Administrators of Special Education, Inc. ED 247741

Hayes-Roth, F., Waterman, D.A., & Lenat, D.B. (Eds.), (1983). Building expert systems. Reading, MA: Addison-Wesley.

Hink, R.F., & Woods, D.L. (1987). How humans process uncertain knowledge: An introduction for knowledge engineers. <u>AI Magazine</u>, <u>8</u>(3), 41-53.

Hoffman, R.R. (1987). The problem of extracting the knowledge of experts from the perspective of experimental psychology. <u>AI Magazine</u>, <u>8</u>(2), 53-67.

Mack, J.H. (1980). <u>An analysis of state definitions for severely emotionally disturbed</u>. Reston, VA: Council for Exceptional Children, Policy Research Center. ED 201 135.

Martin, W.A., & Fateman, F.J. (1971). The MACSYMA system. In <u>Proceedings of the second symposium on symbolic and algebraic manipulation</u>. Los Angeles, 59-75.

McCoy, M.S., & Levary, R.R. (1988). Augmenting knowledge acquisition processes of expert systems with human performance modeling techniques. <u>IEEE Transactions on Systems, Man, and Cybernetics</u>, <u>18</u>(3), 467-472.

McGinnis, E., Kiraly, J., Jr., & Smith, C.R. (1984). The types of data used in identifying public school students as behaviorally disordered. <u>Behavioral Disorders</u>, <u>9</u>, 239-246.

O'Keefe, R.M., Balci, O., & Smith, E.P. (1987). Validating expert system performance. <u>IEEE Expert</u>, <u>2</u>(4), 81-89.

Olson, J.R., & Reuter, H.H. (1987). Extracting expertise from experts: Methods for knowledge acquisition. Expert Systems, 4(3), 152-168.

Ott, L. (1988). An introduction to statistical methods and data analysis. Boston: PWS-Kent Publishing Company.

Parry, J.D. (1986). Mandate Consultant: An expert system for reviewing regulatory procedures of IEP development [computer program]. Logan, UT: Utah State University.

Shortliffe, E.H. (1976). Computer-based medical consultation: MYCIN. New York: American Elsevier.

Shortliffe, E.H., & Buchanan, B.G. (1984). A model of inexact reasoning in medicine. In B.G. Buchanan & E.H. Shortliffe (Eds.), Rule-based expert systems. Reading, MA: Addison-Wesley.

Swartz, S.L., & Mosley, W.J. (1986). Diagnosing behavior disorders: An analysis of Illinois criteria. Paper presented at the Conference of the Illinois Federation of the Council for Exceptional Children, Rosement, IL, November 6-8, 1986. ED 278 206.

Triggs, T.J. (1988). The ergonomics of decision-making in large scale systems: Information displays and expert knowledge elicitation. Ergonomics, 31(5), 711-719.

Utah State Office of Education (1988). <u>State board of education special education rules.</u> Salt Lake City, UT: Utah State Office of Education.

Walker, H.M., Reavis, H.K., Rhode, G., & Jenson, W.R. (1985). A conceptual model for delivery of behavioral services to behavior disordered children in educational settings. In P.H. Bornstein and A.E. Kazdin (Eds.), <u>Handbook of clinical behavior therapy with children.</u> (pp. 700-741). Homewood, IL: The Dorsey Press.

APPENDICES

Appendix A

Class.BD Rules for Determining the Seriousness

of Problem Behavior

```
/*********control stuff**********/

/*---analysis_from_direct_observation----*/

multivalued(analysis_from_direct_observation).

if cached(TYPE-behavior = BEHAVIOR)
and not(BEHAVIOR == 'not observed')
and not(continuation = no)
and  'child`s problems'-BEHAVIOR-TYPE is sought
then analysis_from_direct_observation.

/*---analysis_from_incidental data---*/

multivalued(analysis_from_incidental_data).

if not(continuation = no)
and  'incidental data' is sought
then analysis_from_incidental_data.

/*----- behavior -----*/

rule-610:
if  'behavior  evaluation'  =  'severe  enough'  cf
BAD_CERTAINTY
and (100-BAD_CERTAINTY) = DIFFERENCE
and (DIFFERENCE*15)/10 = GOOD_CERTAINTY
and GOOD_CERTAINTY <= 100
then behavior = ok cf GOOD_CERTAINTY.

if  'behavior  evaluation'  =  'severe  enough'  cf
BAD_CERTAINTY
and (100-BAD_CERTAINTY) = DIFFERENCE
and (DIFFERENCE*15)/10 = GOOD_CERTAINTY
and GOOD_CERTAINTY > 100
then behavior = ok.

/*---'behavior evaluation'---*/

if temp_eval = X
then 'behavior evaluation' = X.

/*-----continuation -----*/

nocache(continuation).

if temp_eval = 'severe enough' cf CF
    and CF > 60
    and more_information-CF =
            'Go directly to the EDUCATIONAL performance
section.'
then continuation = no.
```

```
   if cached(more_information-CF = 'Consider more BEHAVIOR
data.')
   and do(reset more_information-CF)
then continuation = yes.

continuation = yes.

/*----current_cf----*/
nocache(current_cf).

if cached(TYPE-behavior = BEHAVIOR)
   and cached('child`s problems'-BEHAVIOR-TYPE = 'severe
enough' cf CF)
   then current_cf = bad cf CF.

/*---final_lookup(NUM)---*/

nocache(final_lookup(NUM)).

if slope_final = [TOP,BOTTOM,SLOPE,INTERCEPT]
and NUM < TOP
and NUM >= BOTTOM
and (NUM*SLOPE)+INTERCEPT = CF
then final_lookup(NUM) = CF.

/*--slope_final---*/
multivalued(slope_final).
nocache(slope_final).

slope_final = [101,94,0,100].
slope_final = [94,50,1.5,-40].
slope_final = [50,35,1,-15].
slope_final = [35,-10,0.444,4.444].
slope_final = [-10,-100,0,0].

/*----- more-CERTAINTY -----*/
question(more_information-CERTAINTY) = [
'  A behavior disordered student is defined as one whose
BEHAVIOR or ',
'emotional conduct is serious enough to adversely affect
EDUCATIONAL ',
'performance.',nl,nl,
'  Based on the information you have entered describing
the student`s ',
'BEHAVIOR you can be ',CERTAINTY,' percent certain that
the student`s ',
'behaviors are serious enough to justify a BD
classification in Utah.  ',
'Consideration of more information regarding BEHAVIOR
will probably increase ',
'the level of certainty.',nl,nl,
```

```
'  You now have a decision to make.  You may describe
additional behavior ',
'related information or you may go directly to questions
which consider ',
'evidence   regarding   the   student`s   EDUCATIONAL
performance.',nl,nl,
'    Do you wish to:',nl].

legalvals(more_information-CERTAINTY)  =['Consider  more
BEHAVIOR data.',
                  'Go  directly  to  the  EDUCATIONAL
performance section.'].

enumeratedanswers(more_information-CERTAINTY).

automaticmenu(more_information-CERTAINTY).


explanation(more_information-CERTAINTY) = [
'      Class.BD considers a variety factors  (frequency,
severity, ',
'etc.)  relative  to each problematic behavior  in
arriving  at  a ',
'certainty that a student exhibits problematic behavior
sufficient ',
'to  warrant a BD classification. After considering these
factors ',
'for each behavior, the program  reports the overall
certainty and ',
'allows  the user the option of considering additional
behavioral ',
'data  that may increase the  certainty of moving on  to
consider ',
'other important aspects of BD classification.  ',nl,nl].

rule-120:
if analysis_from_direct_observation is sought
  and analysis_from_incidental_data is sought
  and behavior = ok cf CF
  and CF > 70
  and display(['
  If the data which you have entered is correct, the
problems which you ',
'describe are not very severe.',nl,nl])
  and 'continue with a '- 'low severity rating' =
      'no, stop the consultation'
  and do(abort)
 then 'problematic behavior'.

rule-130:
if 'behavior evaluation' is unknown
  and display(['
```

   Based on standardized behavior measures, incidental data and/or ',
'observation results you cannot be certain that the student`s behavior ',
'problems are severe enough to warrant a BD classification in Utah.',
' Consideration of additional information would be silly.',nl,nl])
   and 'continue with a '- 'lack of problematic behavior' = 'no, stop the consultation'
   and do(abort)
then 'problematic behavior'.

/*---temp_eval---*/

nocache(temp_eval).

if cached(measures = 'standardized instruments suggest problems' cf CF)
 then temp_eval = 'severe enough' cf CF.

if current_cf = bad cf CF
   and final_lookup(CF) = CERT
 then temp_eval = 'severe enough' cf CERT.

rule-620:
if cached('incidental data' = 'suggests problems' cf CF)
 then temp_eval = 'severe enough' cf CF.

/********BEHAVIOR ANALYSIS STUFF*********/

/******severity*******/

/*----- (child`s problems) -- degree -----*/

rule-280:
if severity_level = LEVEL-[CF]
and not(cached(level_found-BEHAVIOR = yes))
and LEVEL-BEHAVIOR-TYPE = yes
and do(set level_found-BEHAVIOR = yes)
then 'child`s problems'-BEHAVIOR-TYPE='severe enough' cf CF.

explanation(rule-280) = [
'   Utah rule C.1 requires that before a student can be ',
'classified as behavior disordered observations to document ',
'problematic behavior be made. The present question is designed to ',
'determine the severity of the behaviors exhibited by the target ',

'student. Severity is one factor considered in
determining if the ',
'student`s behavior is problematic enough to
warrant a BD ',
'classification. ',nl,nl,
'Ref. State Board of Education Special Education Rules
(1988), p. ',
'A-30. ',nl,nl].

```
rule-295:
if not(cached(level_found-BEHAVIOR = yes))
   and 'your opinion'-BEHAVIOR-TYPE = OPINION
   and severity_level = OPINION-[CF]
then 'child`s problems'-BEHAVIOR-TYPE ='severe enough' cf
CF.
```

explanation(rule-295) = [
'   Utah rule C.1 requires that before a student
can be ',
'classified as behavior disordered observations to
document ',
'problematic behavior be made. The present question is
designed to ',
'determine the severity of the behaviors exhibited by
the target ',
'student. Severity is one factor considered in
determining if the ',
'student`s behavior is problematic enough to
warrant a BD ',
'classification. ',nl,nl,

'Ref. State Board of Education Special Education Rules
(1988), p. ',
'A-30. ',nl,nl].

```
/*----severity_level-----*/
multivalued(severity_level).

nocache(severity_level).
severity_level = 'serious risk of physical harm'-[60].
severity_level = severe-[40].
severity_level = moderate-[15].
severity_level = mild-[-20].

/*----- 'serious risk of physical harm'-BEHAVIOR-
externalized -----*/

question('serious risk of physical harm'-BEHAVIOR-
externalized) =[
'  Is the student`s "',BEHAVIOR,'" behavior, so serious
that even one occurrence poses a severe threat to
individuals ',
```

```
'?',nl,nl,'    e.g. Fights where students or teachers ',
'are seriously injured.',nl].


/*----- 'serious  risk  of  physical  harm'-BEHAVIOR-
internalized -----*/

question('serious  risk  of  physical  harm'-BEHAVIOR-
internalized) =[
'  Does the student`s "',BEHAVIOR,'" behavior ',
'put the student or others at physically "at risk"?',nl,
'    e.g.  Hallucinating  in  the  middle  of  a  busy
highway.',nl,nl].

legalvals('serious  risk  of  physical  harm'-BEHAVIOR-
TYPE)=[yes,no].
enumeratedanswers('serious  risk  of  physical  harm'-
BEHAVIOR-TYPE).
automaticmenu('serious  risk  of  physical  harm'-BEHAVIOR-
TYPE).

question(severe-BEHAVIOR-externalized) =[
'  Is  the  student`s  "',BEHAVIOR,'"  behavior,  so  serious
'that  even  one  occurrence  poses  a  severe  threat  to
property ',
'and/or some threat to individuals?', nl,nl,
'    e.g.  lighting  fires  or  damaging  ',  'expensive
equipment.',nl].

legalvals(severe-BEHAVIOR-externalized)=[yes,no].
enumeratedanswers(severe-BEHAVIOR-externalized).
automaticmenu(severe-BEHAVIOR-externalized).

/*----- severe-BEHAVIOR-internalized -----*/

question(severe-BEHAVIOR-internalized) =[
'  Does the student`s "',BEHAVIOR,'" behavior ',
'preclude  or  practically  preclude  social  interaction
and/or ',
'suggests severe emotional/upset?',nl,nl,
'    e.g.  hiding  under  the  chair  in  reading  or  trembling
and hiding in ',
'the presence of a math text book.',nl,nl].

legalvals(severe-BEHAVIOR-internalized)=[yes,no].
enumeratedanswers(severe-BEHAVIOR-internalized).
automaticmenu(severe-BEHAVIOR-internalized).

/*----- moderate-BEHAVIOR-externalized -----*/

question(moderate-BEHAVIOR-externalized) = [
'  Does the student`s "',BEHAVIOR,'" behavior ',
'cause a good deal of classroom disruption and/or pose ',
```

```
'a       moderate     threat     to     individuals     and/or
property?',nl,nl,
'    e.g. fights where no one is injured, or marking
furniture.',nl].

/*----- moderate-BEHAVIOR-internalized -----*/

question(moderate-BEHAVIOR-internalized) = [
'  Does the student`s "',BEHAVIOR,'" behavior limit',
' social interaction and/or suggest moderate emotional
upset?',nl,nl,
'   e.g. staying exclusively with 1 or 2 friends or
occasionally',
' crying for no apparent reason.',nl].

legalvals(moderate-BEHAVIOR-TYPE) = [yes,no].
enumeratedanswers(moderate-BEHAVIOR-TYPE).
automaticmenu(moderate-BEHAVIOR-TYPE).

/*----- mild-BEHAVIOR-externalized -----*/

question(mild-BEHAVIOR-externalized) =[
'  Does the student`s "',BEHAVIOR,'"',
' behavior constitute a nuisance but pose no ',
' threat to individuals and/or property?',nl,nl,
'    e.g. being out of seat, calling out.',nl].

/*----- mild-BEHAVIOR-internalized -----*/

question(mild-BEHAVIOR-internalized) =[
'  Does the student`s "',BEHAVIOR,
'" behavior only slightly limit ',
'social interaction and/or suggest only mild emotional
upset?',nl,nl,
'  e.g. somewhat more frequent absence from school than
',
'average or has more than average (but only occasional)
stomach ',
'or headaches',nl].

legalvals(mild-BEHAVIOR-TYPE)=[yes,no].
enumeratedanswers(mild-BEHAVIOR-TYPE).
automaticmenu(mild-BEHAVIOR-TYPE).

/*----- 'your opinion'-BEHAVIOR -----*/

question('your opinion'-BEHAVIOR-TYPE) = [
'    How would you characterize the student`s
',BEHAVIOR,'?',nl].

legalvals('your opinion'-BEHAVIOR-TYPE) =
[mild,moderate,severe].
```

```
automaticmenu('your opinion'-BEHAVIOR-TYPE).
enumeratedanswers('your opinion'-BEHAVIOR-TYPE).

/******Generality*******/

rule-270:
if TYPE-behavior = BEHAVIOR
   and generality-TYPE-BEHAVIOR = CF
 then 'child`s problems'-BEHAVIOR-TYPE = 'severe enough'
cf CF.

if 'number of problem'(area)-BEHAVIOR = PA
   and PA*2 = PRODUCT
 then 'child`s problems'-BEHAVIOR-TYPE = 'severe enough'
cf PRODUCT.

/*---'generality percentage'-BEHAVIOR---*/

rule-355:
if 'number of problem'(class)-BEHAVIOR = ST
   and 'number of subjects' = SN
   and not(SN = 0)
   and (ST/SN)*100 = X
 then 'generality percentage'-BEHAVIOR = X.

/*----- generality-TYPE-BEHAVIOR -----*/

if 'generality percentage'-BEHAVIOR = PERCENT
   and              generality_slope_facts              =
[TOP,BOTTOM,SLOPE,INTERCEPT]
   and PERCENT < TOP
   and PERCENT >= BOTTOM
   and ((PERCENT*(SLOPE))+INTERCEPT) = CERTAINTY
 then generality-TYPE-BEHAVIOR = CERTAINTY.

/****generality_slope_facts*****/

nocache(generality_slope_facts).

generality_slope_facts = [101,0,0.4,-20].

/*---'number of problem'(CLASS_OR_AREA)-BEHAVIOR---*/

if listof(PLACE,cached(setting-BEHAVIOR = PLACE) and
                  not cached(subjects = PLACE)) = []
then 'number of problem'(area)-BEHAVIOR = 0.

if listof(PLACE,cached(setting-BEHAVIOR = PLACE) and
                  not cached(subjects = PLACE)) = LIST
and length(LIST) = N
then 'number of problem'(area)-BEHAVIOR = N.
```

```
rule-335:
if listof(CLASS,cached(setting-BEHAVIOR = CLASS) and
                    cached(subjects = CLASS)) = []
then 'number of problem'(class)-BEHAVIOR = 0.

rule-340:
if listof(CLASS,cached(setting-BEHAVIOR = CLASS) and
                    cached(subjects = CLASS)) = LIST
and length(LIST) = N
then 'number of problem'(class)-BEHAVIOR = N.


/*---'number of subjects'---*/

if listof(subjects) = []
then 'number of subjects' = 0.

if listof(subjects) = LIST
and length(LIST) = N
then 'number of subjects' = N.

/********** INTENSITY *********/

rule-265:
if mostlikely('recording method'-BEHAVIOR) =  METHOD
and METHOD-evaluation_of-BEHAVIOR = VALUE
then 'child`s problems'-BEHAVIOR-TYPE = 'severe enough'
cf VALUE.

/*---'DURATION (1)'-evaluation_of-BEHAVIOR ----*/

rule-800:
if 'average percentage of'-BEHAVIOR-'total time'-'target
student' = TS
and 'average  percentage  of'-BEHAVIOR-'total  time'-
'comparison student' = CS
and not(TS==0)
and CS/TS = DIV
and ratio_lookup(DIV) = CERT
then 'DURATION (1)'-evaluation_of-BEHAVIOR = CERT.

/*---'DURATION (2)'-evaluation_of-BEHAVIOR ----*/

rule-805:
if 'average observed length of time'-BEHAVIOR-'target
student'= TS
and  'average  observed  length  of  time'-BEHAVIOR-
'comparison student'= CS
and not(TS==0)
and CS/TS = DIV
and ratio_lookup(DIV) = CERT
then 'DURATION (2)'-evaluation_of-BEHAVIOR = CERT.
```

```
/*---EVENT_OR_RATE-evaluation_of-TYPE-BEHAVIOR----*/

rule-835:
if (METHOD=='RATE'
or METHOD=='EVENT')
   and 'general reported frequency'-BEHAVIOR = 'several
times/hour or more'
   and 'observed place'-BEHAVIOR is sought
   and 'observed place'-BEHAVIOR is definite
   and lookup_frequency_cf(['several times/hour or
more',BEHAVIOR]) = BASE
   and 'comparison adjustment'-BEHAVIOR = ADJ
   and (BASE)+(ADJ) = CERT
then METHOD-evaluation_of-BEHAVIOR = CERT.

if (METHOD=='RATE'
or METHOD=='EVENT')
   and 'general reported frequency'-BEHAVIOR = LEVEL
   and lookup_frequency_cf([LEVEL,BEHAVIOR]) = BASE
   and 'comparison adjustment'-BEHAVIOR = ADJ
   and (BASE)+(ADJ) = CERT
then METHOD-evaluation_of-BEHAVIOR = CERT.

/*----'INTERVAL'-evaluation_of-BEHAVIOR----*/

rule-815:
if 'average percentage of'-BEHAVIOR-intervals-'target
student' = TS
and 'average percentage of'-BEHAVIOR-intervals-
'comparison student' = CS
and not(TS==0)
and CS/TS = DIV
and ratio_lookup(DIV) = CERT
then 'INTERVAL'-evaluation_of-BEHAVIOR = CERT.

/*---'PERCENT OF OCCURRENCE'-evaluation_of-TYPE-BEHAVIOR
----*/

rule-845:

if 'average percentage of'-BEHAVIOR-opportunities-'target
student' = TS
and 'average percentage of'-BEHAVIOR-opportunities-
'comparison student' = CS
and not(TS==0)
and CS/TS = DIV
and ratio_lookup(DIV) = CERT
then 'PERCENT OF OCCURRENCE'-evaluation_of-BEHAVIOR =
CERT.

/*---'PERCENT CORRECT'-evaluation_of-TYPE-BEHAVIOR---*/
```

```
rule-850:
if 'average percentage of'-BEHAVIOR-opportunities-'target
student' = TS
and  'average  percentage  of'-BEHAVIOR-opportunities-
'comparison student' = CS
and not(CS==0)
and TS/CS = DIV
and ratio_lookup(DIV) = CERT
then 'PERCENT CORRECT'-evaluation_of-BEHAVIOR = CERT.


/*----'TIME SAMPLE'-evaluation_of-BEHAVIOR---*/

rule-855:
if 'average percentage of'-BEHAVIOR-'time samples'-
'target student' = TS
and 'average percentage of'-BEHAVIOR-'time samples'-
'comparison student' = CS
and not(TS==0)
and CS/TS = DIV
and ratio_lookup(DIV) = CERT
then 'TIME SAMPLE'-evaluation_of-BEHAVIOR = CERT.

/*----ratio_lookup(DIV)----*/

nocache(ratio_lookup(DIV)).

if slope_lookup_ratio = [TOP,BOTTOM,SLOPE,INTERCEPT]
and DIV < TOP
and DIV >= BOTTOM
and (SLOPE*DIV)+INTERCEPT = CERT
then ratio_lookup(DIV) = CERT.

/*---slope_lookup_ratio ----*/
multivalued(slope_lookup_ratio).
nocache(slope_lookup_ratio).

slope_lookup_ratio = [1.1,0.5,-80,40].
slope_lookup_ratio = [0.5,0,-40,20].

/**************PERCENT STUFF**************/

/*-----'average percentage of'-BEHAVIOR-LABEL-STUDENT----
---*/

rule-785:
if listof(PERCENT,'observed place'-BEHAVIOR = PLACE
  and 'specific  percentage  of'-BEHAVIOR-LABEL-STUDENT-
PLACE = PERCENT) = LIST
and not(LIST == [])
and length(LIST) = N
and sum-LIST = T
```

```
and (T/N) = P
then 'average percentage of'-BEHAVIOR-LABEL-STUDENT = P.

/*----'percentage of'-BEHAVIOR-PLACE -----*/

question('specific percentage of'-BEHAVIOR-LABEL-STUDENT-
PLACE) =[
'    In what percentage of ',LABEL,' during the ',PLACE,
'  observation  did  the  ',  STUDENT,'  exhibit  the
',BEHAVIOR,
' behavior?',nl,nl].

legalvals('specific   percentage   of'-BEHAVIOR-LABEL-
STUDENT-PLACE) = integer(0,100).

explanation('specific  percentage  of'-BEHAVIOR-LABEL-
STUDENT-PLACE) = [
'    Utah  rule  C.1  requires  that  before  a   student
can  be ',
'classified  as  behavior  disordered  observations  to
document ',
'problematic behavior be made. The present question is
designed to ',
'determine if the percent exhibited by the target ',
'student   suggests  that  the  ',BEHAVIOR,'  behavior  is
problematic ',
'enough  to  warrant  a  BD classification. ',nl,nl,

'Ref. State Board of Education Special Education Rules
(1988),  p. ',
'A-30.  ',nl,nl].

/************LENGTH STUFF***********/

/*-----'average observed length'-BEHAVIOR-LABEL-STUDENT--
-----*/

if listof(TIME,'observed place'-BEHAVIOR = PLACE and
 'specific observed length of time'-BEHAVIOR-STUDENT-
PLACE=TIME)=LIST
and not(LIST == [])
and length(LIST) = N
and sum-LIST = T
and (T/N)*100 = P
  then 'average observed length of time'-BEHAVIOR-STUDENT
= P.

/*--duration_lookup(LENGTH)--*/
multivalued(duration_lookup(LENGTH)).

nocache(duration_lookup(LENGTH)).
```

```
duration_lookup('less than a minute') = [1,second].
duration_lookup('more than a minute but less than an
hour') =
[60,minute].
duration_lookup('more than an hour') = [3600,hour].


/*---'general length'-BEHAVIOR-STUDENT-PLACE---*/

question('general length'-BEHAVIOR-STUDENT-PLACE) = [
'   When the ',STUDENT,' was observed in the ',PLACE,'
setting ',
'about how long did it take to complete the ',BEHAVIOR,
' behavior?',nl,nl].

legalvals('general length'-BEHAVIOR-STUDENT-PLACE) = [
'less than a minute',
'more than a minute but less than an hour',
'more than an hour'].

automaticmenu('general length'-BEHAVIOR-STUDENT-PLACE).
enumeratedanswers('general length'-BEHAVIOR-STUDENT-
PLACE).

/*----observed-'duration of'-BEHAVIOR-'in'-TIME_BLOCK----
*/

question(observed-'duration of'-BEHAVIOR-'in'-TIME_BLOCK)
= [
'What was the average observed duration of the
',BEHAVIOR,' behavior ',
'in ',TIME_BLOCK,'s?',nl,'NOTE: If the average duration
was less than ',
'one ',TIME_BLOCK,' enter 0.',nl,nl].

legalvals(observed-'duration         of'-BEHAVIOR-'in'-
TIME_BLOCK) = integer.

/*-----'specific length'-BEHAVIOR-STUDENT-PLACE-UNIT----
*/

question('specific length'-BEHAVIOR-STUDENT-PLACE-UNIT) =
[
'   When the ',STUDENT,' was observed in the ',PLACE,'
setting ',
'how many ',UNIT,'s did it take him/her to do the
',BEHAVIOR,
' behavior?',nl,'NOTE: Please round to the nearest
',UNIT,'.',nl,nl].

legalvals('specific length'-BEHAVIOR-STUDENT-PLACE-UNIT)
= integer.
```

```
/*-----'specific observed length of time'-BEHAVIOR-
STUDENT-PLACE=TIME-----*/

if 'general length'-BEHAVIOR-STUDENT-PLACE = LENGTH
 and duration_lookup(LENGTH) = [MULT,UNIT]
 and 'specific length'-BEHAVIOR-STUDENT-PLACE-UNIT = N
 and MULT*N = TIME
then 'specific observed length of time'-BEHAVIOR-STUDENT-
PLACE=TIME.


/*************FREQUENCY STUFF*************/

/*---'average frequency'-BEHAVIOR-UNIT-STUDENT---*/

if listof(FR,'observed place'-BEHAVIOR = PLACE
 and 'specific frequency'-BEHAVIOR-hour-STUDENT-PLACE  =
FR)=LIST
 and not(LIST == [])
 and length(LIST) = N
 and sum-LIST = SUM
 and SUM/N = AVG
then 'average frequency'-BEHAVIOR-hour-STUDENT = AVG.


if 'specific frequency'-BEHAVIOR-UNIT-STUDENT-all = FR
then 'average frequency'-BEHAVIOR-UNIT-STUDENT = FR.


/*--- 'comparison adjustment'-BEHAVIOR ----*/

if cached('specific frequency'-BEHAVIOR-UNIT-'target
student'-PLACE = 0)
then  'comparison adjustment'-BEHAVIOR = -10.

if cached('average frequency'-BEHAVIOR-UNIT-'target
student' =TS)
and not(TS == 0)
and  'average  frequency'-BEHAVIOR-UNIT-'comparison
student'=CS
and  CS/TS = DIV
and -10*DIV = ADJ
then  'comparison adjustment'-BEHAVIOR = ADJ.

/*-----'general reported frequency'-BEHAVIOR -----*/

question('general reported frequency'-BEHAVIOR) = [
'  In general, how frequently did the ',BEHAVIOR,'
occur?',nl].

legalvals('general reported frequency'-BEHAVIOR) = [
'several times/hour or more',
'a few times/day',
```

```
'one or more times per week',
'less than once/week',
'less than once/month'].

automaticmenu('general reported frequency'-BEHAVIOR).
enumeratedanswers('general reported frequency'-BEHAVIOR).

explanation('general reported frequency'-TYPE-BEHAVIOR) =
[' Utah  rule  C.1  requires  that  before  a  student
can  be ',
'classified  as  behavior  disordered  observations  to
document ',
'problematic behavior be made. The present question is
designed to ',
'determine the frequency of the behaviors exhibited by
the  target ',
'student.  Frequency  is  one  factor  considered  in
determining if the ',
'student`s  behavior  is  problematic  enough  to
warrant  a  BD ',
'classification.  ',nl,nl,

'Ref. State Board of Education Special Education Rules
(1988),  p. ',
'A-30.  ',nl,nl].

/*---lookup_frequency_cf([LEVEL,BEHAVIOR])---*/

lookup_frequency_cf(['less than once/week',BEHAVIOR]) = -
20.
lookup_frequency_cf(['less than once/month',BEHAVIOR]) =
-40.

if  slope_frequency(LEVEL) = [UNIT,SLOPE,INTERCEPT]
and 'average frequency'-BEHAVIOR-UNIT-'target student' =
TIMES
and  (TIMES*SLOPE)+INTERCEPT = CF
then lookup_frequency_cf([LEVEL,BEHAVIOR]) = CF.

/*---slope_frequency(LEVEL)---*/

slope_frequency('one  or  more  times  per  week')  =
[week,6.25,-21.25].
slope_frequency('a few times/day') = [day,1,9].
slope_frequency('several  times/hour  or  more')
=[hour,0.263,14.737].
slope_frequency('a few times/day') = [hour,0.263,14.737].


/*---'specific  frequency'-BEHAVIOR-UNIT-STUDENT-PLACE---
*/
```

```
  question('specific frequency'-BEHAVIOR-UNIT-STUDENT-all)
=[
'   What was the total number of ',BEHAVIOR,' behaviors
per ',UNIT,
'  which were recorded for the ',STUDENT,'in all
settings?',nl,nl].

  question('specific      frequency'-BEHAVIOR-UNIT-STUDENT-
PLACE) =[
'   How many ',BEHAVIOR,' behaviors per ',UNIT,' were
recorded ',
'for the ',STUDENT,' in the "',PLACE,'" setting?',nl,nl].

  legalvals('specific      frequency'-BEHAVIOR-UNIT-STUDENT-
PLACE) = integer.

/********* duration ********/

if duration-BEHAVIOR-TYPE = LENGTH
and cf_lookup_for_duration-LENGTH = CF
then 'child`s problems'-BEHAVIOR-TYPE= 'severe enough' cf
CF.

/******cf_lookup_for_duration-LENGTH******/

cf_lookup_for_duration-'over a year' = 15.
cf_lookup_for_duration-'6 months to 1 year' = 10.
cf_lookup_for_duration-'3 to 6 months' = 7.
cf_lookup_for_duration-'1 to 2 months' = -20.
cf_lookup_for_duration-'less than 1 month' = -40.

/*----- duration-BEHAVIOR-TYPE -----*/

question(duration-BEHAVIOR-TYPE) = [
'   How long has the ',BEHAVIOR,' behavior been
occurring?',nl].


legalvals(duration-BEHAVIOR-TYPE) = ['less than 1
month','1 to 2 months',
'3 to 6 months','6 months to 1 year','over a year'].
automaticmenu(duration-BEHAVIOR-TYPE).

enumeratedanswers(duration-BEHAVIOR-TYPE).


explanation(duration-BEHAVIOR-TYPE) = [
'    Utah rule C.1 requires that before a student
can be ',
'classified as behavior disordered observations to
document ',
'problematic behavior be made. The present question is
```

```
designed to ',
'determine the duration (for how long they have been
occurring) of ',
'the behaviors exhibited by the target student. Duration
is one ',
'factor   considered   in  determining  if  the  student`s
behavior is ',
'problematic enough to warrant a BD classification.
',nl,nl,

'Ref. State Board of Education Special Education Rules
(1988), p. ',
'A-30.  ',nl,nl].

/*----- (child`s problems) -- in_class_peer_data -----*/

if percent_peers-BEHAVIOR = PERCENT
and peer_slope_lookup =[TOP,BOTTOM,SLOPE,INTERCEPT]
and PERCENT < TOP
and PERCENT >= BOTTOM
and (PERCENT*SLOPE)+INTERCEPT = CF
 then 'child`s problems'-BEHAVIOR-TYPE = 'severe enough'
cf CF.

/*----'other students'-BEHAVIOR-CLASS---*/

question('other students'-BEHAVIOR-CLASS) = [
'  How many other students were in ',CLASS,' class',
' when the "',BEHAVIOR,'" behavior was observed?',nl,nl].

legalvals('other students'-BEHAVIOR-CLASS) = integer.


/*---peer_slope_lookup---*/

multivalued(peer_slope_lookup).

nocache(peer_slope_lookup).

peer_slope_lookup = [101,30,-0.143,-45.714].
peer_slope_lookup = [30,10,-2,10].
peer_slope_lookup = [10,0,-2.222,12.222].

/*----- 'problem kids'-BEHAVIOR-CLASS-COUNT -----*/

question('problem kids'-BEHAVIOR-CLASS-COUNT) = [
'  How many (if any) of the ',COUNT,' other students in
',CLASS,' class',
' were exhibiting the "',BEHAVIOR,'" behavior at about
the',
'  same  (or  a  higher)  level  as  the  target
student?',nl,nl].
```

```
legalvals('problem kids'-BEHAVIOR-CLASS-COUNT) = integer.


explanation('problem kids'-BEHAVIOR) = [
'        Utah  rule  C.1  requires that  before  a  student
can  be ',
'classified  as  behavior  disordered  observations  to
document ',
'problematic  behavior  be made.  Further,  the  rule
requires  that ',
'observations  also  be made on at least one  comparison
student. ',
'Class.BD expands on this rule and asks for information
regarding ',
'how many other  students  in the  class exhibit the  same
problematic ',
'behaviors  as  the  target  student.  It uses  the
information  to ',
'determine  how  typical  the problematic  behaviors  are
of  all ',
'students  in the class. How atypical the behaviors of
the  target ',
'student  are  is  one factor considered in  determining
if  the ',
'student`s  behavior  is  problematic  enough  to
warrant  a  BD ',
'classification.',nl,nl,

'Ref.  State  Board of Education Special Education  Rules
(1988), ',
'p. A-30.  ',nl,nl].

/*---percent_peers----*/

rule-275:
if total-'problem kids'-BEHAVIOR = BAD
   and total-'other students'-BEHAVIOR = OTHERS
   and not(OTHERS == 0)
   and (BAD/OTHERS)*100 = PERCENT
then percent_peers-BEHAVIOR = PERCENT.

/*----- sum-LIST -----*/

nocache(sum-LIST).

sum-[] = 0.

if sum-TAIL = SUM_OF_REST
   and SUM_OF_REST + HEAD = SUM
then sum-[HEAD|TAIL] = SUM.
```

```
/*******observation info**********/

/*---'observed place'-BEHAVIOR---*/
multivalued('observed place'-BEHAVIOR).

if  listof(SETTING,cached(setting-BEHAVIOR=  SETTING))  =
LIST
and length(LIST) = N
and N > 1
and 'data collected in'-LIST-BEHAVIOR = PLACE
then 'observed place'-BEHAVIOR = PLACE.

rule-795:
if  setting-BEHAVIOR is unique
and setting-BEHAVIOR = SETTING
and BEHAVIOR-'data collected'-SETTING = yes
then 'observed place'-BEHAVIOR  = SETTING.

/*--- 'data collected in'-LIST-BEHAVIOR----*/

multivalued('data collected in'-LIST-BEHAVIOR).

question('data collected in'-LIST-BEHAVIOR) = [
'In which of the settings listed below were data ',
'describing the ',BEHAVIOR,' behavior collected?',nl,
' Note: type "u" for unknown if no data were
collected',nl,nl].

legalvals('data collected in'-LIST-BEHAVIOR) = LIST.
enumeratedanswers('data collected in'-LIST-BEHAVIOR).
automaticmenu('data collected in'-LIST-BEHAVIOR).

explanation('data collected'-LIST-BEHAVIOR) = [
'Data describing the ',BEHAVIOR,' behavior may not have
been collected ',
'in all possible settings.  The computer program is
seeking information ',
'about where ',BEHAVIOR,' data was collected.',nl,nl].

/*---BEHAVIOR-'data collected'-SINGLE_MEMBER---*/

question(BEHAVIOR-'data collected'-SINGLE_MEMBER) = [

'Do you have access to data on ',BEHAVIOR,' in the
"',SINGLE_MEMBER,'" ',
'setting?',nl,nl].

legalvals(BEHAVIOR-'data collected'-SINGLE_MEMBER)
=[yes,no].

enumeratedanswers(BEHAVIOR-'data collected'-
SINGLE_MEMBER).
```

```
automaticmenu(BEHAVIOR-'data collected'-SINGLE_MEMBER).

explanation(BEHAVIOR-'data collected'-SINGLE_MEMBER) = [
'Data describing the ',BEHAVIOR,' behavior may not have
been collected ',
'in all possible settings.  The computer program is
seeking information ',
'about where ',BEHAVIOR,' data was collected.',nl,nl].

/*---- total-'other students'-BEHAVIOR ----*/

if listof(CLASS_KIDS, cached('observed place'-BEHAVIOR =
CLASS) and
                        cached(subjects = CLASS) and
                        'other students'-BEHAVIOR-CLASS =
                        CLASS_KIDS) = LIST
  and sum-LIST = SUM
then total-'other students'-BEHAVIOR = SUM.


if listof(BAD_KIDS, cached('observed place'-BEHAVIOR =
CLASS) and
cached(subjects = CLASS) and
'other students'-BEHAVIOR-CLASS =
COUNT and 'problem kids'-BEHAVIOR-CLASS-COUNT
=BAD_KIDS) = LIST
and sum-LIST = SUM
then total-'problem kids'-BEHAVIOR = SUM.



/******************INCIDENTAL STUFF**************/

/*----- agency reports -----*/
question('agency reports') = [
'  During the last six months, how many agency reports
citing the student`s ',
'problem behaviors outside of school have been
made?',nl].

legalvals('agency reports') = integer.

explanation('agency reports') = [
'       In addition to observations in school, reports
from  social ',
'service agencies can be used to document that a student
exhibits ',
'problematic  behavior. Class.BD uses information gained
from  the ',
'present  question to increase its confidence that  the
student`s ',
```

```
'behavior is problematic enough to warrant a BD
classification.',nl,nl].


/*----- incidental data -----*/

rule-680:
if 'agency reports'= X
  and X>0
 then 'incidental data' = 'suggests problems' cf 5.

rule-685:
if law = yes
 then 'incidental data' = 'suggests problems' cf 5.

rule-690:
if parents = yes
 then 'incidental data' = 'suggests problems' cf 5.

/*----- law -----*/

question(law) = [
'    During the last six months, has the student`s
behavior problem resulted in trouble with the law?',nl].

legalvals(law) = [yes,no].
enumeratedanswers(law).
automaticmenu(law).

explanation(law) = [
'    In addition to observations in school, reports
from law ',
'enforcement agencies can be used to document that a
student ',
'exhibits problematic behavior. Class.BD uses
information gained ',
'from the present question to increase its confidence
that the ',
'student`s behavior is problematic enough to
warrant a BD ',
'classification.  ',nl,nl].

/*----- parents -----*/

question(parents) = [
'  During the last two months, have the student`s parents
 reported trouble at home?',nl].

legalvals(parents) = [yes,no].
enumeratedanswers(parents).
automaticmenu(parents).
```

```
explanation(parents) = [
'        In addition to observations in school, parental
reports can ',
'be used to document that a student exhibits problematic
behavior. ',
'Class.BD  uses information gained from  the present
question to ',
'increase    its    confidence    that    the    student`s
behavior  is ',
'problematic enough to warrant a BD classification.
',nl,nl].


if sum-TAIL = SUM_OF_REST
  and SUM_OF_REST + HEAD = SUM
 then sum-[HEAD|TAIL] = SUM.


/*---- total-'other students'-BEHAVIOR ----*/

if listof(CLASS_KIDS, cached('observed place'-BEHAVIOR =
CLASS) and
                    cached(subjects = CLASS) and
                    'other students'-BEHAVIOR-CLASS =
                    CLASS_KIDS) = LIST
  and sum-LIST = SUM
then total-'other students'-BEHAVIOR = SUM.


if listof(BAD_KIDS, cached('observed place'-BEHAVIOR =
CLASS) and
                    cached(subjects = CLASS) and
                    'other students'-BEHAVIOR-CLASS =
COUNT and 'problem kids'-BEHAVIOR-CLASS-COUNT
= BAD_KIDS) = LIST
and sum-LIST = SUM
then total-'problem kids'-BEHAVIOR = SUM.
```

Appendix B

Public Law 94-142 Definition

Serious Emotional Disturbance

Public Law 94-142 Definition of

Serious Emotional Disturbance

A condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree, which adversely affects educational performance:

A. An inability to learn which cannot be explained by intellectual, sensory, or health factors;

B. An inability to build or maintain satisfactory interpersonal relationships with peers or teachers;

C. Inappropriate types of behavior or feelings under normal circumstances;

D. A general pervasive mood of unhappiness or depressions; or

E. A tendency to develop physical symptoms or fears associated with personal school problems.

The term does not include children who are socially maladjusted unless it is determined that they are seriously emotionally disturbed.

Appendix C

Utah Definition of Behavior Disordered

CATEGORY: **BEHAVIOR DISORDERED**

student as primarily behavior disordered, it must be determined that:

1.   The student is not <u>primarily</u> identified as manifesting one of the other handocapping conditions described in these Rules.

2.   The student is not behaving as a behavior disordered student because of vision or hearing impairment.

3.   The stuent is not behaving as a behavior disordered student because of (1) an inappropriate classroom discipline system, (2) breakdown of classroom discipline, or (3) inappropriate academic instruction or materials.

Disclaimer information may include:

a.   Data in the cumulative records.

b.   Interviews and classroom observations.

c.   Evaluations.

C. Assessment For Classification. A complete formal and informal evaluation covering behavior-social-educational areas is required before classficaiton as behavior disordered. In addition to the requirement for the composition of the team (see Rules III.E.4 and III.G), one member of the team must be a certified school psychologist, a certified social worker, a certified school counselor, or a qualified teacher for the type of program for which the student is being referred.

1.   Classroom observations of the student should include at least three fifteen-minute observations on referring behavior pinpoints. A student who is non-handicapped and who is not being referred must be selected and observed in the same setting on the same behavior pinpoints as the referred student for comparision. It is expected that the classroom observations will be made by an assigned member of the assessment team.

2.   Every student classified as behavior disordered will have complete documentation in his/her

records regarding each of the following areas:

a.   Educational behavior.

Information may include:

(1) classroom academic screeings and tests.

(2) achievement tests.

(3) report cards/cumulative records.

b.   Social/adamptive behavior checklists or scales, and sociometric devices.

Information may include:

(1) student's past and present patterns of interaction with peers, family, teachers, adults, etc.

(2)   teacher/parent checklists.

Many observation, checklist and other assessment instruments address issues across adaptive and social areas. Scale and checklist must be age appropriate for the student being assessed.

D.   IEP.   Behavior objectives for which the student is referred (initial referrals or referrals to more restrictive settings) must be addressed in the IEP goals and objectives.

When a student is classified as behavior disordered or is referred to a more restrictive setting, a plan to return the student to a less restrictive or regular class must be developed specifying appropriate re-entry behaviors and classroom environment.

Appendix D

Definitions for Severity Levels and

Externalized/Internalized

Definitions for Severity Levels and Externalized/Internalized

High

<u>Externalized</u> - even one occurrance poses a severe threat to individuals and/or property - ex. lighting fires, damaging expensive equipment.

<u>Internalized</u> - precludes or practically precludes social interaction and/or suggests severe emotional upset.

Moderate

<u>Externalized</u> - causes a good deal of classroom disruption and/or poses a moderate threat to individuals and/or property eg. fights where no one is injured or marking furniture.

<u>Internalized</u> - limits social interaction and/or suggests moderate emotional upset.

Mild

<u>Externalized</u> - muisance; poses no or only a minimal threat to individuals and/or property, eg. being out of seat, calling out.

<u>Internalized</u> - slightly limits social interaction and/or suggests mild emotional upset.

IGNORE FREQUENCY, DURATION, PERCENT OF PEERS AND GENERALITY.

Appendix E

Problem Behavior Descriptions

Jeff has been setting fires on school property, eg., next to storage buildings, in the parking lot, etc.
Jeff has been setting fires on school property approximately twice per week for 2 months.
Jeff sets fires during recess and lunch, and has skipped two of his 5 classes to start fires.
About 10 percent of the students in Jeff's class have participated in fire setting on school grounds.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                      moderate                    serious
```


Comments:

E S-H F-M D-M G-M %-M

Linda is often non-complicant, publicly defying the teacher (talks back and refuses to cooperate), which her teachers feel is very disruptive.
Linda has been non-compliant 5-6 times per week on average, and has been behaving this way on a regular basis for the past 3 months.
Linda has refused to comply with teacher directions in 5 of the 9 classes she is taking (math, reading, phy. ed., music, and social studies).
Approximately 13% of the students in Linda's grade have exhibited similar behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                  serious
```


Comments:

E S-M F-M D-M G-M %-M

Kevin has trouble staying seated during class time.
He is out of seat 4 to 5 times per week, a problem which
has been going on for 4 months.
Approximately 17% of his classmates exhibit similar
difficulty in staying seated.
Jeff has particular difficulty staying seated during
math, reading, and science classes (he has less
difficulty during his other three classes).

Please rate the seriousness of the behavior on the
following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                  serious
```
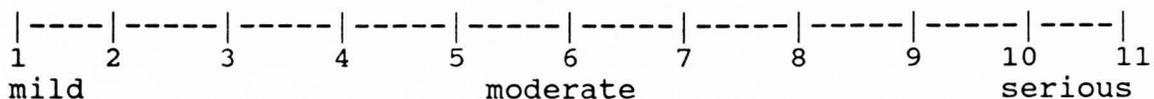
Comments:

E S-L F-M D-M G-M %-M

Jenny has been getting into fights with her classmates. Most of the time, the fights are verbal, but she has gotten into physical fights on occassion. Generally, no one involved gets hurt very seriously.
Jenny has been getting into fights approximately twice per day over the last 2 months.
Jenny has gotten into fights during lunch, recess, reading, and science classes (she is enrolled in 3 other subjects).
Approximately 11% of Jenny's classmates have initiated fights in school.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                      moderate                    serious
```
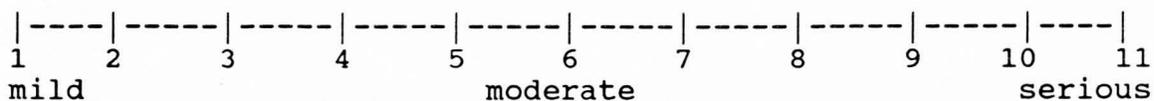

Comments:

E S-M F-H D-M G-M %-M

Rick has been bringing toys to class and playing with them during class time. He refuses to cooperate when the teacher asks him to put the toys away. Other children watch him and talk to him, which the teacher finds very disruptive.
Rick has done this approximately once per month for the last 5 months.
He has done this in 4 of his 7 classes.
Approximately 13% of the other student's in Rick's class have also disrupted class by bringing in balls and other toys.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                      moderate                    serious
```
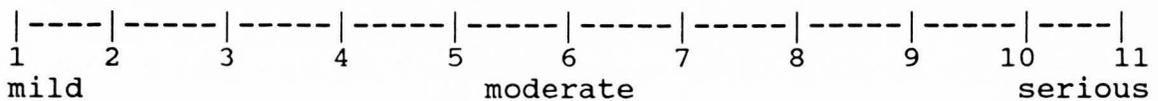

Comments:

E S-M F-L D-M G-M %-M

Lisa refuses to be quiet during class. She talks out in class, and will not raise her hand and wait to be called upon.   This is against class rules, she bothers other students, and the teacher finds Lisa's behavior very disruptive.
Lisa talks out in class 8 times per week on average, and has been doing so for approximately 7 months.
Lisa talks out in 4 of her 7 classes. Approximately 10% of the other student's in Lisa's class talk out from time to time.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                      moderate                    serious
```


Comments:

E S-M F-M D-H G-M %-M

Karen talks loudly to the other students sitting next to her during class. She has succeeded in disrupting the entire class schedule by enlisting the other students in disruptive activities.
Karen has been doing this about 4 times per week on average for approximately 3 weeks.
Karen talks to her neighbors during 3 of her 6 classes.
Approximately 11% of Karen's classmates talk to other students while the teacher is lecturing.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1     2     3     4     5     6     7     8     9     10   11
mild                          moderate                   serious
```

Comments:

E S-M F-M D-L G-M %-M

Mark writes in his school books and in library books, and sometimes tears out pages when he likes the pictures on them.
Mark has been doing this 4 to 5 times per week for the last 4 months.
Mark has done this in all of his classes in which school-owned materials are used. Approximately 10% of his classmates have damaged books at one time or another.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9    10   11
mild                        moderate                 serious
```
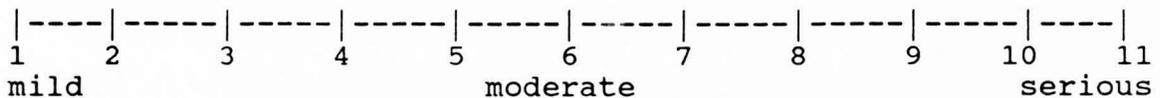
Comments:

E S-M F-M D-M G-H %-M

Theresa kicks the seat of the student seated in front of her when the teacher is trying to lecture. The noise disrupts the class; this behavior has also resulted in the other student getting bruised, and even falling out of their seat on occassion.
Theresa has been doing this 4 to 5 times per week for the last two months.
Theresa does this in math class, one of her 6 classes. Approximately 12% of the other students in Theresa's class exhibit the same behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1     2     3     4     5     6     7     8     9     10   11
mild                          moderate                   serious
```
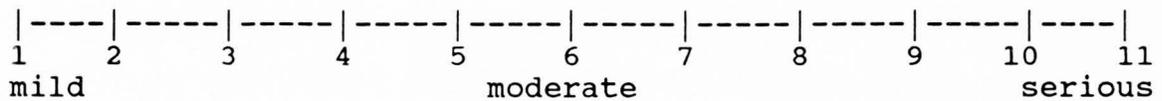
Comments:

E S-M F-M D-M G-L %-M

Bob gets out of his seat and talks to other students when they are supposed to be doing seatwork, which is against class rules and the teacher finds this disruptive. The teacher has had to adjust his lesson plans to accomodate for class time lost due to these disruptions.
Bob has been doing this 5 to 6 times per week for the last 2 months.
Bob talks during seatwork time in 4 of his 7 classes. None of the other students initiate conversations with each other during seatwork time.

Please rate the seriousness of the behavior on the following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   serious
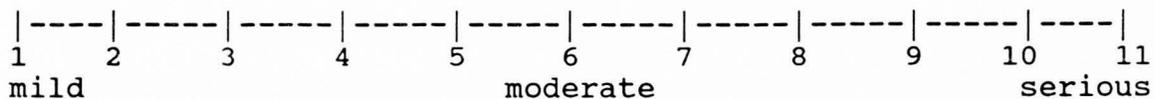
Comments:

E S-M F-M D-M G-M %-H

Angela talks back to the teacher when the class is given
a direction. Some of the other students tend to mimic
her behavior, and the teacher finds it difficult to
regain control of the class.
Angela has been talking back to the teacher 3-5 times per
week over the last 3 months.
Angela has talked back to the teacher during 3 of her six
class period.
Approximately 25% of the other students in Angela's class
engage in similar behavior.

Please rate the seriousness of the behavior on the
following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                  serious
```


Comments:

E S-M F-M D-M G-M %-L

Jeff has been setting fires in the school and on the school grounds.
He attempts to set a fire several times per day. He has been doing this for approximately six months.
Jeff has attempted to do this in all of his classes.
None of Jeff's classmates engage in fire starting behavior.

Please rate the seriousness of the behavior on the following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   serious

Comments:

E S-H F-H D-H G-H %-H

Mark is easily distracted when in class; when asked a question he will talk about whatever is on his mind at the time, regardless of the question; at times he will talk out in class, bringing everyones attention to what he is attending to. The teacher finds this very disruptive.
Mark has been doing this several times per day for the last 7 months.
He does it in all of his classes; none of the other students in Mark's class engage in this kind of behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9    10   11
mild                        moderate                  serious
```

Comments:

E S-M F-H D-H G-H %-H

Margaret does not pay attention in class. Rather than attending to the lessons and doing her seat work, Margaret prefers to read a library book.
Margaret has been doing this several times per day for over 6 months.
She reads when she is supposed to be paying attention in 5 of her 6 classess.
None of Margaret's classmates engage in this type of behavior.

Please rate the seriousness of the behavior on the following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   serious

Comments:

E S-L F-H D-H G-H %-H

Mike fights with other students on the playground and in
class; these fights usually result in injury to the other
student (black eyes, etc.).
Mike picks fights with other students an average of 4
times per week.
This has been going on for the last 6 months.
Mike has picked fights with other students in each one of
his classes; he has not done this in the lunch room.
Approximately 2% of the students in Mike's grade in his
school initiate fights.

Please rate the seriousness of the behavior on the
following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   serious

Comments:

E S-H F-M D-H G-H %-H

Judy has been stealing things from the teachers lounge
and the supply room - things like money and car keys.
The school staff are especially concerned because Judy
has stolen cars and been involved in accidents.
Judy has been caught stealing four times in the last six
months.
Judy skipped out of all of her classes; she has been
caught in the teachers lounge or supply room at all hours
of the school day.
None of Judy's peers exhibit similar behavior.

Please rate the seriousness of the behavior on the
following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9    10    11
mild                        moderate                   serious
```
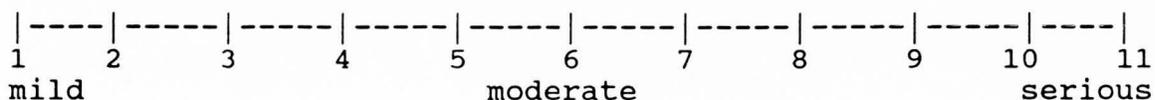

Comments:

E S-H F-L D-H G-H %-H

Anthony has been systematically taking apart the computer equipment in the lab and the teacher suspects he is selling the parts. He has ruined 3 computers, and four others are unusable until replacement parts arrive.
Somehow, Anthony manages to do this when he is in the lab (under supervision), which is attended in 4 of his 5 classes. This has been occuring once or twice per day for the last 2 months.
None of Anthony's classmates attempt to dismantle the computers.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                          moderate                  serious
```
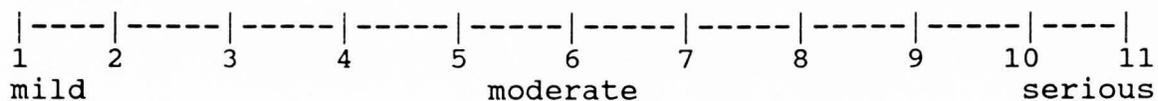
Comments:

E S-H F-H D-M G-H %-H

Gail has been throwing violent temper tantrums in class;
this usually occurs after she has been given a direction
to do her work. During these episodes she does not
appear to control herself or pay attention to her
surroundings, and often gets serious cuts and bruises.
Gail has tantrummed in 5 of her 6 classes (all but phy.
ed.).
This has been occuring 2 to 4 times per day for the last
2 weeks.
None of Gail's peers engage in similar behavior.

Please rate the seriousness of the behavior on the
following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                  serious
```


Comments:

E S-H F-H D-L G-H %-H

When Doug gets angry, with teachers or other students, he
has often threaten them with a knife; when he doesn't
have a knife (5 have been taken away from him so far), he
finds other sharp objects and threatens to hurt people
with them.   He has, on occassion, attempted to follow
through on these threats; so far, an adult has been able
to intervene before anyone was seriously hurt.
Doug as done this at least once per day over the last six
months.   It has occurred in only two of his five classes
(math and science) and on the playground and in the
halls.
Approximately 3% of the other students in Doug's class
have engaged in similar behavior.

Please rate the seriousness of the behavior on the
following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   serious

Comments:

E S-H F-H D-H G-M %-H

When Joyce is upset she screams and throws objects at the teacher and other students (often resulting in injuries). This has been occuring at least once per day for the last 6 months.
Joyce only does this in her math class.
None of the other students in Joyce's class exhibit similar behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                  serious
```

Comments:

E S-H F-H D-H G-L %-H

When angry, Alan has been physically attacking teachers
and other school personnel. The reasons for his anger
are not always obvious. During these attacks, Alan
typically kicks the individual and tries to bite them
(usually with success).
These attacks have been occuring once or twice per day;
he has been attacking school personnel for approximately
6 months.
Alan has done this during all of his classes and in the
lunchroom and principal's office.
Approximately 10% of the student's in Alan's grade have
attacked school personnel.

Please rate the seriousness of the behavior on the
following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                  serious
```
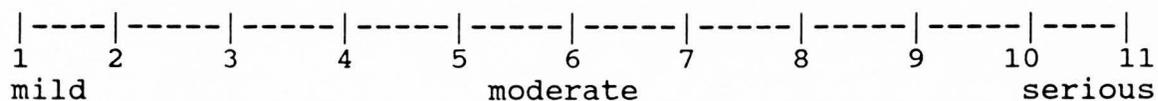

Comments:

E S-H F-H D-H G-H %-M

Beth has been breaking windows in the school, usually
when left unsupervised for even the briefest amount of
time.
She has been doing this almost daily for 6 months, and no
interventions in the classroom have been successful.
She has done this in 5 of her six classes and in two of
the restrooms in the school.
Approximately 20% of Beth's classmates break windows in
the school and other buildings.

Please rate the seriousness of the behavior on the
following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9     10    11
mild                        moderate                   serious
```

Comments:

E S-H F-H D-H G-H %-L

Caroline has been caught chewing gum and eating candy in class, which is against school rules.
She was caught doing this once during the last two months.
This has only occured in one class, reading.
30% of the other students in Caroline's class have been caught chewing gum and eating candy.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9     10    11
 mild                    moderate                     serious
```


Comments:

E S-L F-L D-L G-L %-L

When angry with teachers or other students, Terrie hits
and bites other students and the teacher.
This has occurred twice in the last three months.
This has occurred in two of Terrie's classes, and never
occurred in settings outside the classroom.
20% of Terrie's peers have engaged in this type of
behavior at one time or another during the past two
school years.

Please rate the seriousness of the behavior on the
following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9     10    11
mild                        moderate                  serious
```


Comments:

E S-H F-L D-L G-L %-L

To amuse himself when bored during class, Dick makes funny, humming noises, and gets out of his seat and takes other students study materials away from them. This attracts the attention of the other students and the teacher finds this behavior disruptive.
Dick has done this once in the last month, during social studies class.
Approximately 40% of Dick's classmates engage in similar behavior.

Please rate the seriousness of the behavior on the following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   serious

Comments:

E S-M F-L D-L G-L %-L

Joe does not pay attention in class, preferring to doodle and color with his marking pens.
This occurs several times per day, and has been going on for the last three and a half weeks.
Joe does this in two of his 6 classes, and attends to adults in other school settings.
Approximately 35% of Joe's classmates engage in similar behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1     2     3     4     5     6     7     8     9     10    11
mild                          moderate                  serious
```


Comments:

E S-L F-H D-L G-L %-L

Diane has been caught lying, saying that she has her homework done when it is not complete. Her teacher finds this behavior irritating, but not disruptive to the class.
This has happened approximately three times per week for the last month.
She has lied about her homework in one of her classes, math.
Approximately 25% of Diane's classmates have been caught lying about having their homework done when it wasn't.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9     10    11
mild                      moderate                    serious
```


Comments:

E S-L F-M D-L G-L %-L

Candice has been tearing up her corrected tests and
putting the pieces in her desk, which frequently spill
out and make a mess on the floor. She cannot be disuaded
from tearing up the tests, and never remembers to take
the pieces to the garbage.
This has happened approximately once per month for the
last six months.
20% of the other student's in Candices class engage in
similar behavior.

Please rate the seriousness of the behavior on the
following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   serious

Comments:

E S-L F-L D-H G-L %-L

Leslie cracks his knuckles during class, which the teacher finds annoying.
He has done this only during spelling tests.
This behavior has occurred approximately once per month for the last 4 months.
Approximately 35% of the other kids in Leslie's class engage in similar behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9    10    11
mild                        moderate                   serious
```


Comments:

E S-L F-L D-M G-L %-L

Colleen has been working on her art project, without the teacher's permission, during seatwork time in reading and math classes. She has not tried to do this during spelling class.
This has been occuring for approximately one month, and happened once each in math and reading class.
Approximately 20% of the other student's in Colleen's class have tried to leave their seats and work in the art area without permission.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9    10    11
mild                        moderate                    serious
```
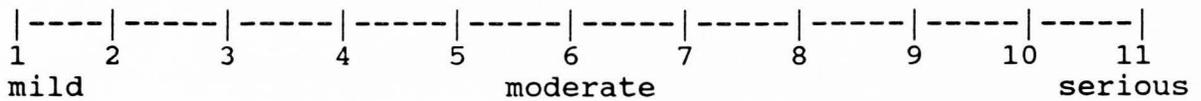

Comments:

E S-L F-L D-L G-H %-L

Billy has been caught writing and sending notes to other students during class.
This has occurred twice in the last two months, in two of Billy's three classes.
Approximately 20% of the other students write and send notes during class.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9     10    11
mild                        moderate                    serious
```
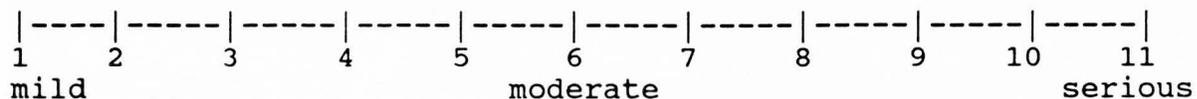

Comments:

E S-L F-L D-L G-H %-L

Anne swears in class when she makes a mistake, which is against classroom rules.
This has occurred once in the last two months, and only in one of Anne's classes.
None of Anne's classmates engage in similar behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9     10    11
mild                        moderate                   serious
```


Comments:

E S-L F-L D-L G-L %-H

Darrel leaves the classroom to go to the restroom without asking permission and paying the required number of tokens for going to the restroom during class time.
This has happened once in the last month, during his reading class.
Approximately 15% of Darrel's classmates engage in similar behavior.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
1    2     3     4     5     6     7     8     9    10    11
mild                      moderate                   serious
```

Comments:

E S-L F-L D-L G-L %-M

Andrew will not speak to anyone, and appears very
frightened when teachers ask him a question in class.
This happens 3 to 5 times per week, and has been going on
for the last 4 months.
Andrew has responded in this manner to three of his five
teachers.
Approximately 10% of the other kids in his grade exhibit
similar behavior.

Please rate the seriousness of the behavior on the
following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                    severe
```


Comments:

I S-H F-M D-M G-M %-M

Darlene cries whenever she doesn't complete her assignments on time, arrives late, etc. - she behaves as if she is afraid of being punished.
This occurs approximately 2 to 3 times per week, and has been happening since the start of the school year 3 months ago.
She has behaved this way in 2 of her 5 classes.
Approximately 11 percent of Darlene's classmates have behaved in this manner during the present school year.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                    severe
```
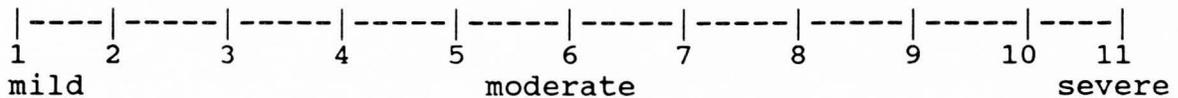
Comments:

I S-M F-M D-M G-M %-M

Karl gets stomach aches before each math and science test
- he doubles over, and looks as if he is genuinely in
pain.
This happens approximately twice per week, and has been
occuring for the last five months.
Karl gets sick before tests in two of his five classes.
Approximately 12% of the students in Karl's grade
complain of stomach or other aches before exams.

Please rate the seriousness of the behavior on the
following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   severe
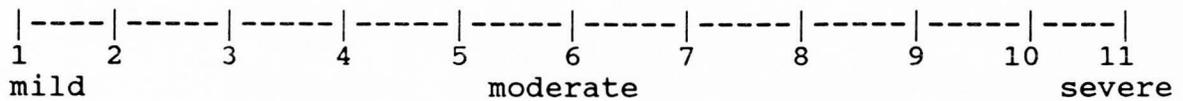
Comments:

I S-L F-M D-M G-M %-M

Nancy has days when she appears to be very insecure, and will stay very near the adult school personnel.
This happens almost every day, and has been happening consistently for the last two months.
She usually behaves this way during phy. ed., recess, lunch, and art. She does not follow her math or reading teacher around, but will keep a vigilant eye on them.
Approximately 13% of the students in Nancy's grade behave in this manner from time to time.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                       moderate                     severe
```
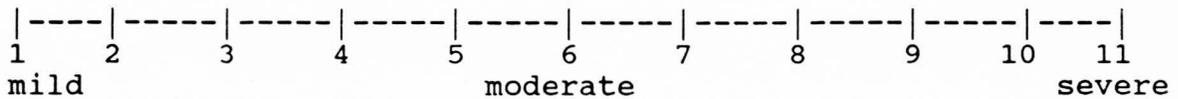

Comments:

I S-M F-H D-M G-M %-M

Ed will not play or talk with most of the students in his class. Although he seems to be able to communicate with one other student, he refuses to be involved in any group activities requiring conversation. He does not respond and seems to be trying to ignore the other students who try to be friendly to him.
This happens approximately 3 times per month, and has been occurring for approximately five months. It usually occurs on the play ground, during phy. ed., at lunch, and during science class group projects. He has not done this in math or reading.
Approximately 10% of the students in Ed's class have been observed to behave in this manner.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9    10   11
mild                        moderate                   severe
```
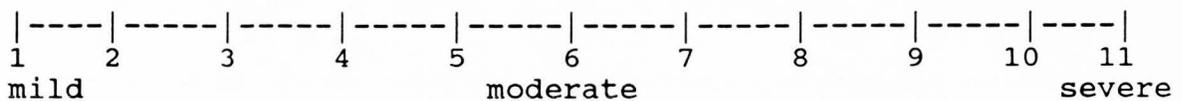

Comments:

I S-M F-L D-M G-M %-M

Sometimes, Catherine refuses to go to class; she sits down on one end of a hallway behind the inside door, where she is not easily seen.
This occurs approximately four times per week, and has been occuring for the last six months.
She has skipped out of 3 of the 6 classes she is enrolled in.
Approximately 15% of the students in Catherine's grade cut class regularly.

Please rate the seriousness of the behavior on the following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   severe
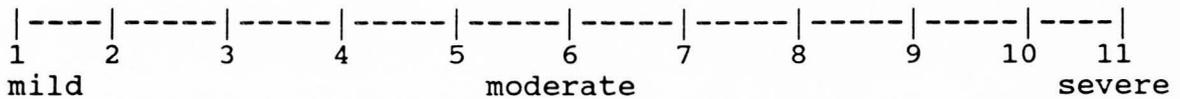
Comments:

I S-M F-M D-H G-M %-M

Paul does not seem to be paying attention to what is going on in the classroom, and sometimes gives silly answers that have nothing to do with the topic at hand when asked questions by the teacher or when other kids talk to him.
The daydreaming or fantasizing occurs approximately 3 times per week and has been occuring for 3 weeks.
He has behaved in this manner in 3 of his five classes and in the lunch room.
Approximately 12% of the other students in Paul's grade space off and daydream during school activities.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                          moderate                  severe
```


Comments:

I S-M F-M D-L G-M %-M

Gerri refuses to join in with the other students in group activities, and tends to isolate herself in a corner of the room or the playground to avoid doing the group activity.
This happens approximately 4 times per week, and has been occuring for 4 months. She has isolated herself, at one time or another, in all of her classes and on the playground.   She eats lunch with her sister.
Approximately 10% of the kids in Gerri's class isolate themselves from time to time.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                    severe
```


Comments:

I S-M F-M D-M G-H %-M

Donald has frequent panic attacks, during which he is
unable to stay in the classroom and is taken to the
school nurse.
The panic attacks occur approximately three times per
week, just before quizes, and has been occurring for
approximately 3 months.   This has only occured in
spelling and math, two of his six classes.
Approximately 10% of the other students in Donald's class
have panic attacks before quizes and tests.

Please rate the seriousness of the behavior on the
following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
mild   moderate   severe
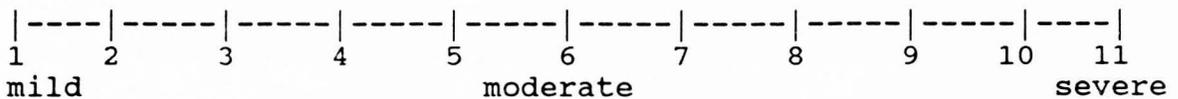
Comments:

I S-M F-M D-M G-L %-M

Cindy rocks in her chair, apparently soothing herself or providing herself with stimulation. This behavior has kept Cindy from timely and accurate completion of her assignments.
This occurs approximately 4 times per week, and has been occuring for 5 months. She has rocked herself in 3 of her 5 classes.
Approximately 25% of the other students in Cindy's grade exhibit this kind of behavior.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1     2     3     4     5     6     7     8     9    10   11
mild                          moderate                  severe
```
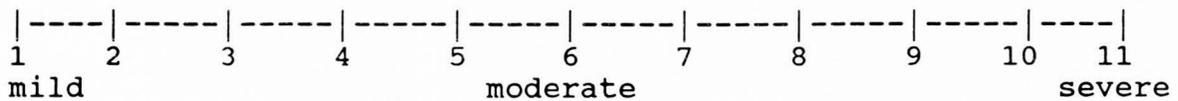

Comments:

I S-M F-M D-M G-M %-L

Kyle appears to be depressed; he cries very easily, and does not seem to be very energetic and is less interested in playing with the other students than he used to be. He looks sad and near tears approximately 5 times per week, and has been observed to behave in this manner for the last 3 months. He has behaved in this manner in 2 of his 5 classes and at lunch.
None of Kyle's classmates exhibit this kind of behavior.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                    severe
```


Comments:

I S-M F-M D-M G-M %-H

Carmen has completely isolated herself from others at
school; she does not do any work or interact with her
teacher or the other students. She sits with her head
down, and appears to be frightened or sad.
This occurs continually, and has been happening for
approximately 6 months.
She behaves this way in all her classes, in the lunch
room and on the playground.
None of the other students in Carmen's grade behave in
this manner.

Please rate the seriousness of the behavior on the
following scale:

|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1   2   3   4   5   6   7   8   9   10   11
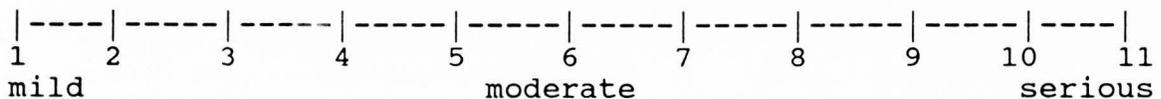mild   moderate   severe
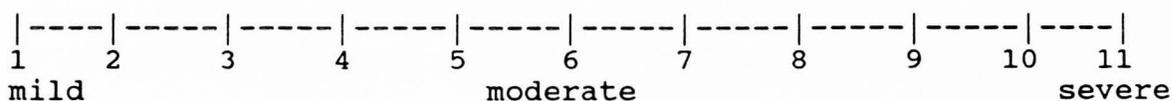
Comments:

I 8-H F-H D-H G-H %-H

Bruce has difficulty communicating. He speaks in a very quiet, almost inaudible voice. Teachers and peers frequently have to ask him to repeat what he has said, and he appears to become more uncomfortable and embarrassed when asked to repeat himself.
He does this several times per day, and has been behaving in this way for the last seven months.
He speaks in this manner during all of his class, but is quite loud when on the playground.
None of the other students in Bruce's grade behave in this manner.

Please rate the seriousness of the behavior on the following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1     2     3     4     5     6     7     8     9    10   11
mild                        moderate                    severe
```


Comments:

I S-M F-H D-H G-H %-H

Naomi appears to be uncomfortable playing games with the
other students, though she will stand by and talk with
them while they are playing - she appears to fear being
hurt or making physical contact with them.
This happens several times per day, particularly on the
playground, but she is also vigilant while walking down
the halls and when walking through her classroom. This
has been occurring for the last 8 months.
None of the other students in Naomi's class behave in
this manner.

Please rate the seriousness of the behavior on the
following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                      moderate                      severe
```

Comments:

I S-L F-H D-H G-H %-H

Approximtely 4 times each week, Curt will appear to be
talking or responding to someone (or something) that is
not actually there and which he finds frightening - the
teacher thinks he is hallucinating.   The other students
think he is very weird because of this and don't interact
with him much.
Curt has been "hallucinating" for approximately 6 months;
this behavior has occured in all of his classes as well
as during lunch and recess.
None of the other students in Curt's class behave this
way.

Please rate the seriousness of the behavior on the
following scale:


```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1     2     3     4     5     6     7     8     9     10   11
mild                          moderate                  severe
```
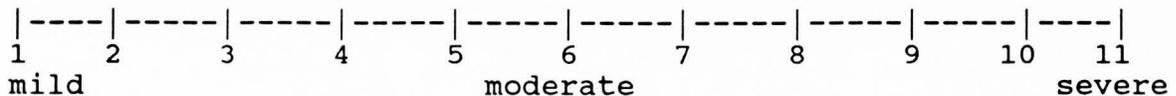

Comments:

I S-H F-M D-H G-H %-H

Kathleen has disabling panic attacks, during which she cannot catch her breath and becomes very frightened; this behavior continues to occur despite repeated interventions with the school counselor. She has panic attacks approximately once per month. This has been going on for 8 months, and has occurred in 5 of Kathleen's six classes. None of the other students in Kathleen's class behave in this manner.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9    10   11
mild                        moderate                    severe
```

Comments:

I S-H F-L D-H G-H %-H

Russell has been having a very difficult time concentrating and reports that his "thoughts get all mixed up"; he frequently says things that don't make sense, and, according to his teacher, does not seem to have "any common sense anymore". He has been behaving in this peculiar manner for approximately 2 months; this behavior has occurred in all of this classes and other school settings. None of the other students in his class behave in this manner.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1     2     3     4     5     6     7     8     9    10    11
mild                        moderate                    severe
```

Comments:

I S-H F-H D-M G-H %-H

Patrice has been truant alot lately, approximately 3 times per week. She runs away whenever the teacher makes academic demands on her, and tends to spend the time away from school by herself. She began skipping school approximately 3 weeks ago. When truant, she spends the entire day away from the school. Only one other student in Patrice's grade has been truant this year.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1     2     3     4     5     6     7     8     9    10   11
mild                        moderate                    severe
```
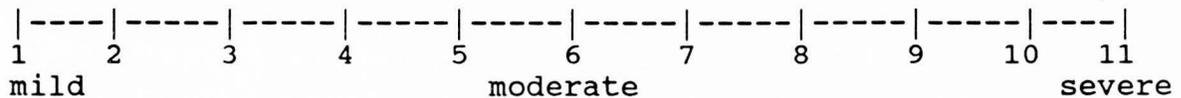
Comments:

I S-H F-H D-L G-H %-H

Brent communicates with no one. He appears frightened of others unable to speak coherently and appearing to be near tears when spoken to.
This occurs approximately 3 times per day, and has been occurring for 7 months.
He usually behaves this way in 2 if his five classes (math and science).
None of the other students in Brent's class behave in this manner.

Please rate the seriousness of the behavior on the following scale:

```
|----|-----|-----|-----|-----|-----|-----|-----|-----|----|
1    2     3     4     5     6     7     8     9     10   11
mild                        moderate                    severe
```

Comments:

I S-H F-H D-H G-M %-H

Kelly speaks with no one. He frequently puts his face in his hands, refuses to move and refuses to speak with the teacher, other students, or the principal.
This occurs almost daily, and has been happening for 8 months.
Kelly only behaves this way in art class.
None of the other kids in Kelly's class behave in this manner.

Please rate the seriousness of the problem behavior on the following scale:


```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                        moderate                   severe
```


Comments:

I S-H F-H D-H G-L %-H

Carrie is unable to follow teacher directions. She cries
when given a direction by the teacher, and acts as if her
feelings have been hurt.
This has been occuring daily for the last 6 months.
Carrie has done this in all of her classes and on the
playground.
Approximately 21 percent of Carrie's classmates engage in
similar behavior.

Please rate the seriousness of the problem behavior on
the following scale:


```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                          moderate                  severe
```


Comments:

I S-H F-H D-H G-H %-L

David is unable to maintain friendships with any of his
schoolmates; he will find someone who will play and talk
with him for a few days, then he will spend days talking
with no one. He exhibits sudden mood swings, during
which he will not respond to peers, teachers, or his
parents.
This has happened in all of David's classes, and has been
going on for more than 6 months on a daily basis.
Approximately 12 percent of the students in David's grade
do things to alienate another child.

Please rate the seriousness of the problem behavior on
the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                        moderate                   severe
```

Comments:

I S-H F-H D-H G-H %-M

Jessica appears to be very uncomfortable interacting with some of her teachers; she speaks very softly when called on to answer questions in class and whispers when asking the teacher questions. She is able to speak clearly and audibly, as evidenced in her interactions with other students.
Jessica has spoken this way 2 to 3 times in the last month, and only in her math class.
Approximately 25% of the students in Jessica's class appear uncomfortable talking to the teacher.

Please rate the seriousness of the problem behavior on the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                        moderate                    severe
```

Comments:

I S-L F-L D-L G-L %-L

Dale tried to commit suicide last week by taking a large dose of aspirin. This is the first time he has attempted suicide.
Approximately 20% of Dale's classmates have attempted suicide in the last year.

Please rate the seriousness of the problem behavior on the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                         moderate                  severe
```
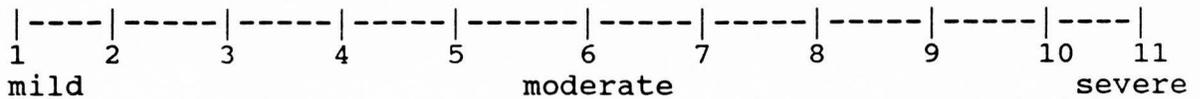
Comments:

I S-H F-L D-L G-L %-L

JoAnn does not communicate with the other students in her class.
She appears to be very uncomfortable with same age peers; she speaks softly and responds briefly to their questions and turns down offers to play or do homework with them.
This has happened a few times during the last month, and only in JoAnn's science class.
Approximately 20% of the students in JoAnn's class are very shy.

Please rate the seriousness of the problem behavior on the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                        moderate                    severe
```
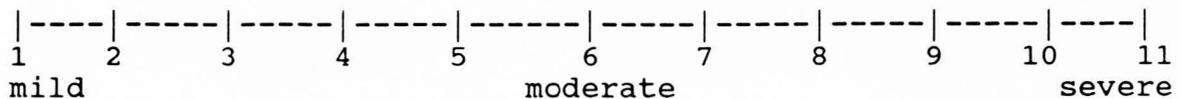
Comments:

I S-M F-L D-L G-L %-L

Daniel has a habit of picking at his clothes, pulling strings out of his sweaters, shoe laces, and other items of apparel; he seems to do this when bored. Daniel picks at his clothes several times per day, and usually in his math and science classes. Daniel has been doing this for about 3 weeks. Approximately 20% of the students in Daniel's classes do this.

Please rate the seriousness of the problem behavior on the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                         moderate                  severe
```
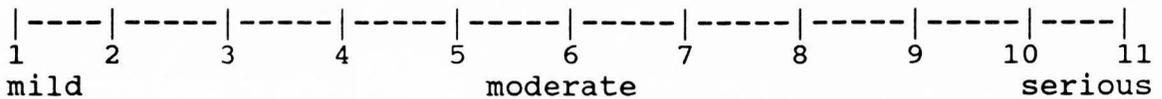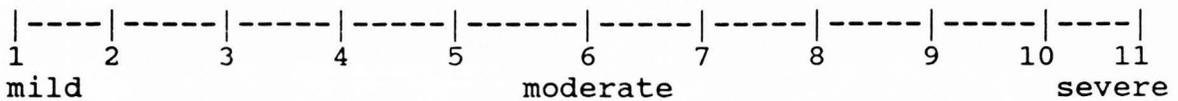
Comments:

I S-L F-H D-L G-L %-L

Charlene appears to be uncomfortable talking to her
teachers; she blushes, her eyes water, and she stutters.
This occurs several times per week, in Charlene's reading
class, and has been occuring for approximately one month.
Approximately 20% of the students in Charlene's class
appear uncomfortable when responding to the teacher's
questions.

Please rate the seriousness of the problem behavior on
the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9    10   11
mild                        moderate                    severe
```

Comments:

I S-L F-M D-L G-L %-L

When assigned to do a group project, Martin does not appear comfortable contributing to the project, and is very quiet, staying on the periphery of the activity.
Martin behaves in this manner approximately once per month, and has been doing so for approximately 7 months.
He behaves in this manner in his science class; he seems to do fine, appearing comfortable and getting involved, when in his other classes.
Approximately 20% of Martin's classmates behave in this manner.

Please rate the seriousness of the problem behavior on the following scale:


```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9    10   11
mild                          moderate                  severe
```


Comments:

I S-L F-L D-H G-L %-L

Occassionally Marie has episodes during which she appears
to be uncomfortable with teacher demands; during these
times she responds by crying quitely and appears as if
her feelings were hurt. She behaves this way every 1 to
2 months, and has done so for the last 4 to 5 months.
This has occured only during phy. ed. class.
Approximately 25% of Marie's classmates exhibit this type
of behavior.

Please rate the seriousness of the problem behavior on
the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1     2     3     4     5      6     7     8     9     10   11
mild                          moderate                  severe
```
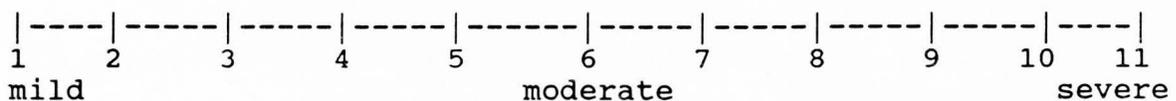
Comments:

I S-L F-L D-M G-L %-L

Last week, Charles became upset when his teacher asked him to turn in his overdue homework. He refused to speak or move for an entire morning - he appeared angry and upset, but would not, and as yet, has not, talked about what specifically he was angry about. This was the only time that Charles has behaved in this manner. Approximately 20% of Charles classmates have gone through periods where they refuse to speak.

Please rate the seriousness of the problem behavior on the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9     10   11
mild                          moderate                  severe
```
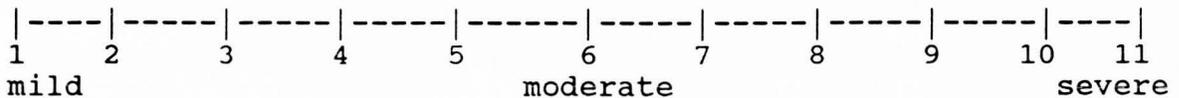
Comments:

I S-L F-L D-L G-M %-L

Kevin has skipped out of his math class twice in the last month and a half. This is the only class from which Kevin has been absent. None of Kevin's classmates have skipped out of class during the present school year.

Please rate the seriousness of the problem behavior on the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1    2     3     4     5      6     7     8     9    10   11
mild                             moderate                severe
```

Comments:

I S-L F-L D-L G-L %-H

Allissa recently announced that she is no longer going to participate in phy. ed. class, and refuses to tell anyone why. When asked, she refuses to talk until the subject is changed, except to repeat that she is not going any more. This happened one week ago, and she has missed one phy. ed. class period. Approximately 7% of Allissa's classmates have refused to participate in phy. ed.

Please rate the seriousness of the problem behavior on the following scale:

```
|----|-----|-----|-----|------|-----|-----|-----|-----|----|
1     2     3     4     5      6     7     8     9    10    11
mild                          moderate                    severe
```
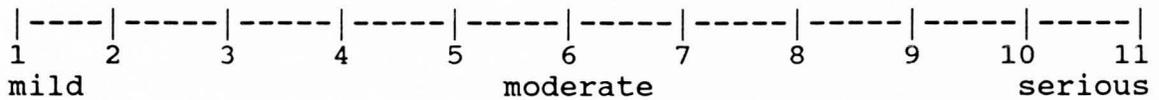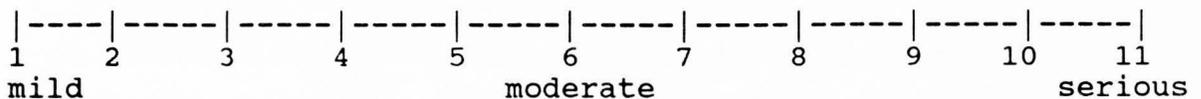
Comments:

I S-L F-L D-L G-L %-M

Appendix F

Instructions to Study Participants

Instructions

The purpose of this research is to clarify some of the underlying rules by which assessment team members assess problem behaviors in school children. Your responses will be used to assist the staff at the Technology Division of the Developmental Center for Handicapped Persons at Utah State University to develop an expert system for assisting school personnel in classifying students with behavioral disorders.

Please read each vignette and provide your rating of the seriousness of the behavior on the eleven point scale below each vignette. The scale is set up such that 1 indicates a mild problem, 6 indicates a moderate problem, and 11 indicates a problem that would be considered quite serious.

I realized that psychologists and teachers utilize much more information than that presented in the vignettes to determine whether a problem behavior is serious enough for a BD classification. However, in this study we are concerned only with the impact of the student's behavior itself on professional judgements regarding the seriousness of a behavior problem.

There may be some instances when you would want to know more about the behavior itself before making a judgement. When that occurs, please 1) circle the number you think is the most appropriate based on the information provided, the 2) write on the lower half of the page those additional things you would like to have known about in determining the seriousness of the problem behavior. At any time, feel free to write your thoughts, comments, criticisms, and so forth, on the lower half of the pages.

Please remember to fill out the Consultation Invoice; be sure to provide your social security number and the address you would like the honorarium check sent to. Also, please remember to sign your name at the bottom of the form behind "CONSULTANT".

If you have any questions, please feel free to call me (Sheila) at any time. My phone numbers are:

Technology Division (USU) 750-3734
Home   752-0198

I will be contacting you on the 25th to make arrangements to pick up the materials.
Thank you so much for your participation.

Appendix G

Semi-Structured Interview Form

Interview Questions

Case # _____

Interviewee _____

Of the information provided in this case, which pieces of information were most important in helping you determine your rating in this case?

Rank order: _____ Frequency
              _____ Severity
              _____ Duration
              _____ Generality
              _____ Percent of Peers/Similar Behavior

Why did your rank (the one of the above factors which was varied in the case) where you did? How did it enter into your decision?

Comment on your approach to determining the seriousness of this problem behavior.

Appendix H

Scatterplots

Correlations Between Subjects' Ratings on the

Problem Behavior Descriptions

Scatterplot of Subject 1 vs Subject 2
Ratings on Problem Behavior Descriptions

$y = -0.48637 + 0.89146x$ $R^2 = 0.599$
$y = -0.11400 + 1.2866x - 0.15183x^2 + 1.0915e\text{-}2x^3$ $R^2 = 0.625$
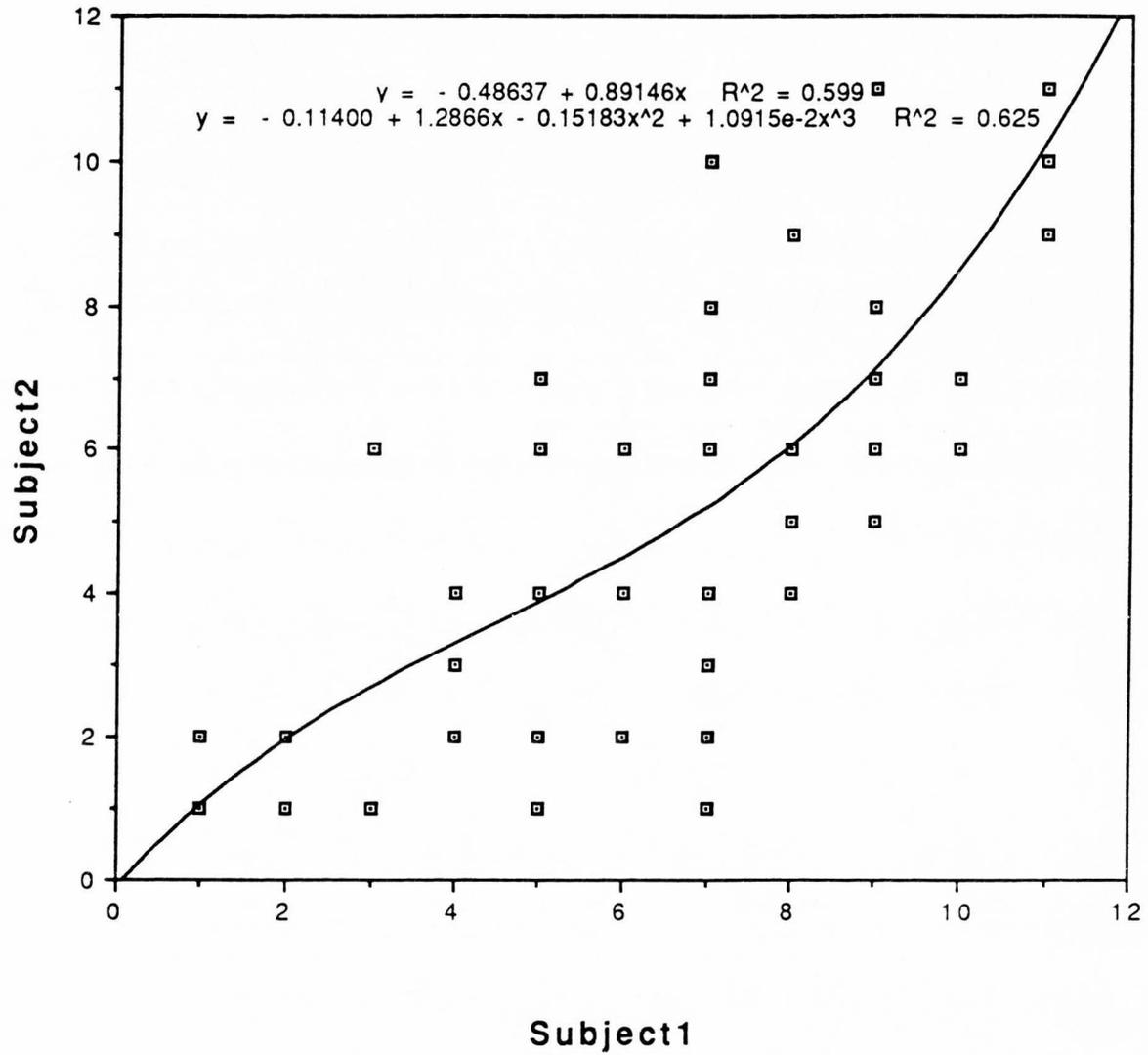
Subject2

Subject1

Scatterplot of Subject 1 vs Subject 3
Ratings on Problem Behavior Descriptions

$y = 1.0117 + 0.20170x - 1.1605\text{e-}2x^2 + 5.6521\text{e-}3x^3 \quad R^2 = 0.613$
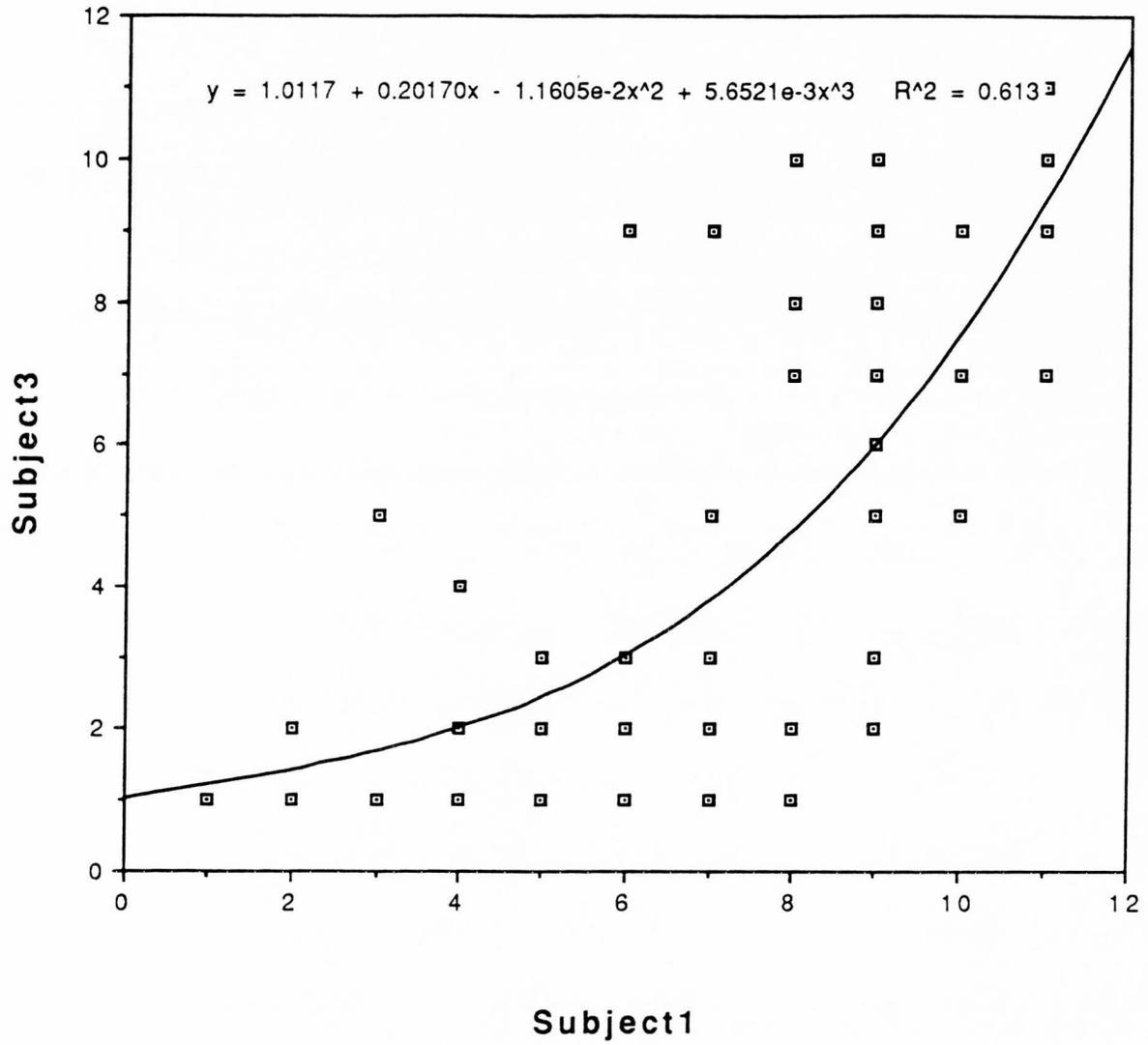
Scatterplot of Subject 1 vs Subject 4
Ratings on Problem Behavior Descriptions
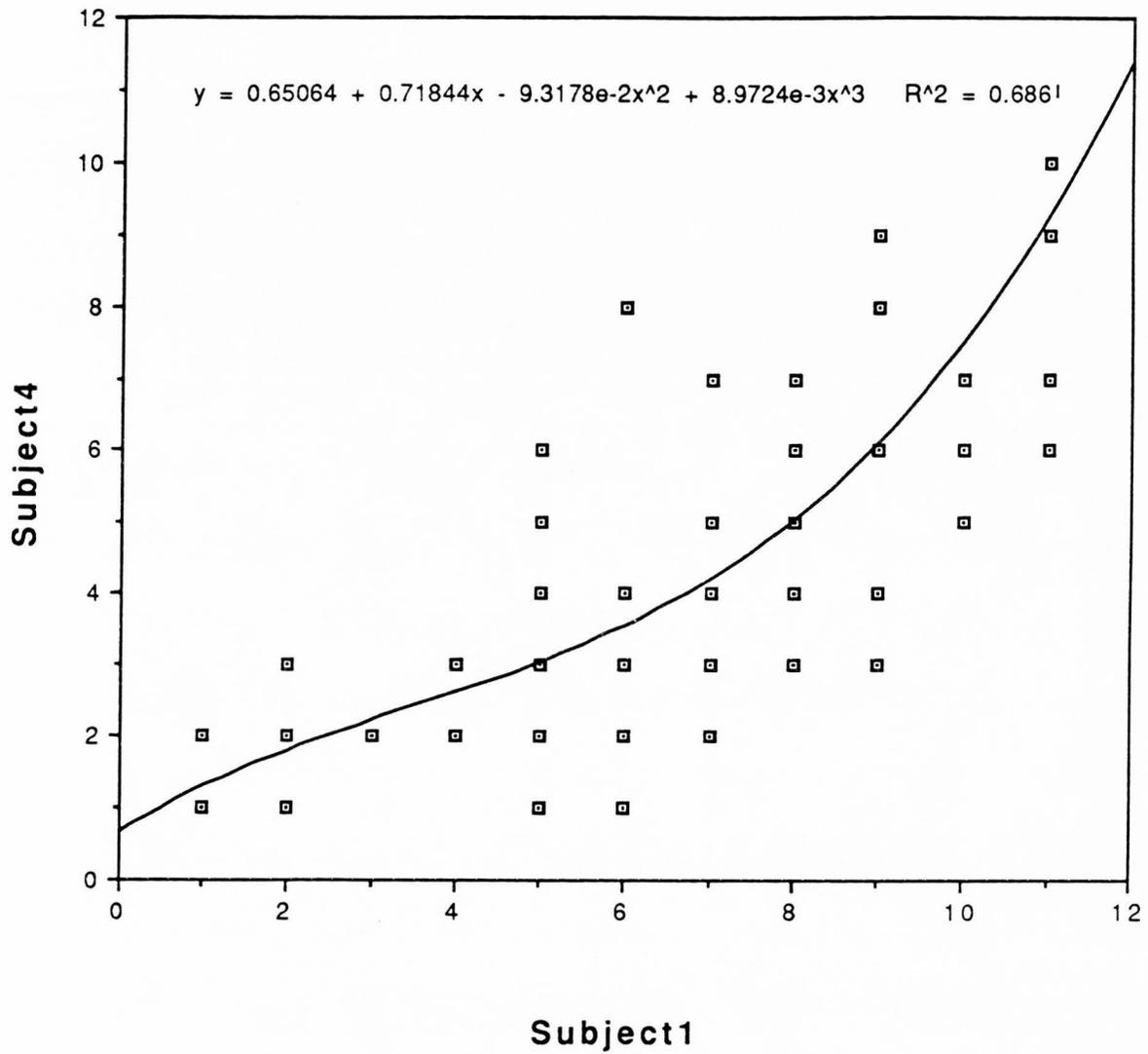
Scatterplot of Subject 1 by Subject 5
Ratings on Problem Behavior Descriptions

$y = -0.19252 + 1.0953x - 8.7595e\text{-}2x^2 + 6.4420e\text{-}3x^3 \quad R^2 = 0.600$
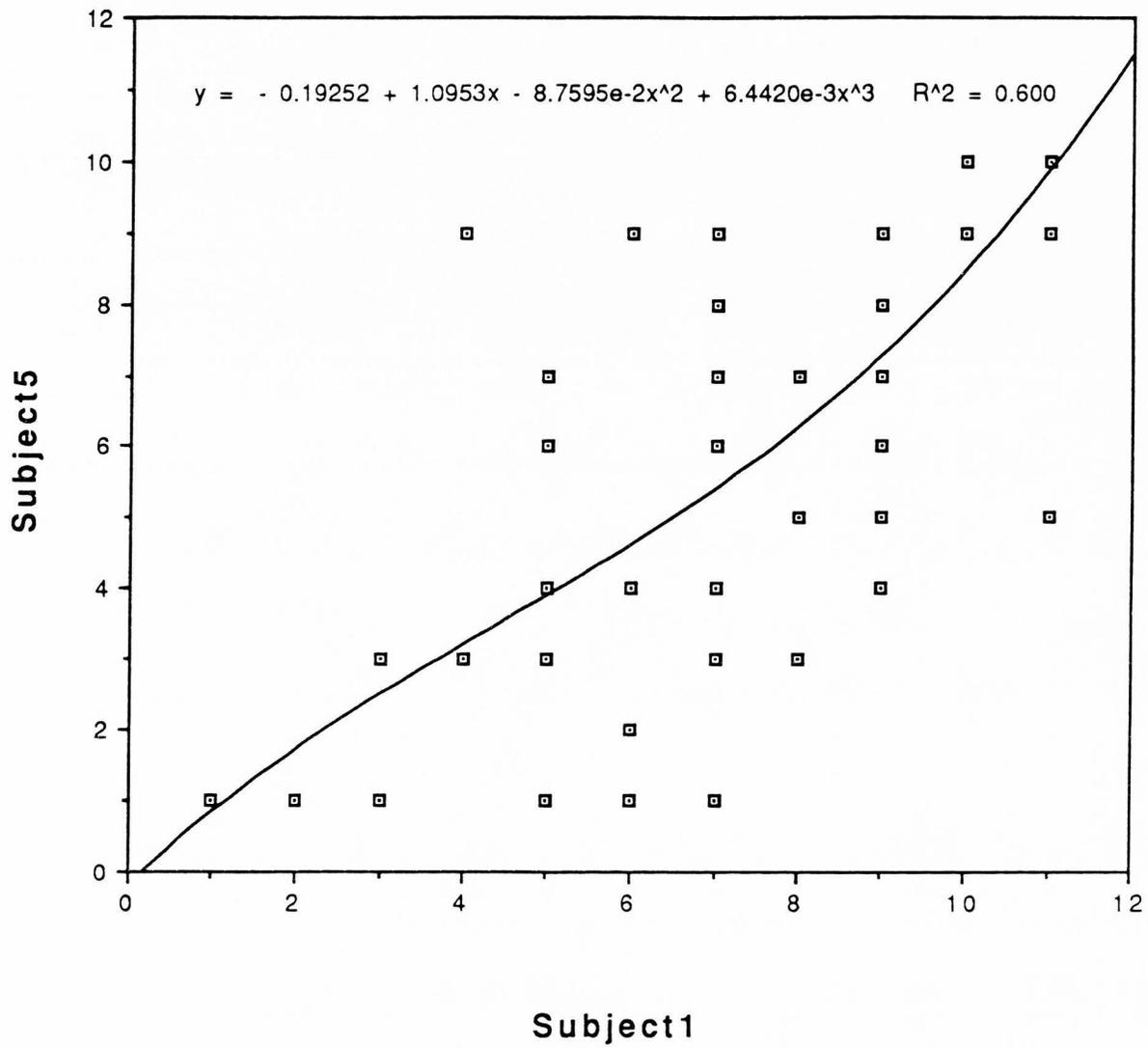
Subject5

Subject1

**Scatterplot of Subject 1 by Subject 6
Ratings on Problem Behavior Descriptions**

$y = 0.61888 + 0.34993x - 3.1762e-2x^2 + 7.1659e-3x^3 \quad R^2 = 0.737$[1]
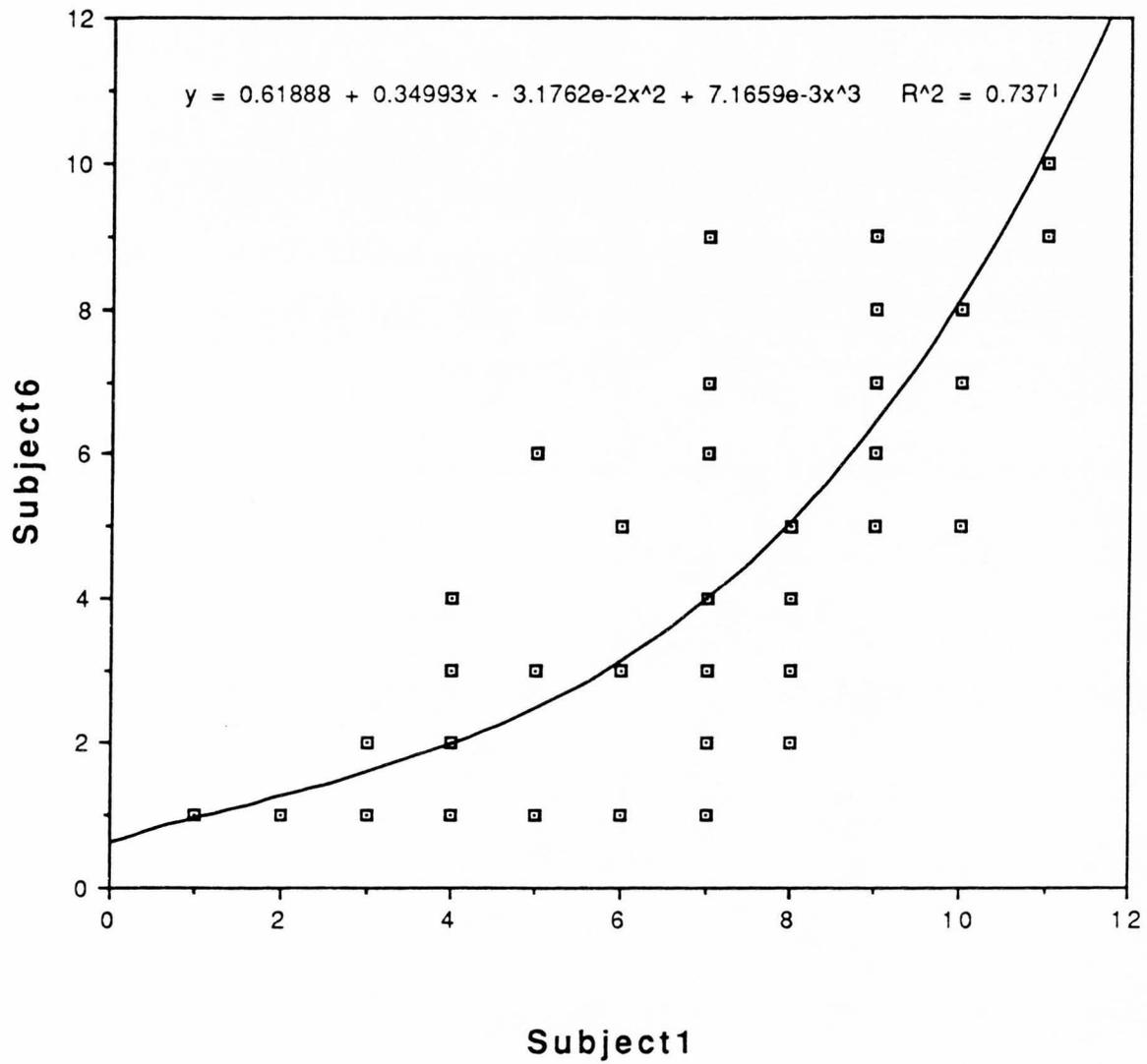
Scatterplot of Subject 2 by Subject 3
Ratings on Problem Behavior Descriptions

y = 0.35852 + 0.66703x + 2.0486e-2x^2 - 1.1371e-3x^3   R^2 = 0.582!
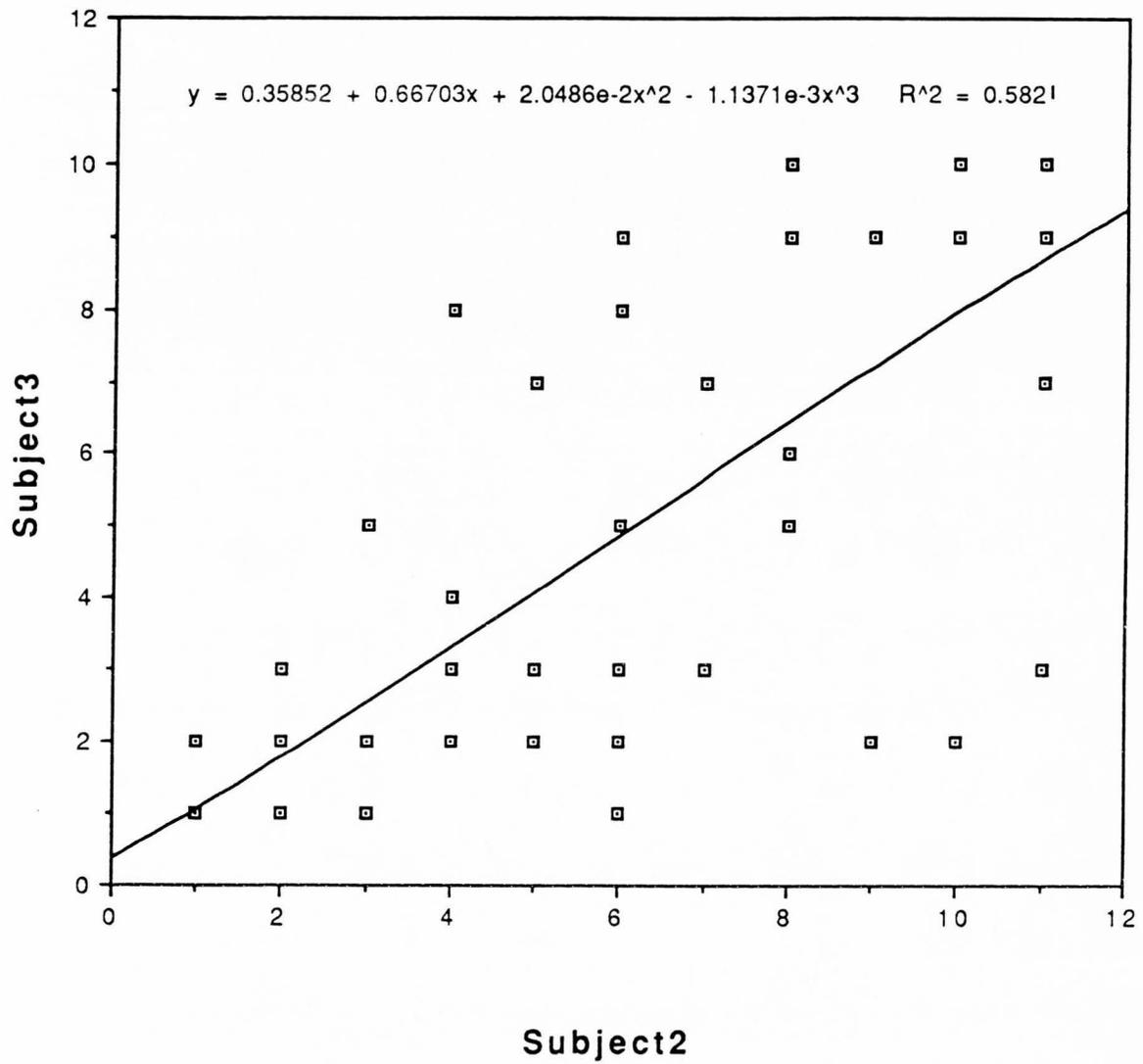
Scatterplot of Subject 2 vs Subject 4
Ratings on Problem Behavior Descriptions

$y = 1.7392 - 0.10527x + 0.14643x^2 - 6.9471\text{e-}3x^3 \quad R^2 = 0.767$
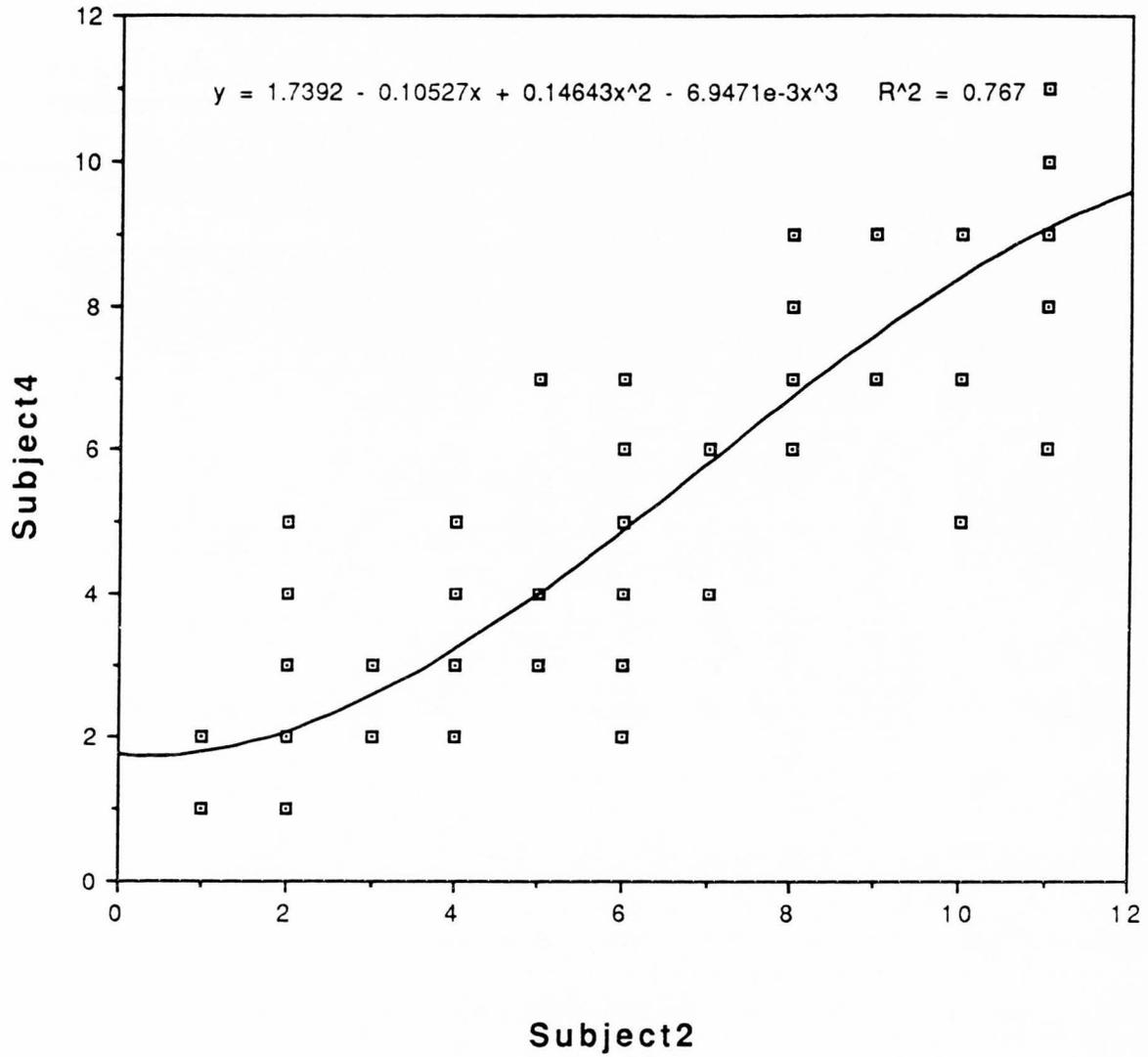
Scatterplot of Subject 2 vs Subject 5
Ratings on Problem Behavior Descriptions

$y = -0.43634 + 1.7911x - 0.12431x^2 + 3.5803e\text{-}3x^3 \quad R^2 = 0.600$
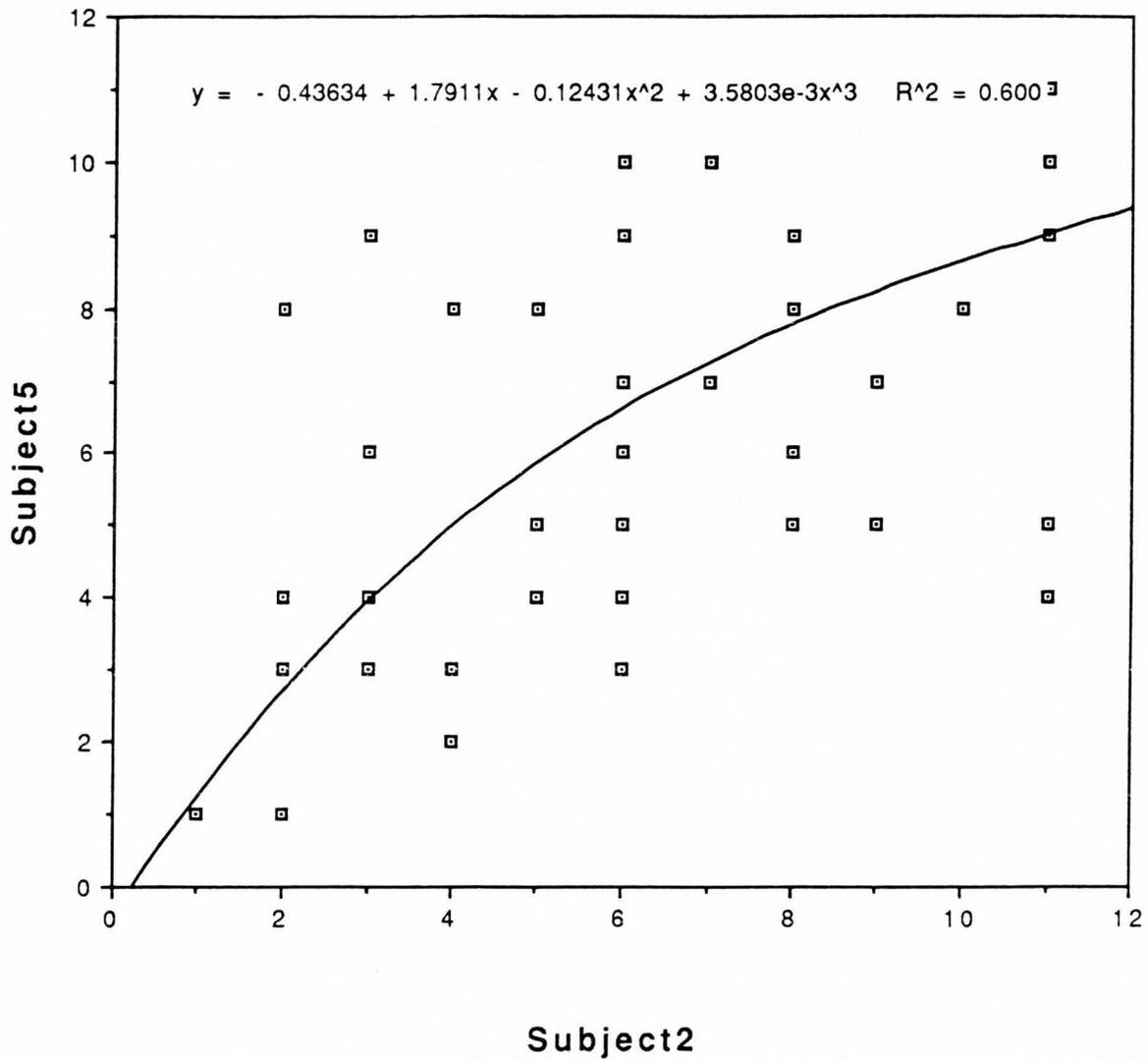
**Scatterplot of Subject 2 vs Subject 6
Ratings on Problem Behavior Descriptions**

$y = 1.3352 - 0.29639x + 0.22559x^2 - 1.2085e-2x^3 \quad R^2 = 0.717$
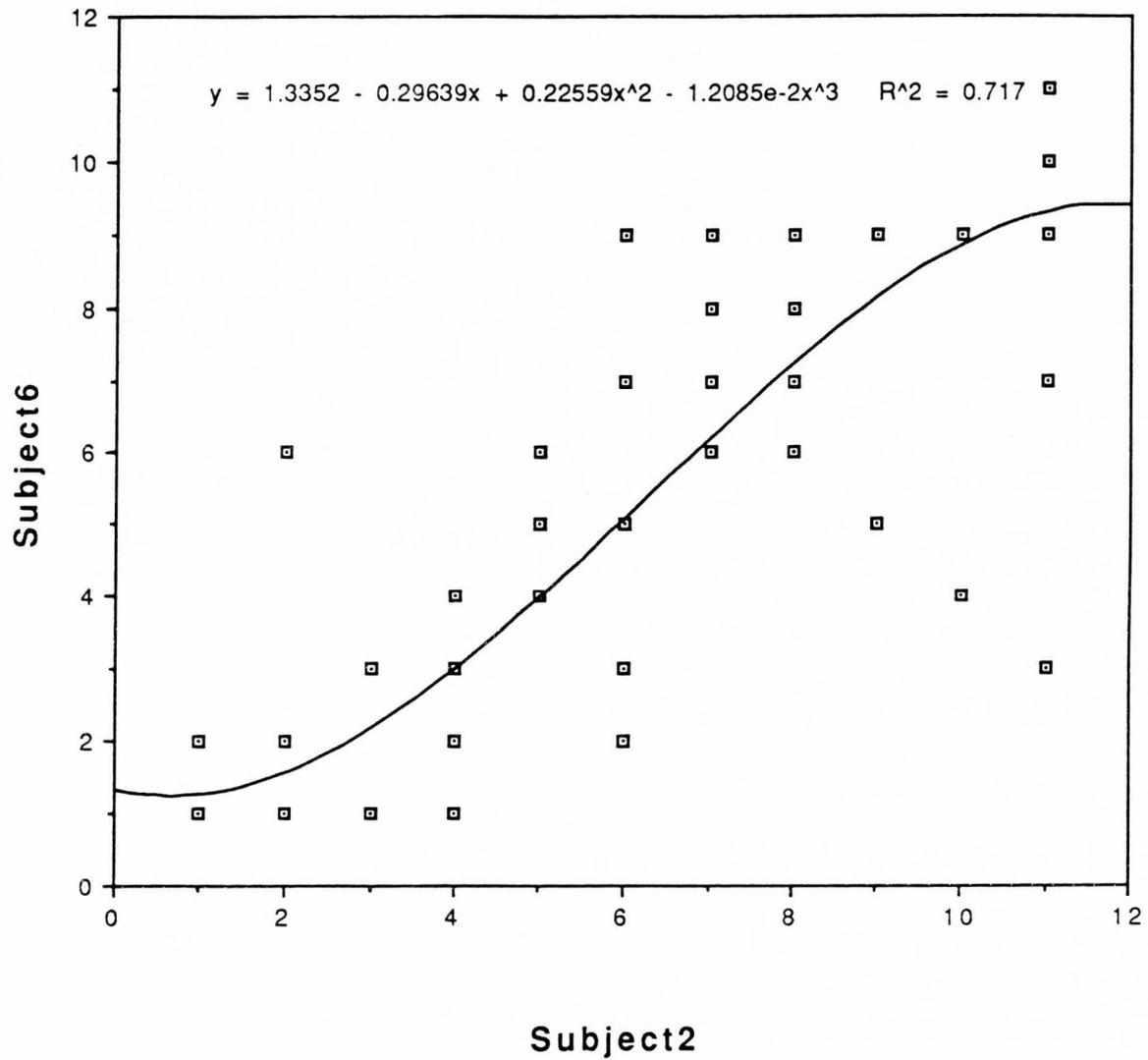
Subject6

Subject2

Scatterplot of Subject 3 vs Subject 4
Ratings on Problem Behavior Descriptions

$y = 0.60403 + 1.9015x - 0.29559x^2 + 1.8913e-2x^3$ $R^2 = 0.622$
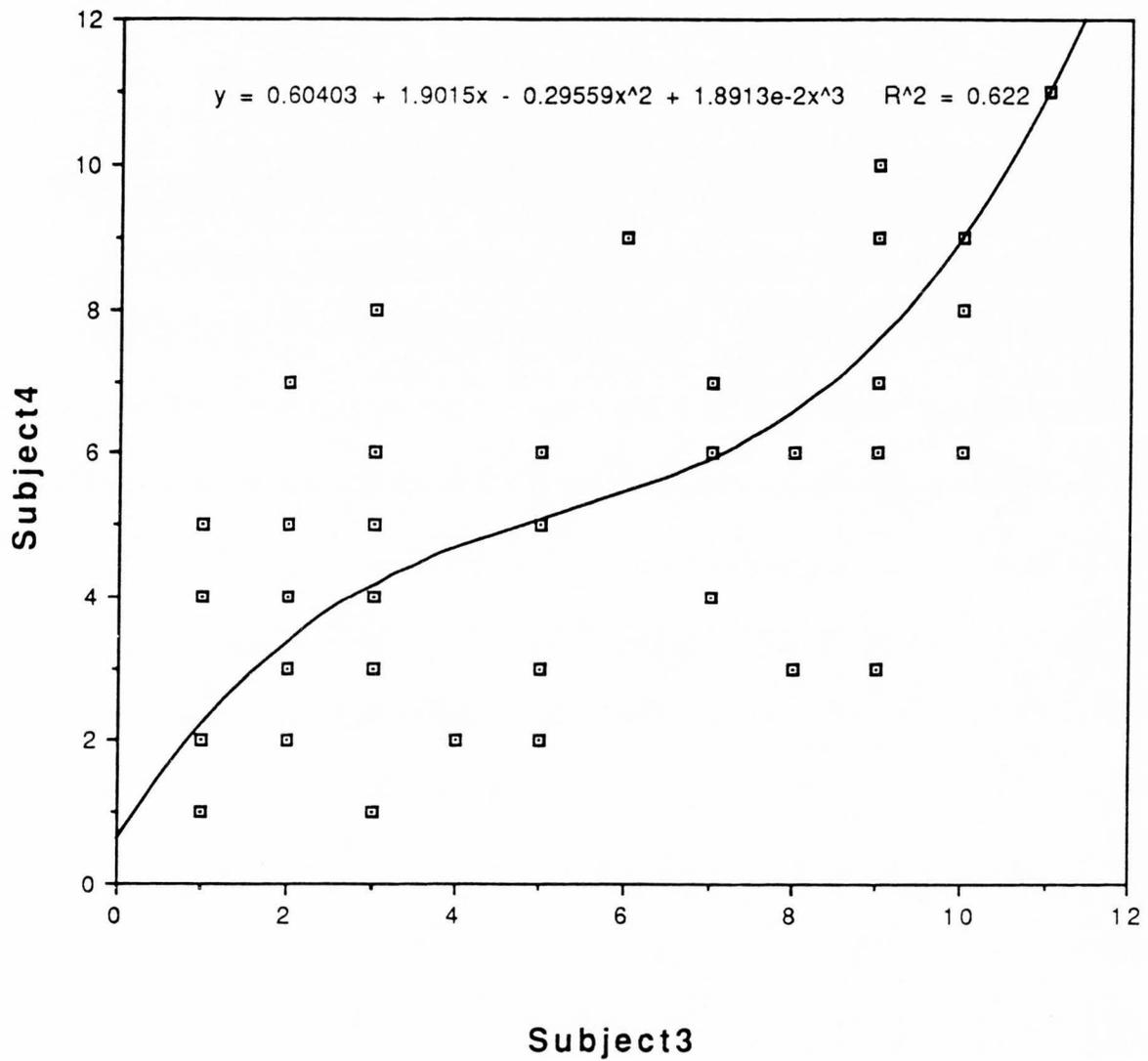
Subject4

Subject3

Scatterplot of Subject 3 vs Subject 5 Ratings on Problem Behavior Descriptions

$y = 0.35996 + 2.5934x - 0.34539x^2 + 1.7112e\text{-}2x^3 \quad R^2 = 0.447$
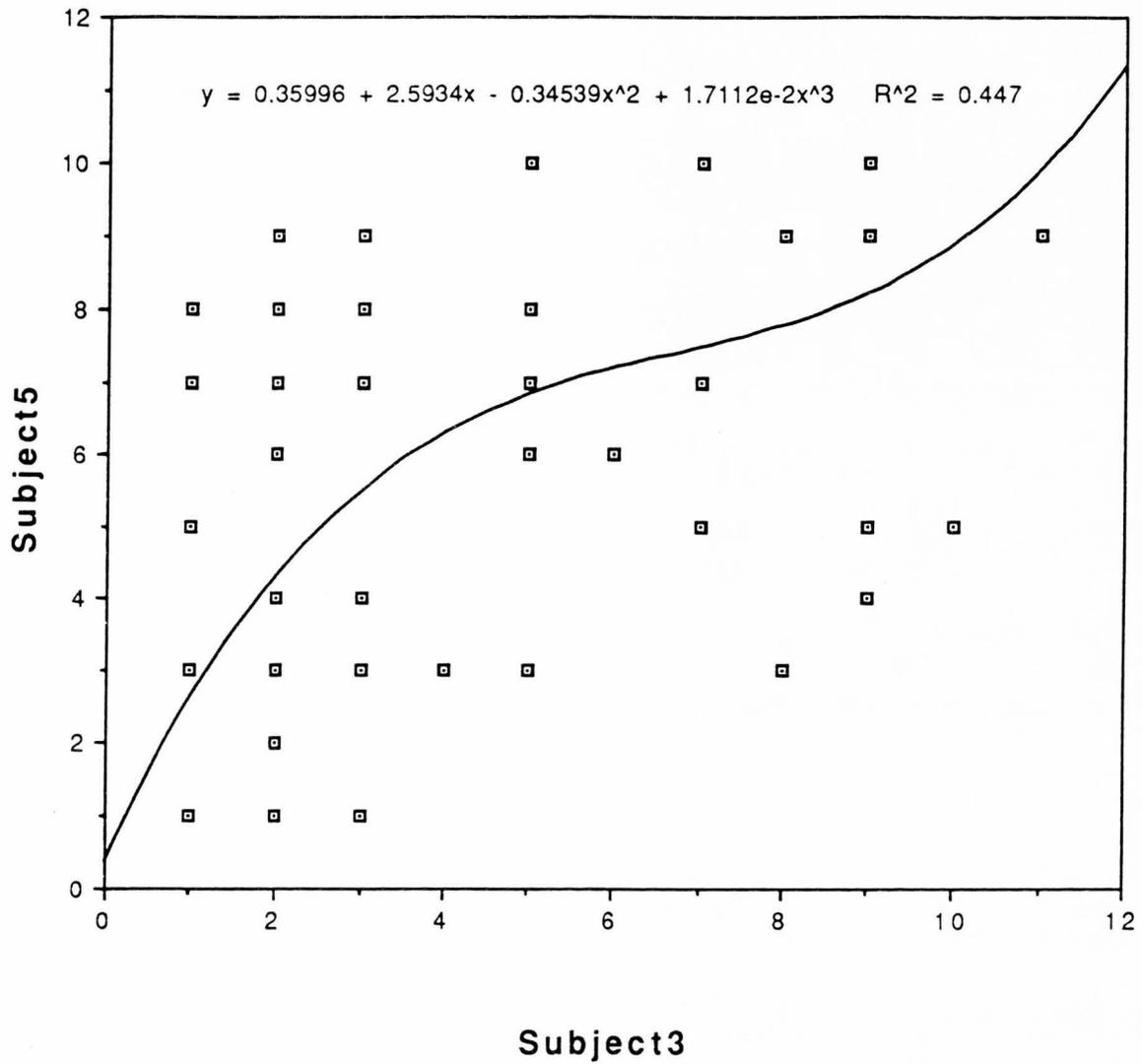
Subject5

Subject3

Scatterplot of Subject 3 vs Subject 6
Ratings on Problem Behavior Descriptions

$y = 3.6841e\text{-}2 + 1.8169x - 0.22154x^2 + 1.3544e\text{-}2x^3 \quad R^2 = 0.673$
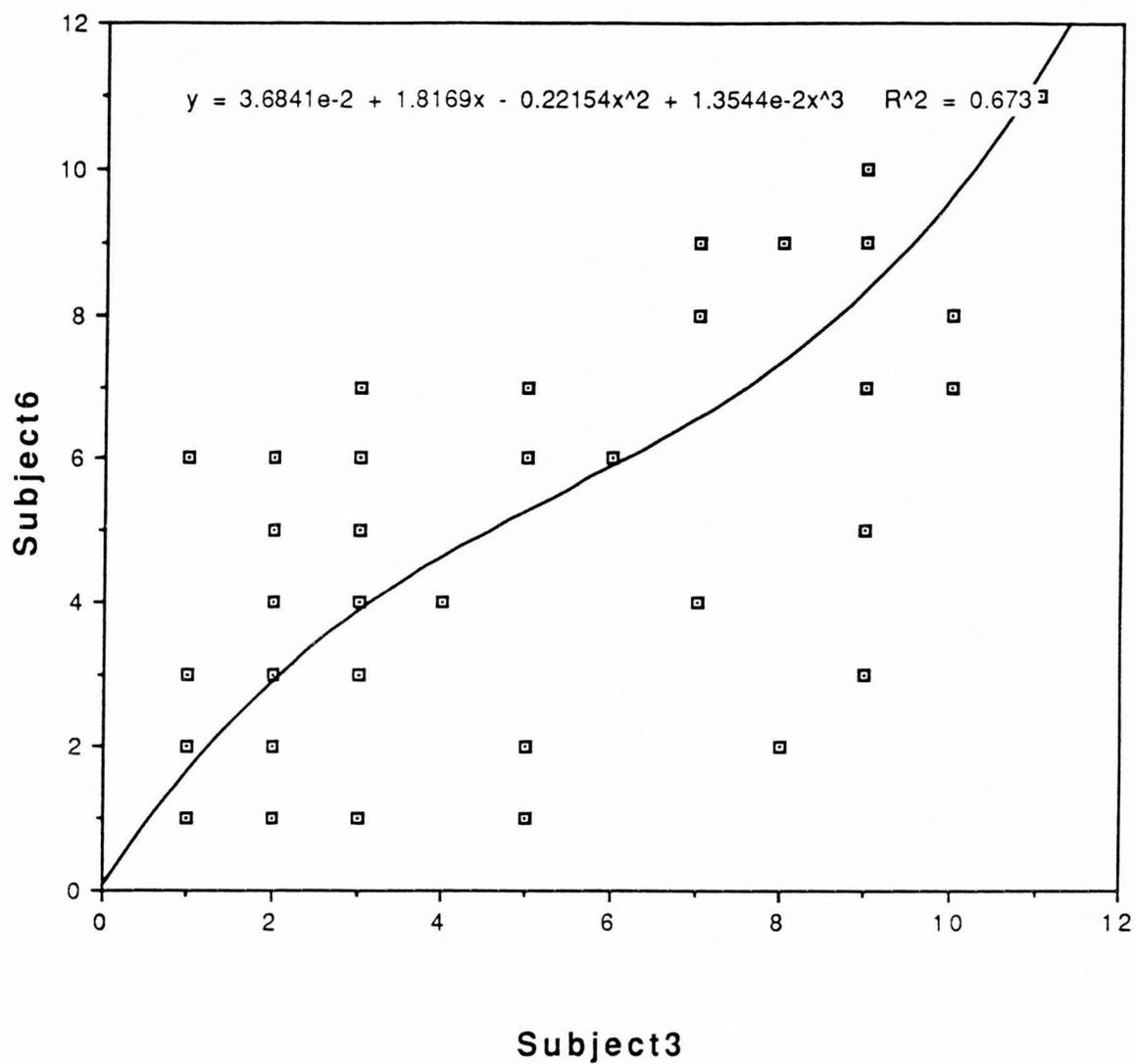
Scatterplot of Subject 4 vs Subject 5
Ratings on Problem Behavior Descriptions

$y = -3.2679 + 4.0357x - 0.51569x^2 + 2.3513e\text{-}2x^3 \quad R^2 = 0.643$
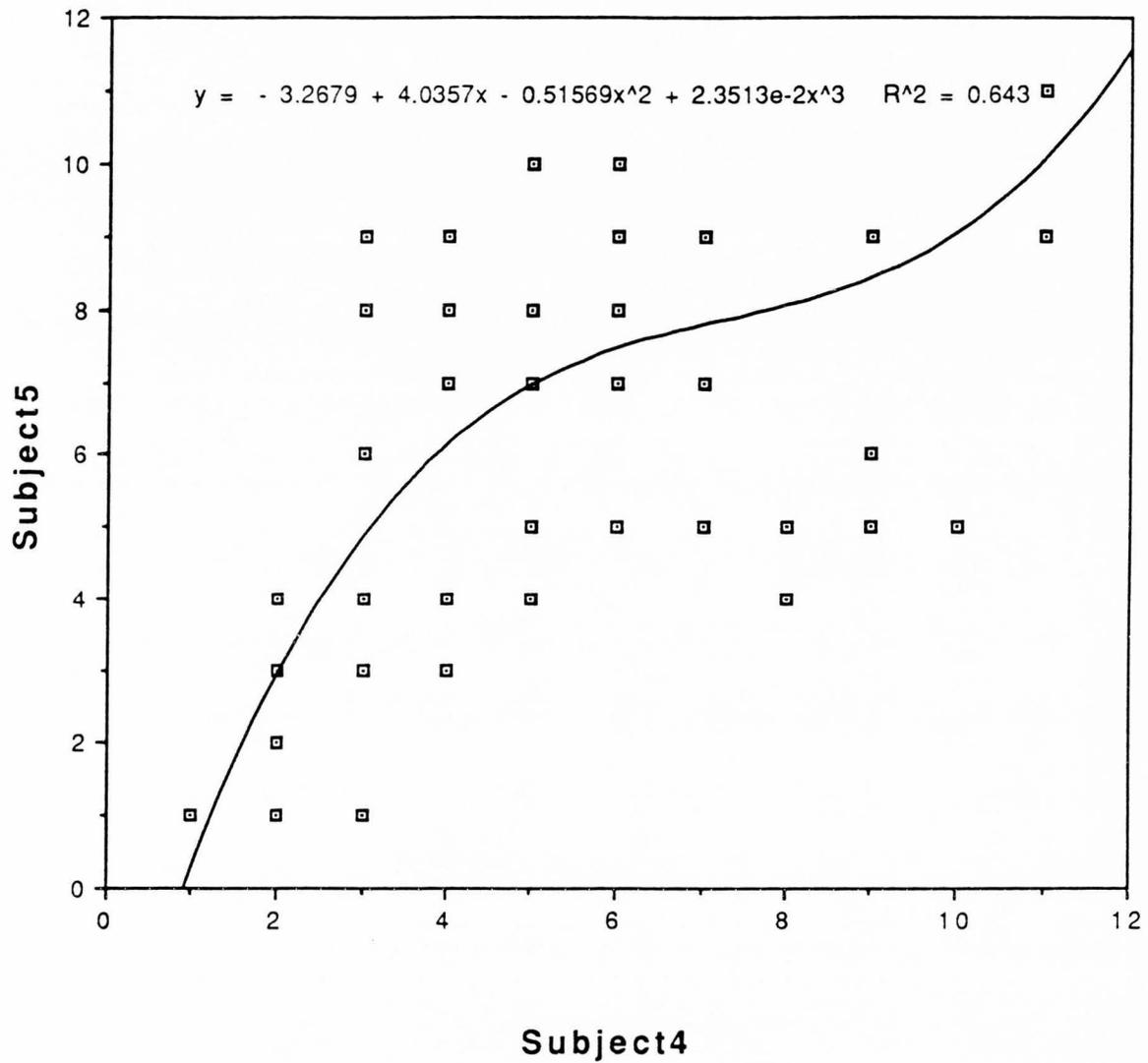
Subject5

Subject4

Scatterplot of Subject 4 vs Subject 6
Ratings on Problem Behavior Descriptions

$y = -1.3302 + 1.9720x - 0.17210x^2 + 8.4934e\text{-}3x^3 \quad R^2 = 0.711$
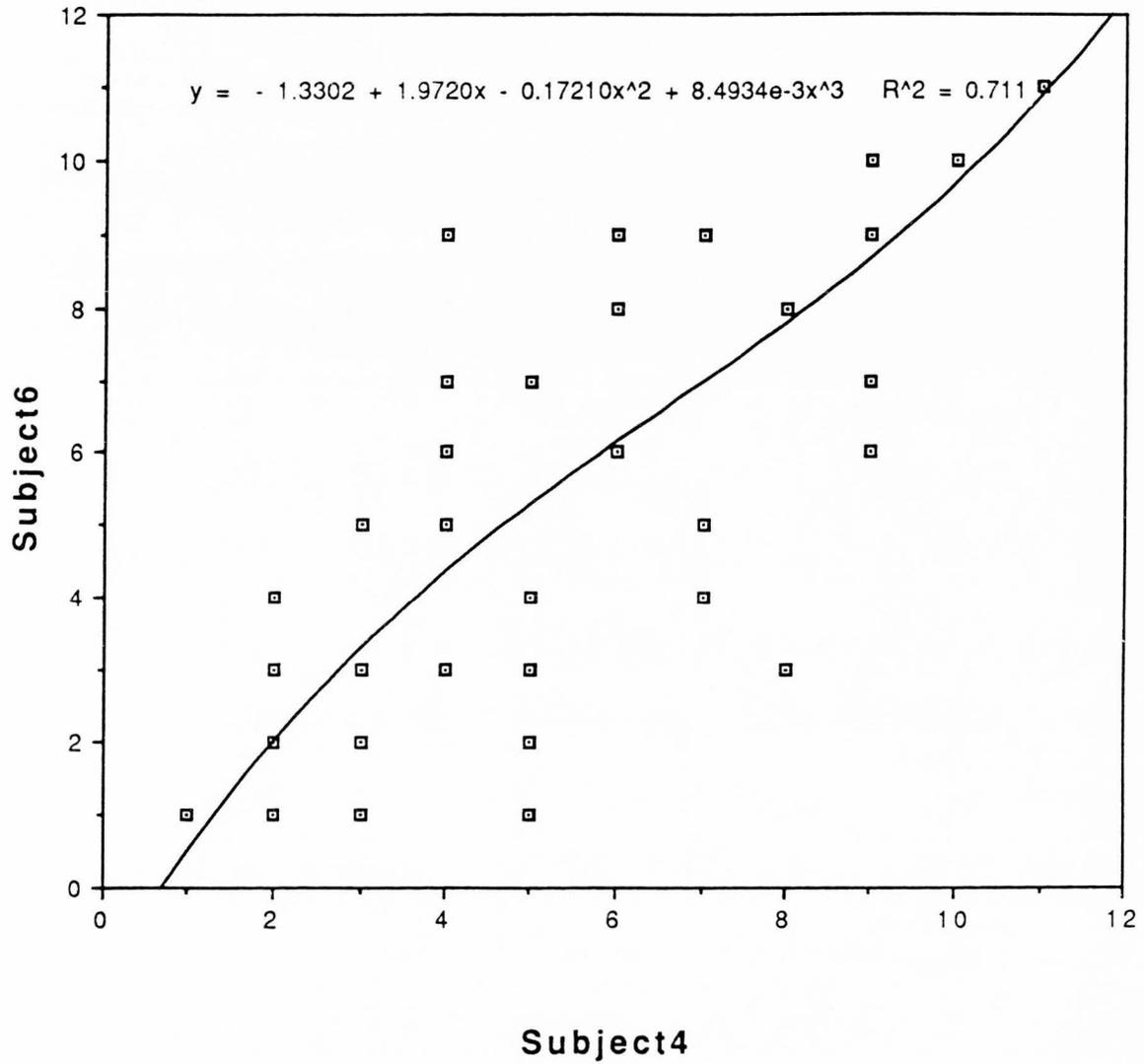
Subject6

Subject4

**Scatterplot of Subject 5 vs Subject 6
Ratings on Problem Behavior Descriptions**

$y = -1.0219 + 2.0764x - 0.27557x^2 + 1.6095e{-}2x^3 \quad R^2 = 0.660$
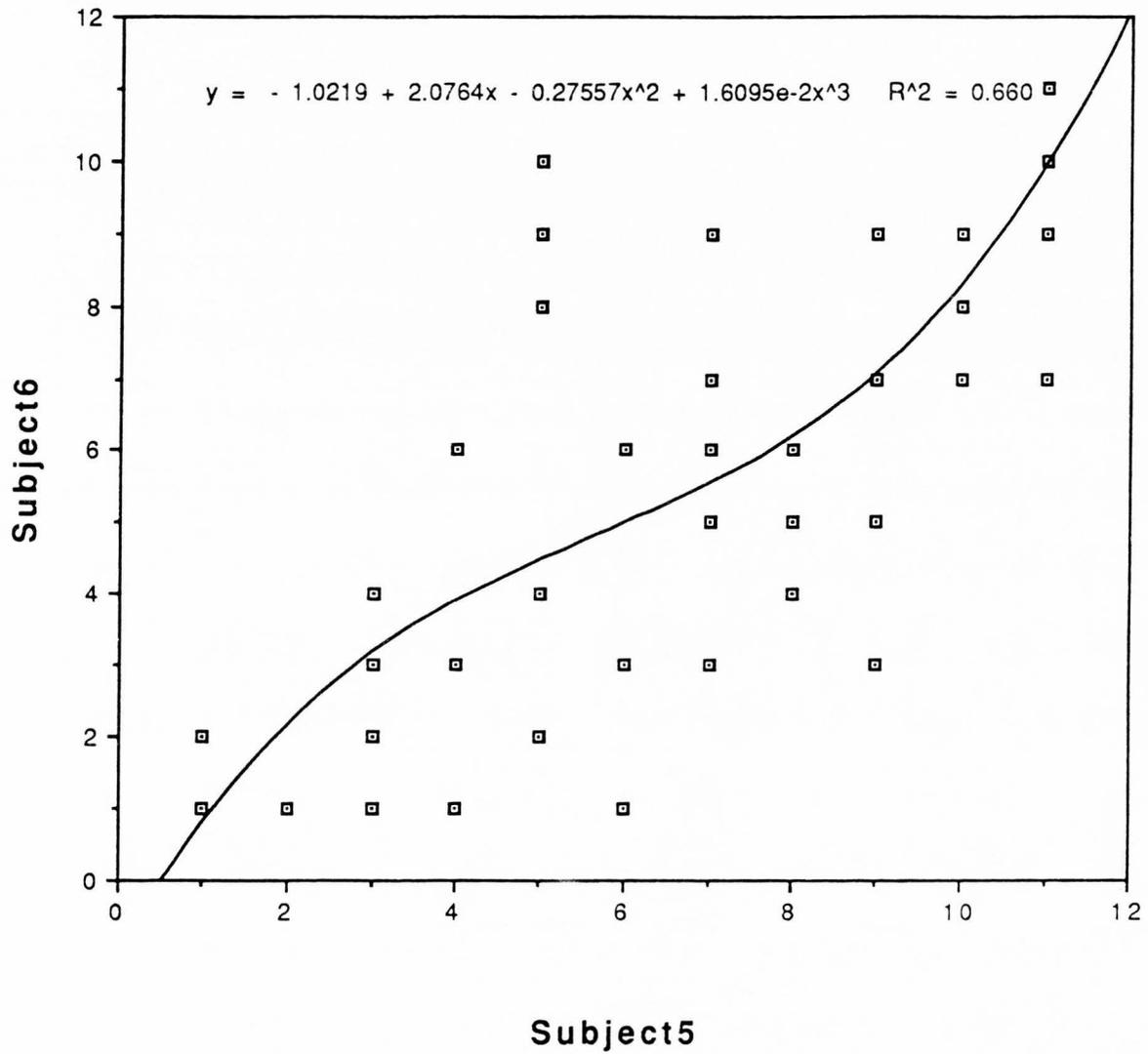
Appendix I

Analysis of Variance Tables

Analysis of Variance for Severity

	df	F	Sig. Level
Ex/In	1	0.13	.724
Severity Level	2	112.94	<.001
Other 4 Factors	2	38.46	<.001
Ex/In x Severity	2	3.82	.026
Ex/In x Other 4 Factors	2	16.50	<.001
Severity by Other 4 Fac.	4	6.14	<.001
Ex/In by Sever. x Other 4	4	6.90	<.001
Error	85	2.65	

Analysis of Variance for Frequency

	df	F	Sig. Level
External/Internal	1	9.94	.002
Frequency	2	25.37	<.001
Other 4 Factors	2	299.57	<.001
Ex/In x Frequency	2	6.81	.002
Ex/In x Other 4 Factors	2	3.65	.030
Frequency x Other 4 Fac.	4	3.81	.007
Ex/In x Frequency x O4F	4	4.20	.004
Error	85	1.72	

Analysis of Variance for Duration

	df	F	Sig. Level
Externalized/Internalized	1	2.90	.092
Level of Duration	2	5.06	.008
Level of Other Four Factors	2	197.63	<.001
Ex/In x Duration	2	2.66	.076
Ex/In x Other 4 Factors	2	4.76	.011
Duration x Other 4 Factors	4	6.20	<.001
Ex/In x Duration x Other 4 Fac.	4	3.98	.005
Error	85	2.33	

Analysis of Variance for Generality

	df	F	Sig. Level
Externalized/Internalized	1	13.31	<.001
Level of Generality	2	4.95	.009
Level of Other Four Factors	2	127.92	<.001
Ex/In x Generality	2	.24	.790
Ex/In x Other Four Factors	2	10.94	<.001
Generality x Other Four Factors	4	2.29	.067
Ex/In x Generality x Other 4 Fac.	4	3.24	.026
Error	80	2.21	

Analysis of Variance for Percent of Peers

	df	F	Sig. Level
Externalized/Internalized	1	2.12	.149
Level of Percent of Peers	2	14.05	<.001
Level of Other Four Factors	2	192.38	<.001
Ex/In x Percent of Peers	2	10.56	<.001
Ex/In x Other Four Factors	2	21.55	<.001
% of Peers by Other 4 Factors	4	1.64	.171
Ex/In x % of Peers x Other 4 Fac.	4	3.97	.005
Error	85	2.12	

Appendix J

Means and Fisher's LSD Comparisons -

Severity ANOVA

Appendix J

Means and Fisher's LSD Comparison

Severity ANOVA

Main Effect for Externalized/Internalized
    F=0.13   p=.724

                        <u>Mean</u>
        Externalized   5.78
        Internalized   5.67

Fisher's LSD   difference=.11   nonsignificant

Main Effect for Levels of Severity
    F=112.94   p,.001

                        <u>Mean</u>
Severity High   8.94
Severity Moderate   4.83
Severity Low   3.39

Fisher's LSD
                        <u>difference</u>
High vs. Moderate   4.11   significant
Moderate vs. Low   1.44   significant
High vs. Low   5.56   significant

Main Effect for Levels of Other Four Factors
    F=38.46   p<.001

                        <u>Mean</u>
Other 4 High   7.19
Other 4 Moderate   6.08
Other 4 Low   3.89

Fisher's LSD Comparisons
                        <u>difference</u>
High vs. Moderate   1.11   significant
High vs. Low   3.31   significant
Moderate vs. Low   2.19   significant

Interaction between Externalized/Internalized and Level of Severity
    F=3.82   p=.026

                        <u>Mean</u>
Exter/High Severity   9.06
Exter/Mod. Severity   5.39
Exter/Low Severity   2.89
Inter/High Severity   8.83
Inter/Mod. Severity   4.28
Inter/Low Severity   3.89

Fisher's LSD

		difference	
Exter/High vs. Exter/Mod		3.67	significant
Exter/High vs. Exter/Low		6.17	significant
Exter/High vs. Inter/High		0.22	nonsignificant
Exter/Mod vs. Exter/Low		2.50	significant
Exter/Mod vs. Inter/Mod		1.11	significant
Exter/Low vs. Inter/Low		-1.00	nonsignificant
Inter/High vs. Inter/Mod		4.56	significant
Inter/High vs. Inter/Low		4.94	significant
Inter/Mod vs. Inter/Low		0.39	nonsignificant

Interaction between Externalized/Internalized and Level of Other Four Factors
$F=16.49$ $p<.001$

	Mean
Exter/Other 4 High	7.56
Exter/Other 4 Moderate	7.06
Exter/Other 4 Low	2.72
Inter/Other 4 High	6.83
Inter/Other 4 Moderate	5.11
Inter/Other 4 Low	5.06

Fisher's LSD Comparisons

	difference	
Exter/High vs. Exter/Mod	0.50	nonsignificant
Exter/High vs. Exter/Low	4.83	significant
Exter/High vs Inter/High	0.72	nonsignificant
Exter/Mod vs. Exter/Low	4.33	significant
Exter/Mod vs. Inter/Mod	1.94	significant
Exter/Low vs. Inter/Low	-2.33	significant
Inter/High vs. Inter/Mod	1.72	significant
Inter/High vs. Inter/Low	1.78	significant
Inter/Mod vs. Inter/Low	0.54	nonsignificant

Interaction between Level of Severity and Level of Other Four Factors
$F=6.14$ $p<.001$

	Mean
Severity H, Other 4 H	10.17
Severity H, Other 4 M	8.33
Severity H, Other 4 L	8.33
Severity M, Other 4 H	6.17
Severity M, Other 4 M	6.42
Severity M, Other 4 L	1.92
Severity L, Other 4 H	5.25
Severity L, Other 4 M	3.50
Severity L, Other 4 L	1.42

Fisher's LSD Comparisons

	difference	
SH,O4H vs SH,O4M	1.83	significant
SH,O4H vs SH,O4L	1.83	significant
SH,O4H vs SM,O4H	4.00	significant
SH,O4H vs SL,O4H	4.92	significant
SH,O4M vs SH,O4L	0.00	nonsignificant
SH,O4M vs SM,O4M	1.92	significant
SH,O4M vs SL,O4M	4.83	significant
SH,O4L vs SM,O4L	6.42	significant
SH,O4L vs SL,O4L	6.92	significant
SM,O4H vs SM,O4M	-0.25	nonsignificant
SM,O4H vs SM,O4L	4.25	significant
SM,O4H vs SL,O4H	0.92	nonsignificant
SM,O4M vs SM,O4L	4.50	significant
SM,O4M vs SL,O4M	2.92	significant
SM,O4L vs SL,O4L	0.50	nonsignificant
SL,O4H vs SL,O4M	1.75	significant
SL,O4H vs SL,O4L	3.83	significant
SL,O4M vs SL,O4L	2.08	significant

Interaction between Externalized/Internalized, Level of Severity, and Level of Other Four Factors
$F=6.89$ $p<.001$

	Mean
Ex/SH/O4H	10.83
Ex/SH/O4M	10.33
Ex/SH/O4L	6.00
Ex/SM/O4H	7.17
Ex/SM/O4M	8.00
Ex/SM/O4L	1.00
Ex/SL/O4H	4.67
Ex/SL/O4M	2.83
Ex/SL/O4L	1.17
In/SH/O4H	9.50
In/SH/O4M	6.33
In/SH/O4L	10.67
In/SM/O4H	5.17
In/SM/O4M	4.83
In/SM/O4L	2.83
In/SL/O4H	5.83
In/SL/O4M	4.17
In/SL/O4L	1.67

259

Fisher's LSD Comparisons

	difference	
Ex/SH/O4H vs Ex/SH/O4M	0.50	nonsignificant
Ex/SH/O4H vs Ex/SH/O4L	4.83	significant
Ex/SH/O4H vs Ex/SM/O4H	3.67	significant
Ex/SH/O4H vs Ex/SL/O4H	6.17	significant
Ex/SH/O4H vs In/SH/O4H	1.33	nonsignificant
Ex/SH/O4M vs Ex/SH/O4L	4.33	significant
Ex/SH/O4M vs Ex/SM/O4M	2.33	significant
Ex/SH/O4M vs Ex/SL/O4M	7.50	significant
Ex/SH/O4M vs In/SH/O4M	4.00	signfilcant
Ex/SH/O4L vs Ex/SM/O4L	5.00	significant
Ex/SH/O4L vs Ex/SL/O4L	4.83	significant
Ex/SH/O4L vs In/SH/O4L	-4.67	significant
Ex/SM/O4H vs Ex/SM/O4M	-0.83	nonsignificant
Ex/SM/O4H vs Ex/SM/O4L	6.17	significant
Ex/SM/O4H vs Ex/SL/O4H	2.50	significant
Ex/SM/O4H vs In/SM/O4H	2.00	significant
Ex/SM/O4M vs Ex/SL/O4L	7.00	significant
Ex/SM/O4M vs Ex/SL/O4M	5.17	significant
Ex/SM/O4M vs In/SM/O4M	3.17	significant
Ex/SM/O4L vs Ex/SL/O4L	-0.17	nonsignificant
Ex/SM/O4L vs In/SM/O4L	-1.83	nonsignificant
Ex/SL/O4H vs Ex/SL/O4M	1.83	nonsignificant
Ex/SL/O4H vs Ex/SL/O4L	3.50	significant
Ex/SL/O4H vs In/SL/O4H	-1.67	nonsignificant
Ex/SL/O4M vs Ex/SL/O4L	1.67	nonsignificant
Ex/SL/O4M vs In/SL/O4M	-1.33	nonsignificant
Ex/SL/O4L vs In/SL/O4L	-0.50	nonsignificant
In/SH/O4H vs In/SH/O4M	3.17	significant
In/SH/O4H vs In/SH/O4L	-1.17	nonsignificant
In/SH/O4H vs In/SM/O4H	4.33	significant
In/SH/O4H vs In/SL/O4H	3.67	significant
In/SH/O4M vs In/SH/O4L	-4.33	significant
In/SH/O4M vs In/SM/O4M	1.50	nonsignificant
In/SH/O4M vs In/SL/O4M	2.17	significant
In/SH/O4L vs In/SM/O4L	7.83	significant
In/SH/O4L vs In/SL/O4L	9.00	significant
In/SM/O4H vs In/SM/O4M	0.33	nonsignificant
In/SM/O4H vs In/SM/O4L	2.33	significant
In/SM/O4H vs In/SL/O4H	-0.67	nonsignificant
In/SM/O4M vs In/SM/O4L	2.00	significant
In/SM/O4M vs In/SL/O4M	0.67	nonsignificant
In/SM/O4L vs In/SL/O4L	1.17	nonsignificant
In/SL/O4H vs In/SL/O4M	1.67	nonsignificant
In/SL/O4H vs In/SL/O4L	4.17	significant
In/SL/O4M vs In/SL/O4L	2.50	signficant

Appendix K

Means and Fisher's LSD Comparisons -

Frequency ANOVA

Main Effect for Externalized/Internalized
F=9.94   p=.002

	Mean
Externalized	6.15
Internalized	5.35

Fisher's LSD   difference=.796   signficant

Main Effect for Level of Frequency
F=25.37   p<.001

	Mean
Frequency High	6.06
Frequency Moderate	6.67
Frequency Low	4.53

Fisher's LSD Comparisons

	difference	
High vs Moderate	-0.61	nonsignificant
High vs Low	1.53	signficant
Moderate vs Low	2.14	significant

Main Effect for Level of Other Four Factors
F=299.57   p<.001

	Mean
Other Four High	9.75
Other Four Moderate	5.28
Other Four Low	2.22

Fisher's LSD Comparisons

	difference	
High vs Moderate	4.47	signficant
High vs Low	7.52	signficant
Moderate vs Low	3.06	significant

Interaction between Externalized/Internalized and Level of Frequency
F=6.81   p=.002

	Mean
Ex/Frequency High	7.00
Ex/Frequency Moderate	7.11
Ex/Frequency Low	4.33
In/Frequency High	5.11
In/Frequency Moderate	6.22
In/Frequency Low	4.72

Fisher's LSD Comparisons

		difference	
Ex/FH vs Ex/FM		-0.11	nonsignificant
Ex/FH vs Ex/FL		2.67	significant
Ex/FH vs In/FH		1.89	significant
Ex/FM vs Ex/FL		2.78	significant
Ex/FM vs In/FM		0.89	significant
Ex/FL vs In/FL		-0.39	nonsignificant
In/FH vs In/FM		-1.11	significant
In/FH vs In/FL		0.39	nonsignificant
In/FM vs In/FL		1.50	significant

Interaction between Externalized/Internalized and Level of Other Four Factors
$F=3.65$ $p=.030$

	Mean
Ex/Other 4 High	10.11
Ex/Other 4 Moderate	6.11
Ex/Other 4 Low	2.22
In/Other 4 High	9.39
In/Other 4 Moderate	4.44
In/Other 4 Low	2.22

Fisher's LSD Comparisons

		difference	
Ex/O4H vs Ex/O4M		4.00	significant
Ex/O4H vs Ex/O4L		7.89	significant
Ex/O4H vs In/O4H		0.72	nonsignificant
Ex/O4M vs Ex/O4L		3.89	significant
Ex/O4M vs In/O4M		1.67	significant
Ex/O4L vs In/O4L		0.00	nonsignificant
In/O4H vs In/O4M		4.94	significant
In/O4H vs In/O4L		7.17	significant
In/O4M vs In/O4L		2.22	significant

Interaction between Level of Frequency and Level of Other Four Factors
$F=3.81$ $p=.007$

	Mean
FH/O4H	10.17
FH/O4M	6.17
FH/O4L	1.83
FM/O4H	10.17
FM/O4M	6.42
FM/O4L	3.42
FL/O4H	8.92
FL/O4M	3.25
FL/O4L	1.42

Fisher's LSD Comparisons

		difference	
FH/O4H vs	FH/O4M	4.00	significant
FH/O4H vs	FH/O4L	8.33	significant
FH/O4H vs	FM/O4H	0.00	nonsignificant
FH/O4H vs	FL/O4H	1.25	significant
FH/O4M vs	FH/O4L	4.33	significant
FH/O4M vs	FM/O4M	-0.25	nonsignificant
FH/O4M vs	FL/O4M	2.92	significant
FH/O4L vs	FM/O4L	-1.58	significant
FH/O4L vs	FL/O4L	0.42	nonsignificant
FM/O4H vs	FM/O4M	3.75	significant
FM/O4H vs	FM/O4L	6.75	significant
FM/O4H vs	FL/O4H	1.25	significant
FM/O4M vs	FM/O4L	3.00	significant
FM/O4M vs	FL/O4M	3.17	significant
FM/O4L vs	FL/O4L	2.00	significant
FL/O4H vs	FL/O4M	5.67	significant
FL/O4H vs	FL/O4L	7.50	significant
FL/O4M vs	FL/O4L	1.83	significant

Interaction between Externalized/Internalized, Level of Frequency, and Level of Other Four Factors

$F=4.19$ $p=.004$

	Mean
Ex/FH/O4H	10.83
Ex/FH/O4M	7.83
Ex/FH/O4L	2.33
Ex/FM/O4H	10.17
Ex/FM/O4M	8.00
Ex/FM/O4L	3.17
Ex/FL/O4H	9.33
Ex/FL/O4M	2.50
Ex/FL/O4L	1.17
In/FH/O4H	9.50
In/FH/O4M	4.50
In/FH/O4L	1.33
In/FM/O4H	10.17
In/FM/O4M	4.83
In/FM/O4L	3.67
In/FL/O4H	8.50
In/FL/O4M	4.00
In/FL/O4L	1.67

Fisher's LSD Comparisons

		difference	
Ex/FH/O4H vs Ex/FH/O4M	3.00	significant
Ex/FH/O4H vs Ex/FH/O4L	8.50	significant
Ex/FH/O4H vs Ex/FM/O4H	0.67	nonsignificant
Ex/FH/O4H vs Ex/FL/O4H	1.50	nonsignificant
Ex/FH/O4H vs In/FH/O4H	1.33	nonsignificant
Ex/FH/O4M vs Ex/FH/O4L	5.50	significant
Ex/FH/O4M vs Ex/FH/O4L	-0.17	nonsignificant
Ex/FH/O4M vs Ex/FL/O4M	5.33	significant
Ex/FH/O4M vs In/FH/O4M	3.33	significant
Ex/FH/O4L vs Ex/FM/O4L	-0.83	nonsignificant
Ex/FH/O4L vs Ex/FL/O4L	1.17	nonsignificant
Ex/FH/O4L vs In/FH/O4L	1.00	nonsignificant
Ex/FM/O4H vs Ex/FM/O4M	2.17	significant
Ex/FM/O4H vs Ex/FM/O4L	7.00	significant
Ex/FM/O4H vs Ex/FL/O4H	0.83	nonsignificant
Ex/FM/O4H vs In/FM/O4H	0.00	nonsignificant
Ex/FM/O4M vs Ex/FM/O4L	4.83	significant
Ex/FM/O4M vs Ex/FL/O4M	5.50	significant
Ex/FM/O4M vs In/FM/O4M	3.17	significant
Ex/FM/O4L vs Ex/FL/O4L	2.00	significant
Ex/FM/O4L vs In/FM/O4L	-0.50	nonsignificant
Ex/FL/O4H vs Ex/FL/O4M	6.83	significant
Ex/FL/O4H vs Ex/FL/O4L	8.17	significant
Ex/FL/O4H vs In/FL/O4H	0.83	nonsignificant
Ex/FL/O4M vs Ex/FL/O4L	1.33	nonsignificant
Ex/FL/O4M vs In/FL/O4M	-1.50	nonsignificant
Ex/FL/O4L vs In/FL/O4L	-0.50	nonsignificant
In/FH/O4H vs In/FH/O4M	5.00	significant
In/FH/O4H vs In/FH/O4L	8.17	significant
In/FH/O4H vs In/FM/O4H	-0.67	nonsignificant
In/FH/O4H vs In/FL/O4H	1.00	nonsignificant
In/FH/O4M vs In/FH/O4L	3.17	significant
In/FH/O4M vs In/FM/O4M	-0.33	nonsignificant
In/FH/O4M vs In/FL/O4M	0.50	nonsignificant
In/FH/O4L vs In/FM/O4L	-2.33	significant
In/FH/O4L vs In/FL/O4L	-0.33	nonsignificant
In/FM/O4H vs In/FM/O4M	5.33	significant
In/FM/O4H vs In/FM/O4L	6.50	significant
In/FM/O4H vs In/FL/O4H	1.67	significant
In/FM/O4M vs In/FM/O4L	1.67	nonsignificant
In/FM/O4M vs In/FL/O4M	0.83	nonsignificant
In/FM/O4L vs In/FL/O4L	2.00	significant
In/FL/O4H vs In/FL/O4M	4.50	significant
In/FL/O4H vs In/FL/O4L	6.83	significant
In/FL/O4M vs In/FL/O4L	2.33	significant

Appendix L

Means and Fisher's LSD Comparisons -

Duration ANOVA

Main Effect for Externalized/Internalized
      F=2.89   p=.092

	Mean
Externalized	5.37
Internalized	4.87

Fisher's LSD   difference=0.50   nonsignificant

Main Effect for Level of Duration
      F=5.06   p=.008

	Mean
Duration High	5.56
Duration Moderate	5.33
Duration Low	4.47

Fisher's LSD Comparisons

	difference	
High vs Moderate	0.22	nonsignificant
High vs Low	1.08	significant
Moderate vs Low	0.86	significant

Main Effect for Level of Other Four Factors
      F=197.63   p<.001

	Mean
Other Four High	8.81
Other Four Moderate	4.89
Other Four Low	1.67

Fisher's LSD Comparisons

	difference	
High vs Moderate	3.91	significant
High vs Low	7.14	significant
Moderate vs Low	3.22	significant

Interaction between Externalized/Internalized and Level of Duration
      F=2.66   p=.076

	Mean
Ex/Duration High	5.33
Ex/Duration Moderate	5.89
Ex/Duration Low	4.89
In/Duration High	5.78
In/Duration Moderate	4.78
In/Duration Low	4.06

```
Fisher's LSD Comparisons
                                 difference
Ex/DH vs Ex/DM                     -0.56        nonsignificant
Ex/DH vs Ex/DL                      0.44        nonsignificant
Ex/DH vs In/DH                     -0.44        nonsignificant
Ex/DM vs Ex/DL                      1.00        nonsignificant
Ex/DM vs In/DM                      1.11        significant
Ex/DL vs In/DL                      0.83        nonsignificant
In/DH vs In/DM                      1.00        nonsignificant
In/DH vs In/DL                      1.72        significant
In/DM vs In/DL                      0.72        nonsignificant
```

Interaction between Externalized/Internalized and Level of Other Four Factors

$F=4.76$ $p=.011$

```
                Means
Ex/O4H           9.33
Ex/O4M           5.50
Ex/O4L           1.28
In/O4H           8.28
In/O4M           4.28
In/O4L           2.06
```

```
Fisher's LSD Comparisons
                                 difference
Ex/O4H vs Ex/O4M                    3.83        significant
Ex/O4H vs Ex/O4L                    8.06        significant
Ex/O4H vs In/O4H                    1.06        significant
Ex/O4M vs Ex/O4L                    4.22        significant
Ex/O4M vs In/O4M                    1.22        significant
Ex/O4L vs In/O4L                   -0.78        nonsignificant
In/O4H vs In/O4M                    4.00        significant
In/O4H vs In/O4L                    6.22        significant
In/O4M vs In/O4L                    2.22        significant
```

Interaction between Level of Duration and Level of Other Four Factors

$F=6.19$ $p<.001$

```
                Mean
DH/O4H          10.17
DH/O4M           4.50
DH/O4L           2.00
DM/O4H           8.00
DM/O4M           6.42
DM/O4L           1.58
DL/O4H           8.25
DL/O4M           3.75
DL/O4L           1.42
```

Fisher's LSD Comparisons

		difference	
DH/O4H vs DH/O4M		5.67	significant
DH/O4H vs DH/O4L		8.17	significant
DH/O4H vs DM/O4H		2.17	significant
DH/O4H vs DL/O4H		1.92	significant
DH/O4M vs DH/O4L		2.50	significant
DH/O4M vs DM/O4M		-1.92	significant
DH/O4M vs DL/O4M		0.75	nonsignificant
DH/O4L vs DM/O4L		0.42	nonsignificant
DH/O4L vs DL/O4L		0.58	nonsignificant
DM/O4H vs DM/O4M		1.58	significant
DM/O4H vs DM/O4L		6.42	significant
DM/O4H vs DL/O4H		-0.25	nonsignificant
DM/O4M vs DM/O4L		4.83	significant
DM/O4M vs DL/O4M		2.67	significant
DM/O4L vs DL/O4L		0.17	nonsignificant
DL/O4H vs DL/O4M		4.50	significant
DL/O4H vs DL/O4L		6.83	significant
DL/O4M vs DL/O4L		2.33	significant

Interaction between Externalized/Internalized, Level of Duration and Level of Other Four Factors
$F=3.98$ $p=.005$

	Mean
Ex/DH/O4H	10.83
Ex/DH/O4M	3.50
Ex/DH/O4L	1.67
Ex/DM/O4H	8.67
Ex/DM/O4M	8.00
Ex/DM/O4L	1.00
Ex/DL/O4H	8.50
Ex/DL/O4M	5.00
Ex/DL/O4L	1.17
In/DH/O4H	9.50
In/DH/O4M	5.50
In/DH/O4L	2.33
In/DM/O4H	7.33
In/DM/O4M	4.83
In/DM/O4L	2.17
In/DL/O4H	8.00
In/DL/O4M	2.50
In/DL/O4L	1.67

Fisher's LSD Comparisons

		difference	
Ex/DH/O4H vs Ex/DH/O4M		7.33	significant
Ex/DH/O4H vs Ex/DH/O4L		9.17	significant
Ex/DH/O4H vs Ex/DM/O4H		2.17	significant
Ex/DH/O4H vs Ex/DL/O4H		2.33	significant
Ex/DH/O4H vs In/DH/O4H		1.33	nonsignificant
Ex/DH/O4M vs Ex/DH/O4L		1.83	significant
Ex/DH/O4M vs Ex/DM/O4M		-4.50	significant
Ex/DH/O4M vs Ex/DL/O4M		-1.50	nonsignificant
Ex/DH/O4M vs In/DH/O4M		-2.00	significant
Ex/DH/O4L vs Ex/DM/O4L		0.67	nonsignificant
Ex/DH/O4L vs Ex/DL/O4L		0.50	nonsignificant
Ex/DH/O4L vs In/DH/O4L		-0.67	nonsignificant
Ex/DM/O4H vs Ex/DM/O4M		0.67	nonsignificant
Ex/DM/O4H vs Ex/DM/O4L		7.67	significant
Ex/DM/O4H vs Ex/DL/O4H		0.17	nonsignificant
Ex/DM/O4H vs Ex/DM/O4M		1.33	nonsignificant
Ex/DM/O4M vs Ex/DM/O4L		7.00	significant
Ex/DM/O4M vs Ex/DL/O4M		3.00	significant
Ex/DM/O4M vs In/DM/O4M		3.17	significant
Ex/DM/O4L vs Ex/DL/O4L		-0.17	nonsignificant
Ex/DM/O4L vs In/DM/O4L		-1.17	nonsignificant
Ex/DL/O4H vs Ex/DL/O4M		3.50	significant
Ex/DL/O4H vs Ex/DL/O4L		7.33	significant
Ex/DL/O4H vs In/DL/O4H		0.50	nonsignificant
Ex/DL/O4M vs Ex/DL/O4L		3.83	significant
Ex/DL/O4M vs In/DL/O4M		2.50	significant
Ex/DL/O4L vs In/DL/O4L		-0.50	nonsignificant
In/DH/O4H vs In/DH/O4M		4.00	significant
In/DH/O4H vs In/DH/O4L		7.17	significant
In/DH/O4H vs In/DM/O4H		2.17	significant
In/DH/O4H vs In/DL/O4H		1.50	nonsignificant
In/DH/O4M vs In/DH/O4L		3.17	significant
In/DH/O4M vs In/DM/O4M		0.67	nonsignificant
In/DH/O4M vs In/DL/O4M		3.00	significant
In/DH/O4L vs In/DM/O4L		0.17	nonsignificant
In/DH/O4L vs In/DL/O4L		0.67	nonsignificant
In/DM/O4H vs In/DM/O4M		2.50	significant
In/DM/O4H vs In/DM/O4L		5.17	significant
In/DM/O4H vs In/DL/O4H		-0.67	nonsignificant
In/DM/O4M vs In/DM/O4L		2.67	significant
In/DM/O4M vs In/DL/O4M		2.33	significant
In/DM/O4L vs In/DL/O4L		0.50	nonsignificant
In/DL/O4H vs In/DL/O4M		5.50	significant
In/DL/O4H vs In/DL/O4L		6.33	significant
In/DL/O4M vs In/DL/O4L		0.83	nonsignificant

Appendix M

Means and Fisher's LSD Comparisons -

Generality ANOVA

Main Effect for Externalized/Internalized
         F=13.31   p<.001

                      Mean
Externalized   6.31
Internalized   5.11

Fisher's LSD   difference=1.20   significant

Main Effect for Level of Generality
         F=4.95   p=.009

                      Mean
Generality High   6.25
Generality Moderate   5.86
Generality Low   5.03

Fisher's LSD Comparisons
                              difference
High vs. Moderate   0.39   nonsignificant
High vs. Low   1.22   significant
Moderate vs Low   0.83   significant

Main Effect for Other Four Factors
         F=127.92   p<.001

                      Mean
Other Four High   8.94
Other Four Moderate   6.06
Other Four Low   2.14

Fisher's LSD Comparisons
                              difference
High vs Moderate   2.89   significant
High vs Low   6.81   significant
Moderate vs Low   3.92   significant

Interaction between Externalized/Internalized and Level of Generality
         F=0.23   p=.790

                      Mean
Ex/G High   6.72
Ex/G Moderate   6.61
Ex/G Low   5.61
In/G High   5.78
In/G Moderate   5.11
In/G Low   4.44

Fisher's LSD Comparisons

		difference	
Ex/GH vs Ex/GM		0.11	nonsignificant
Ex/GH vs Ex/GL		1.11	significant
Ex/GH vs In/GH		0.94	nonsignificant
Ex/GM vs Ex/GL		1.00	significant
Ex/GM vs In/GM		1.50	signficant
Ex/GL vs In/GL		1.17	signficant
In/GH vs In/GM		0.67	nonsignificant
In/GH vs In/GL		1.33	nonsignificant
In/GM vs In/GL		0.67	nonsignificant

Interaction between Externalized/Internalized and Level of Other Four Factors
$F=10.94$ $p<.001$

	Mean
Ex/Other 4 High	10.33
Ex/Other 4 Moderate	7.06
Ex/Other 4 Low	1.56
In/Other 4 High	7.56
In/Other 4 Moderate	5.06
In/Other 4 Low	2.72

Fisher's LSD Comparisons

		difference	
Ex/O4H vs Ex/O4M		3.28	significant
Ex/O4H vs Ex/O4L		8.78	significant
Ex/O4H vs In/O4H		2.78	significant
Ex/O4M vs Ex/O4L		5.50	significant
Ex/O4M vs In/O4M		2.00	significant
Ex/O4L vs In/O4L		-1.17	nonsignificant
In/O4H vs In/O4M		2.50	significant
In/O4H vs In/O4L		4.83	significant
In/O4M vs In/O4L		2.33	significant

Interaction between Level of Generality and Level of Other Four Factors
$F=2.29$ $p=.067$

	Mean
GH/O4H	10.17
GH/O4M	5.83
GH/O4L	2.75
GM/O4H	8.92
GM/O4M	6.42
GM/O4L	2.25
GL/O4H	7.75
GL/O4M	5.92
GL/O4L	1.42

Fisher's LSD Comparisons

		difference	
GH/O4H vs GH/O4M		4.33	significant
GH/O4H vs GH/O4L		7.42	significant
GH/O4H vs GM/O4H		1.25	significant
GH/O4H vs GL/O4H		2.42	significant
GH/O4M vs GH/O4L		3.08	significant
GH/O4M vs GM/O4M		-0.58	nonsignificant
GH/O4M vs GL/O4M		-0.08	nonsignificant
GH/O4L vs GM/O4L		0.50	nonsignificant
GH/O4L vs GL/O4L		1.33	nonsignificant
GM/O4H vs GM/O4M		2.50	significant
GM/O4H vs GM/O4L		6.67	significant
GM/O4H vs GL/O4H		1.17	nonsignificant
GM/O4M vs GM/O4L		4.17	significant
GM/O4M vs GL/O4M		0.50	nonsignificant
GM/O4L vs GL/O4L		0.83	nonsignificant
GL/O4H vs GL/O4M		1.83	significant
GL/O4H vs GL/O4L		6.33	significant
GL/O4M vs GL/O4L		4.50	significant

Interaction between Externalized/Internalized, Level of Generality, and Level of Other Four Factors

$F=3.24$ $p=.026$

	Mean
Ex/GH/O4H	10.83
Ex/GH/O4M	7.00
Ex/GH/O4L	2.33
Ex/GM/O4H	10.67
Ex/GM/O4M	8.00
Ex/GM/O4L	1.17
Ex/GL/O4H	9.50
Ex/GL/O4M	6.17
Ex/GL/O4L	1.17
In/GH/O4H	9.50
In/GH/O4M	4.67
In/GM/O4H	7.17
In/GM/O4M	4.83
In/GM/O4L	3.33
In/GL/O4H	6.00
In/GL/O4M	5.67
In/GL/O4L	1.67

```
Fisher's LSD Comparisons
                          difference
Ex/GH/O4H vs Ex/GH/O4M       3.83      significant
Ex/GH/O4H vs Ex/GH/O4L       8.50      significant
Ex/GH/O4H vs Ex/GM/O4H       0.17      nonsignificant
Ex/GH/O4H vs Ex/GL/O4H       1.33      nonsignificant
Ex/GH/O4H vs In/GH/O4H       1.33      nonsignificant
Ex/GH/O4M vs Ex/GH/O4L       4.67      significant
Ex/GH/O4M vs Ex/GM/O4M      -1.00      nonsignificant
Ex/GH/O4M vs Ex/GL/O4M       0.83      nonsignificant
Ex/GH/O4M vs In/GH/O4M       2.33      significant
Ex/GH/O4L vs Ex/GM/O4L       1.17      nonsignificant
Ex/GH/O4L vs Ex/GL/O4L       1.17      nonsignificant
Ex/GH/O4L vs In/GH/O4L      -0.83      nonsignificant
Ex/GM/O4H vs Ex/GM/O4M       2.67      significant
Ex/GM/O4H vs Ex/GM/O4L       9.50      significant
Ex/GM/O4H vs Ex/GL/O4H       1.17      nonsignificant
Ex/GM/O4H vs In/GM/O4H       3.50      significant
Ex/GM/O4M vs Ex/GM/O4L       6.83      significant
Ex/GM/O4M vs Ex/GL/O4M       1.83      significant
Ex/GM/O4M vs In/GM/O4M       3.17      significant
Ex/GM/O4L vs Ex/GL/O4L       0.00      nonsignificant
Ex/GM/O4L vs In/GM/O4L      -2.17      significant
Ex/GL/O4H vs Ex/GL/O4M       3.33      significant
Ex/GL/O4H vs Ex/GL/O4L       8.33      significant
Ex/GL/O4H vs In/GL/O4H       3.50      significant
Ex/GL/O4M vs Ex/GL/O4L       5.00      significant
Ex/GL/O4M vs In/GL/O4M       0.50      nonsignificant
Ex/GL/O4L vs In/GL/O4L      -0.50      nonsignificant
In/GH/O4H vs In/GH/O4M       4.83      significant
In/GH/O4H vs In/GH/O4L       6.33      significant
In/GH/O4H vs In/GM/O4H       2.33      significant
In/GH/O4H vs In/GL/O4H       3.50      significant
In/GH/O4M vs In/GH/O4L       1.50      nonsignificant
In/GH/O4M vs In/GM/O4M      -0.17      nonsignificant
In/GH/O4M vs In/GL/O4M      -1.00      nonsignificant
In/GH/O4L vs In/GM/O4L      -0.17      nonsignificant
In/GH/O4L vs In/GL/O4L       1.50      nonsignificant
In/GM/O4H vs In/GM/O4M       2.33      significant
In/GM/O4H vs In/GM/O4L       3.83      significant
In/GM/O4H vs In/GL/O4H       1.17      nonsignificant
In/GM/O4M vs In/GM/O4L       1.50      nonsignificant
In/GM/O4M vs In/GL/O4M      -0.83      nonsignificant
In/GM/O4L vs In/GL/O4L       1.67      nonsignificant
In/GL/O4H vs In/GL/O4M       0.33      nonsignificant
In/GL/O4H vs In/GL/O4L       4.33      significant
In/GL/O4M vs In/GL/O4L       4.00      significant
```

Appendix N

Means and Fisher's LSD Comparisons -

Percent of Peers ANOVA

Main Effect for Externalized/Internalized
        F=2.12   p=.149

                    Mean
Externalized   5.91
Internalized   5.50

Fisher's LSD Comparison difference=0.41   nonsignificant

Main Effect for Level of Percent of Peers
        F=14.05   p<.001

                    Mean
High %   6.36
Moderate %   6.08
Low %   4.67

Fisher's LSD Comparisons
                              difference
High % vs Moderate %   0.28   nonsignificant
High % vs Low %   1.69   significant
Moderate % vs Low %   1.42   significant

Main Effect for Level of Other Four Factors
        F=192.38   p<.001

                    Mean
Other 4 High   8.86
Other 4 Moderate   6.08
Other 4 Low   2.17

Fisher's LSD Comparison
                              difference
High vs Moderate   2.78   significant
High vs Low   6.69   significant
Moderate vs Low   3.92   significant

Interaction between Externalized/Internalized and Level
    of Percent of Peers
        F=10.55   p,.001

            Mean
Ex/%H   5.67
Ex/%M   6.61
Ex/%L   5.44
In/%H   7.06
In/%M   5.56
In/%L   3.89

```
Fisher's LSD Comparisons
                          difference
Ex/%H vs Ex/%M            -0.94        nonsignificant
Ex/%H vs Ex/%L             0.22        nonsignificant
Ex/%H vs In/%H            -1.39        significant
Ex/%M vs Ex/%L             1.17        significant
Ex/%M vs In/%M             1.06        significant
Ex/%L vs In/%L             1.56        significant
In/%H vs In/%M             1.50        significant
In/%H vs In/%L             3.17        significant
In/%M vs In/%M             1.67        significant
```

Interaction between Externalized/Internalized and Level of Other Four Factors
F=21.55   p<.001

```
               Mean
Ex/O4H        10.17
Ex/O4M         6.33
Ex/O4L         1.22
In/O4H         7.56
In/O4M         5.83
In/O4L         3.11
```

```
Fisher's LSD Comparisons
                          difference
Ex/O4H vs Ex/O4M           3.83        significant
Ex/O4H vs Ex/O4L           8.94        significant
Ex/O4H vs In/O4H           2.61        significant
Ex/O4M vs Ex/O4L           5.11        significant
Ex/O4M vs In/O4M           0.50        nonsignificant
Ex/O4L vs In/O4L          -1.89        significant
In/O4H vs In/O4M           1.72        significant
In/O4H vs In/O4L           4.44        significant
In/O4M vs In/O4L           2.72        significant
```

Interaction between Level of Percent of Peers and Level of Other Four Factors
F=1.65   p=.171

```
               Mean
%H/O4H        10.17
%H/O4M         6.67
%H/O4L         2.25
%M/O4H         9.00
%M/O4M         6.42
%M/O4L         2.83
%L/O4H         7.42
%L/O4M         5.17
%L/O4L         1.42
```

Fisher's LSD Comparisons

		difference	
%H/O4H vs %H/O4M		3.50	significant
%H/O4H vs %H/O4L		7.92	significant
%H/O4H vs %M/O4H		1.17	nonsignificant
%H/O4H vs %L/O4H		2.75	significant
%H/O4M vs %H/O4L		4.42	significant
%H/O4m vs %M/O4M		0.25	nonsignificant
%H/O4M vs %L/O4M		1.50	significant
%H/O4L vs %M/O4L		-0.58	nonsignificant
%H/O4L vs %L/O4L		0.83	nonsignificant
%M/O4H vs %M/O4M		2.58	significant
%M/O4H vs %M/O4L		6.17	significant
%M/O4H vs %L/O4H		1.58	significant
%M/O4M vs %M/O4L		3.58	significant
%M/O4M vs %L/O4M		1.25	significant
%M/O4L vs %L/O4L		1.42	significant
%L/O4H vs %L/O4M		2.25	significant
%L/O4H vs %L/O4L		6.00	significant
%L/O4M vs %L/O4L		3.75	significant

Interaction between Externalized/Internalized, Level of Percent of Peers, and Level of Other Four Factors
F=3.97   p=.005

	Mean
Ex/%H/O4H	10.83
Ex/%H/O4M	5.17
Ex/%H/O4L	1.00
Ex/%M/O4H	10.33
Ex/%M/O4M	8.00
Ex/%M/O4L	1.50
Ex/%L/O4H	9.33
Ex/%L/O4M	5.83
Ex/%L/O4L	1.17
In/%H/O4H	9.50
In/%H/O4M	8.17
In/%H/O4L	3.50
In/%M/O4H	7.67
In/%M/O4M	4.83
In/%M/O4L	4.17
In/%L/O4H	5.50
In/%L/O4M	4.50
In/%L/O4L	1.67

Fisher's LSD Comparisons

		difference	
Ex/%H/O4H vs Ex/%H/O4M		5.67	significant
Ex/%H/O4H vs Ex/%H/O4L		9.83	significant
Ex/%H/O4H vs Ex/%M/O4H		0.50	nonsignificant
Ex/%H/O4H vs Ex/%L/O4H		1.50	nonsignificant
Ex/%H/O4H vs In/%H/O4H		1.33	nonsignificant
Ex/%H/O4M vs Ex/%H/O4L		4.17	significant
Ex/%H/O4M vs Ex/%M/O4M		-2.83	significant
Ex/%H/O4M vs Ex/%L/O4M		-0.67	nonsignificant
Ex/%H/O4M vs In/%H/O4M		-3.00	significant
Ex/%H/O4L vs Ex/%M/O4L		-0.50	nonsignificant
Ex/%H/O4L vs Ex/%L/O4L		-0.17	nonsignificant
Ex/%H/O4L vs In/%H/O4L		-2.50	significant
Ex/%M/O4H vs Ex/%M/O4M		2.33	significant
Ex/%M/O4H vs Ex/%M/O4L		8.83	significant
Ex/%M/O4H vs Ex/%L/O4H		1.00	nonsignificant
Ex/%M/O4H vs In/%M/O4H		2.67	significant
Ex/%M/O4M vs Ex/%M/O4L		6.50	significant
Ex/%M/O4M vs Ex/%L/O4M		2.17	significant
Ex/%M/O4M vs In/%M/O4M		3.17	significant
Ex/%M/O4L vs Ex/%L/O4L		0.33	nonsignificant
Ex/%M/O4L vs In/%M/O4L		-2.67	significant
Ex/%L/O4H vs Ex/%L/O4M		3.50	significant
Ex/%L/O4H vs Ex/%L/O4L		8.17	significant
Ex/%L/O4H vs In/%L/O4H		3.83	significant
Ex/%L/O4M vs Ex/%L/O4L		4.67	significant
Ex/%L/O4M vs In/%L/O4M		1.33	nonsignificant
Ex/%L/O4L vs In/%L/O4L		-0.50	nonsignificant
In/%H/O4H vs In/%H/O4M		1.33	nonsignificant
In/%H/O4H vs In/%H/O4L		6.00	significant
In/%H/O4H vs In/%M/O4H		1.83	significant
In/%H/O4H vs In/%L/O4H		4.00	significant
In/%H/O4M vs In/%H/O4L		4.67	significant
In/%H/O4M vs In/%M/O4M		3.33	significant
In/%H/O4M vs In/%L/O4M		3.67	significant
In/%H/O4L vs In/%M/O4L		-0.67	nonsignificant
In/%H/O4L vs In/%L/O4L		1.83	significant
In/%M/O4H vs In/%M/O4M		2.83	significant
In/%M/O4H vs In/%M/O4L		3.50	significant
In/%M/O4H vs In/%L/O4H		2.17	significant
In/%M/O4M vs In/%M/O4L		0.67	nonsignificant
In/%M/O4M vs In/%L/O4M		0.33	nonsignificant
In/%M/O4L vs In/%L/O4L		2.50	significant
In/%L/O4H vs In/%L/O4M		1.00	nonsignificant
In/%L/O4H vs In/%L/O4L		3.83	significant
In/%L/O4M vs In/%L/O4L		2.83	significant

Appendix O

Kendall's W Coefficient of Concordance

Subject's Rankings per Problem Behavior Description

PBD#	W	Chi Square	Significance	
1	.6222	7.4667	.1132	
2	.3125	5.0000	.2873	
4	.2880	4.6076	.3300	
5	.1905	2.2857	.6834	
6	.4444	5.3333	.2548	
7	.0444	.5333	.9702	
8	.6610	7.8322	.0941	
10	.2121	4.2424	.3742	
11	.2375	3.8000	.4337	
13	.1073	1.2881	.8634	
14	.1361	2.1772	.7032	
16	.7333	8.8000	.0663	
17	.4788	5.7455	.2190	
18	.5200	10.4000	.0342	*
20	.4889	5.8667	.2093	
22	.1111	1.3333	.8557	
24	.1111	1.3333	.8557	
27	.5728	9.1646	.0571	
28	.1190	1.4286	.8392	
30	.2000	2.4000	.6626	
34	.1667	2.0000	.7358	
35	.3000	4.8000	.3084	
36	.2120	3.3924	.4944	
37	.0375	.6000	.9631	
38	.2889	3.4667	.4830	
39	.4625	7.4000	.1162	
40	.2889	3.4667	.4830	
41	.4444	10.6667	.0306	*
42	.3778	4.5333	.3386	
43	.4519	7.2308	.1242	
45	.3778	4.5333	.3386	
46	.4868	7.7895	.0996	
47	.1500	2.4000	.6626	
50	.5000	8.0000	.0916	
51	.4222	5.0667	.2805	
53	.5500	4.4000	.3546	
54	.0847	1.0169	.9072	
55	.7237	11.5789	.0208	*
58	.6000	4.8000	.3084	
62	.8079	9.6949	.0459	*
63	.5375	8.6000	.0719	
65	.2444	2.9333	.5690	

* p<.05

Appendix P

Mean Ranks on the Five Factors

by Subject

Subject 1:   Severity = 2.71
                    Frequency = 2.43
                    Duration = 3.43
                    Generality = 2.89
                    Percent of Peers = 3.54

Subject 2:   Severity = 2.65
                    Frequency = 3.00
                    Duration = 2.87
                    Generality = 2.91
                    Percent of Peers = 3.57

Subject 3:   Severity = 2.17
                    Frequency = 3.04
                    Duration = 3.98
                    Generality = 3.65
                    Percent of Peers = 2.17

Subject 4:   Severity = 2.35
                    Frequency = 2.54
                    Duration = 3.19
                    Generality = 3.58
                    Percent of Peers = 3.35

Subject 5:   Severity = 2.28
                    Frequency = 2.96
                    Duration = 2.76
                    Generality = 3.46
                    Percent of Peers = 3.54

Subject 6:   Severity = 2.25
                    Frequency = 2.83
                    Duration = 3.54
                    Generality = 2.83
                    Percent of Peers = 3.54

Appendix Q

Spearman's Rhos and Interview Comments

PBD 1

Spearman's rho

	Average2	Outlier
Average1	.9*	.1
Average2		.3

Comments:

	Average	Outlier
Severity bad because its damaging/life threateneing	2	1
Frequency made it worse	2	1
Duration of 2 months makes it a concern	1	

PBD 2

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	-.9	.2	-.4
Average2		.5	.7
Outlier1			.9*

Comments:

	Average	Outlier
Behavior problem serious		2
Not a BD behavior, wouldn't refer or self-contain	1	
% of peers - none should do it		1
% of peers - sounds like a classroom management problem, which decreased their rating	2	
Duration - 3 months not that long		1
Duration - 3 months, lead to higher rating	1	
Frequency - lead to rate higher	1	1
Generality - reduced rating	1	
Generality - increased rating		1
Classroom management problem	2	

PBD 3

Spearman's rho

	Average2	Outlier
Average1	-.5	.3
Average2		.3

Comments:

	Average	Outlier
Severity - no big problem	2	1
The problem is the teacher's, probably poor classroom management	1	1
Does the child have academic problems?		1
Frequency - 4 to 5 times per week is not that bad	2	
Duration - interesting that this has gone on so long	1	
Generality - occurs across environments	1	

PBD 4

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.975*	-.5	.6
Average2		-.525	.725
Outlier1			-.9

Comments:

	Average	Outlier
Is the child having family problems, problems at home?	1	
Severity - people getting hurt sometimes	2	1
Duration - too long for this behavior		1
Duration - short	1	
Frequency - twice a day is too often	2	1
% of peers - alot of fighting going on in this school - reduced rating		2
A control issue		1
What's different about the environments it occurs in?		1
What is the student's age?		1

PBD 5

Spearman's rho

	Average2	Outlier
Average1	.3	-.1
Average2		-.6

Comments:

	Average	Outlier
Duration - not long		1
Frequency lead to lower rating	2	
% of peers decreased rating	1	
Behavior problem isn't severe	1	1
What grade is the child in? Would worry if in high school.		1
Previous interventions?	1	

PBD 6

Spearman's rho

	Average2	Outlier
Average1	.075	.075
Average2		.3

Comments:

	Average	Outlier
Duration - teacher shouldn't have let it go on this long		1
Duration - so infrequent that duration doesn't matter	1	
Severity - not bad, but disruptive		1
% of peers - decreased rating	1	
Frequency decreased rating	1	
Teacher problem, needs a better behavior management program		1

PBD 7

Spearman's rho

	Average2	Outlier
Average1	.2	-.3
Average2		.3

Comments:

	Average	Outlier
Duration - hasn't been going on long	2	1
Generality - why doesn't this happen in other classes?	2	1
Severity - involves other students		1
Severity - not severe, easy to fix	2	
Frequency - decreased rating	2	

PBD 8

Spearman's rho

	Average2	Outlier
Average1	.225	.375
Average2		.9*

Comments:

	Average	Outlier
Severity - rater has "a thing" about vandalism, a terrible thing to do		1
Severity - easy problem to correct	1	
Severity - rated as moderate because destructive to school property	1	
Frequency - happens too often	2	1
Duration - a pretty long time		1
% of peers - no impact on rating, even if one does it, it wrong		1
% of peers - rated it more mild because of %	1	
Generality - going on everywhere, makes rater more concerned about it	1	

PBD 10

Spearman's rho

	Average2	Outlier1	Outlier2	Outlier3
Average1	.1	-.625	-.1	.9*
Average2		-.175	.3	.3
Outlier1			-.025	-.825
Outlier2				.2

Comments:

	Average	Outlier
Duration - made it more moderate	2	
Duration - been doing it quite a while		1
Severity - short changing other students	1	
Severity - not a real severe problem, not BD	1	2
Generality - more than half of child's classes, makes it more serious	1	1
Generality - not doing it in all classes		1
Frequency - not high, no big deal	2	2
% of peers - hard to believe he's the only kid who does this	2	
% of peers - more serious as he's the only one		3
Is the instructional environment adequate?		1
Prior interventions?	1	1

PBD 11

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	-.3	.7	.5
Average2		-.8	-.6
Outlier1			.4

Comments:

	Average	Outlier
% of peers - assumes its a teacher problem, poor classroom management	2	1
Severity - not a BD behavior	1	
Severity - a sever problem, challenging authority	1	
Generality - does it about half the time, which shows she has some control over this	1	1
Frequency - not that high		2
Duration - not important in this instance		1
Duration - going on quite awhile	1	
How old is this student?		1

PBD 13

Spearman's rho

	Average2	Outlier
Average1	-.625	.5
Average2		-.825

Comments:

	Average	Outlier
Severity - not a BD behavior	1	
Severity - interfering with learning	1	
% of peers - means it is probably the child's problem (not teacher's)	1	1
Generality - increased rating		1
Duration - too long	1	1
Frequency - makes it more serious	1	
Prior interventions?	1	1
Learning problems?	2	
Attention deficit disorder?	1	

PBD 14

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.875	.3	.3
Average2		-.025	.225
Outlier1			0.0

Comments:

	Average	Outlier
Severity - lesser of alot of evils, could easily correct	2	1
Severity - child not learning anything		1
Generality - makes it more serious	1	1
Frequency - makes it more serious	1	1
% of peers - makes it more serious	1	
% of peers - doesn't fit with the real world	1	
Is the curriculum appropriate?		1
What happend in the last 6 months?		1
Prior interventions?	1	

PBD 16

Spearman's rho

	Average2	Outlier
Average1	.6	.7
Average2		.5

Comments:

	Average	Outlier
Severity - life threatening, very severe	1	1
Generality - doing it everywhere	2	
Frequency - only been caught four times, lack of info on actual frequency	1	1
Frequency - only four times in 6 months, decreased rating	1	
Duration - a reasonable amount of time		1
% of peers - sharp contrast, especially for a female student	1	1

PBD 17

Spearman's rho

	Average2	Outlier
Average1	.5	-.075
Average2		.425

Comments:

	Average	Outlier
Duration - a long time	1	
Duration - not much impact on rating	1	
Duration - mild duration		1
Severity - not much guilt, predelinquent behavior	1	
Severity - a big problem for the school	1	1
% of peers - no body else does it	1	
Frequency - makes it more severe		1
Why hasn't the setting been changed?	1	
Why have no legal charges been brought against the student?	1	

PBD 18

Spearman's rho

	Average2	Outlier1	Outlier2	Outlier3
Average1	.4	.9*	.7	.3
Average2		.7	.9*	-.4
Outlier1			.9*	-.1
Outlier2				-.3

Comments:

	Average	Outlier
Severity - hurting self is severe	2	3
Frequency - high, makes it more severe	1	3
% of peers - makes it more serious, not a teacher problem	2	2
Generality - not able to deal with structure well	1	
Generality - doing it everywhere	1	2
Duration - 2 weeks, situational?		1
Duration - decreased rating	1	2
Appears to be a learned, avoidance behavior		1
Need a medical consultation, neurological?	1	

PBD 20

Spearman's rho

	Average2	Outlier
Average1	-.1	.1
Average2		.7

Comments:

	Average	Outlier
Duration - too long, makes it more serious	2	1
Frequency - too high	1	1
Severity - injuries, disruptive, serious	2	1
Generality - decreased rating, need to do a situational assessment	1	1
% of peers - makes it more serious	1	
Appropriate placement?	1	
Escape/avoidance behavior?	1	

PBD 22

Spearman's rho

	Average2	Outlier
Average1	-.1	-.6
Average2		-.3

Comments:

	Average	Outlier
% of peers - wonder about community standards	1	
% of peers - student shouldn't be singles out if 20% are doing it.		1
% of peers - are they getting adequate supervision?	1	
Duration - too long	2	1
Severity - wasting district money, could hurt someone	1	1
Generality - done in majority of environments	1	1
Frequency - too high	1	
Why aren't police involved?	1	
Prior interventions?	1	

PBD 24

Spearman's rho

	Average2	Outlier
Average1	-.4	.3
Average2		-.9

Comments:

	Average	Outlier
% of peers - high, is there some external reason?	2	
Frequency - not important	1	
Frequency - low		1
Duration - not important	1	
Severity - human bites are nasty, increased rating	1	1
Generality - low, decreased rating	1	1
What's going on in the classroom?	1	
How old is the student?		1

PBD 27

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.875	-.1	.7
Average2		-.025	.575
Outlier1			.6

Comments:

	Average	Outlier
Generality - low		2
Frequency - pretty high, but not important	1	
Frequency - low, decreased rating	1	
Student is going to end up flunking		1
Poor classroom management	2	2
Not an atypical behavior, not a serious, BD behavior	2	1
Academic/learning problem?		1
Curriculum inappropriate?		1

PBD 28

Spearman's rho

	Average2	Outlier
Average1	.9*	.2
Average2		.2

Comments:

	Average	Outlier
Duration - increased rating	1	1
Duration - no impact on rating	1	
Severity - more of an aggravation than a problem	2	1
Frequency - low, decrease rating	2	1
% of peers - decreased rating	1	1
Easy behavior to change		1
How old is the student?	1	

PBD 30

Spearman's rho

	Average2	Outlier
Average1	1.0*	-.8
Average2		-.8

Comments:

	Average	Outlier
Generality - decreases rating	1	1
Generality - no impact on rating	1	
% of peers - majority do it, decreased rating	2	1
Severity - not getting other classwork done, moderate rating		1
Severity - not severe, decreased rating	2	
Frequency - decrease rating	2	
Must be a boring class		1

PBD 34

Spearman's rho

	Average2	Outlier
Average1	.9*	-.7
Average2		-.7

Comments:

	Average	Outlier
Frequency - excessive	2	
Duration - too long	1	
% of peers - decreases rating	2	
Not serious if the student is in a lower grade - age important		1
Too withdrawn, makes it more severe	2	
Maybe the teacher is harsh, intense?	1	

PBD 35

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	0.0	-.7	.3
Average2		.3	.8
Outlier1			-.3

Comments:

	Average	Outlier
Severity - irritating, but not severe	2	2
Generality - decreased rating	1	1
% of peers - do they have a reason to be afraid?	1	
% of peers - lowered rating	1	
Is student manipulative, controlling?		1
Sounds like a discipline problem/classroom management problem	1	
Has a social assessment been done, are there family problem/abuse at home?	1	1
Age of student? If in high school, would be more of a problem		1
Is the student too hard on herself?	1	
Unrealistic expectations about punishment?	1	

PBD 36

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	-.7	.3	.9125*
Average2		-.025	-.775
Outlier1			-.025

Comments:

	Average	Outlier
Severe behavior, but not BD	1	2
Malingering?		1
Normal behavior, decreased rating	1	
Duration - too long	1	1
Generality - is the student appropriately placed?	1	
Generality - specific to 2 classes - are they heavier academically?	1	1
% of peers - not all that unusual, decreased rating	2	
% of peers - lots of malingering		1
Prior interventions?	1	

PBD 37

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	-.9	.3	-.1
Average2		-.6	0.0
Outlier1			-.4

Comments:

	Average	Outlier
Frequency - increases irritation factor		1
Frequency - makes it more serious		1
Frequency - low	1	
Severity - could interfere with relationships		1
Severity - not very severe	2	
Duration - relatively short	1	
Generality - seems restricted to unstructured, not academic settings	1	1
Generality - makes it more serious	1	
% of peers - would hate to have all those kids doing this all day		1
% of peers - most important factor here (13%)		1
% of peers - not deviant	1	
Age of child? Assumed younger, not unusual behavior for little kids		1
Is she terrified of other kids mowing her down?	1	

PBD 38

Spearman's rho

	Average2	Outlier
Average1	-.6	.6
Average2		-.2

Comments:

	Average	Outlier
Frequency - low, not a big issue	1	
Severity - not a very serious behavior	1	
Generality - happens when group participation is expected	1	
Generality - need to look at why it occurs in some settings and not others	1	1
Duration - fairly long	1	
Duration - not very long for this behavior	1	1
% of peers - 10% doesn't seem high for this kind of behavior.	1	
Elective mutism fairly unusual		1
Lack of confidence or social skills?	1	
What grade is child in?	1	

PBD 39

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.6	-.1	.7
Average2		-.3	.9*
Outlier1			-.1

Comments:

	Average	Outlier
Duration - too long for this to be going on	2	2
% of peers - alot, schoolwide problem, makes more mild	2	1
Generality - generalizing to alot of settings		1
Generality - makes it more mild	1	
Assuming its an elementary school		1
Is it inner city?		1
Need a more effective behavior management system	1	
Truancy not a BD referring problem, SPED has no more clout in getting them to attend.	1	
Student's age?	1	
What aspects of certain classrooms promotes this behavior?	1	

PBD 40

Spearman's rho

	Average2	Outlier
Average1	.3	-.6
Average2		.1

Comments:

	Average	Outlier
Rated this too high the first time		1
Duration - not long	2	1
Frequency - relatively high, but not a big deal average for a normal behavior	2	1
Severity - most important, all factors point rating in low direction (not severe)	2	
% of peers - not a very deviant behavior	1	
Generality - makes it a little more serious	1	
All kids have this problem from time to time, normal behavior	1	
Would want to look at duration of episodes	1	

PBD 41

Spearman's rho

	Average2	Average3	Outlier1	Outlier2	Outlier3
Average1	.9*	.7	.3	.3	-.2
Average2		.6	.4	.4	-.1
Average3			0.0	0.0	-.7
Outlier1				1.0*	.7
Outlier2					.7

Comments:

	Average	Outlier
Frequency - low	1	2
Frequency - high, fairly consistent	1	1
Generality - doesn't affect rating		1
Generality - increases rating	3	1
Severity - not a serious problem		1
Severity - could be that she has a serious problem, behavior a symptom of it	2	
Severity - atypical behavior	1	
% of peers - 10% seems about standard for this behavior, decreased rating		2
Duration - decreased rating		1
Duration - increased rating, not going to fix itself	2	
Age of student?		1
Why is she doing this?		1
Language development problem?	1	

PBD 42

Spearman's rho

	Average2	Outlier
Average1	.6	-.1
Average2		-.3

Comments:

	Average	Outlier
Is the curriculum/pace appropriate?	1	
Timing (before quizzes) is important		1
Severity - a serious problem for the student	2	
% of peers - 10% seems pretty standard for this problem	1	
% of peers - uncommon for this % of kids (high)		1
% of peers - doesn't really matter in this case	1	
Generality - just occuring in a couple of classes, gave less weight to decision		1
Generality - a red light that it occurs in 2 subjects	1	
Duration - short; what changed?		1

PBD 43

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.325	.3875	0.0
Average2		.025	-.125
Outlier1			.975*

Comments:

	Average	Outlier
% of peers - a school wide problem?	1	1
Severity - severe behavior	1	
Severity - not severe	1	2
Duration - a long time, makes it more serious	1	1
Frequency - could be worse, not high	1	2
Frequency - high, persistent	1	
Generality - wonder why it doesn't occur in other classes	1	1
Doesn't seem like an inapprop. behavior		1
How old is the student?	1	

PBD 45

Spearman's rho

	Average2	Outlier
Average1	.6	-.3
Average2		-.1

Comments:

	Average	Outlier
Severity - could be crippling keep her from learning, interacting	2	
Severity - moderate		1
Duration - too long	2	1
Frequency - continual, a concern	1	1
Generality - happens everywhere, not attributable to specific environmental variables	1	1
% of peers - wouldn't expect other student to behave in this manner, very severe problem	1	
% of peers - makes it moderately severe		1

PBD 46

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.9*	-.2	.2
Average2		.1	.4
Outlier1			.7

Comments:

	Average	Outlier
Severity - he's going to aggravate people, not get the teacher interaction he needs.		1
Severity - inappropriate level of communication, low frequency in population.		1
Severity - seems like it would easy to change, moderately serious	1	
Duration - a long time, makes it more serious	1	1
% of peers - makes sense, typical		1
% of peers - makes it more serious	1	
Generality - doesn't happen on playground - need to look at this		1
Generality - makes it less severe	2	
Frequency - all the time		1
Frequency - not good data on this	1	
Language problem?	1	
Emotional problem?	1	

PBD 47

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	-.2	-.9	-.2
Average2		.5	0.0
Outlier1			0.0

Comments:

	Average	Outlier
Abuse going on at home? or at school?		2
Age of child?	1	
Duration - too long without effective intervention	1	2
Generality - happens when other people are around her, when not safely contained in desk		1
Generality - increases rating		1
Generality - seems limited to activities she views dangerous	1	
Frequency - seems like she's exposed to unpredictable/ unprotected situations most of the time		1
Frequency - increases rating	1	1
% of peers - deciding factor, increased rating	1	1
Physical disability? Frail?	1	
Severity - not a severe problem	1	

PBD 50

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.5	.35	.5
Average2		-.1	1.0*
Outlier1			-.1

Comments:

	Average	Outlier
Duration - not firmly established, lowered rating	1	2
Duration - indicates something serious	1	
Generality - happens everywhere, doesn't seem like something he can just turn off (increases rating)	2	2
% of peers - makes more serious	2	1
Drug problem?		1
Recent medication change?	1	
Maybe prepsychotic?		1
Head injury?	1	
Problem at home?	1	
Hearing loss?	1	

PBD 51

Spearman's rho

	Average2	Outlier
Average1	-.2	.1
Average2		.5

Comments:

	Average	Outlier
Duration - what happened 3 weeks ago?	2	1
Generality - seems related to academic demands	1	
Frequency - high	1	
Severity - truancy not a BD problem	2	
Severity - severe because avoiding academic demands	1	
Severity - truancy a severe behavior		1
Are classroom demands realistic? the right level for her?	1	1

PBD 52

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	-.3	-.3	.2
Average2		-.8	0.0
Outlier1			.3

Comments:

	Average	Outlier
Generality - reaction to something in 2 of the 5 classes	2	2
Generality - doesn't make sense, should be all day long		1
Frequency - high, indicates something is wrong	2	1
Duration - can't imagine anyone being that miserable for that long	1	1
Severity - very severe problem	1	2
% of peers - makes it more related to student or how he's being dealt with	2	
How old is the student?	1	

PBD 54

Spearman's rho

	Average2	Outlier
Average1	-.2	-.075
Average2		.225

Comments:

	Average	Outlier
% of peers - 25%, a classroom problem?	2	
% of peers - doesn't make sense		1
% of peers - makes it less serious, moderate	1	
Generality - maybe not a classroom management problem	1	
Frequency - classroom problem	1	
Frequency - daily, a concern	1	
Severity - fairly unusual problem		1
Severity - makes it mild to moderate	1	
Duration - extended period of time, makes it a relatively moderate behavior.	1	

PBD 55

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.9*	.5	.9*
Average2		.7	.7
Outlier1			.2

Comments:

	Average	Outlier
Substance abuse or medical problem?		1
Bipolar depression?	1	
Need to look for patterns of occurrance		1
Severity - this can be severe		1
Generality - occurs everywhere, shows that its more a student problem than a classroom management problem	2	1
% of peers - least important	2	
% of peers - a common behavior, not deviant, more mild		1
Duration - a long time, increased rating	1	
Frequency - increased rating	1	

PBD 62

Spearman's rho

	Average2	Outlier
Average1	.075	.175
Average2		.9*

Comments:

	Average	Outlier
Would rate it lower now		1
Sounds like a high school kid - hormones		1
Severity - not a big problem	2	1
Duration - low	1	
Duration - a long time, but infrequent	1	
Frequency - low	2	
Generality - makes it real mild, specific to PE	1	1
% of peers - a common behavior	1	

PBD 63

Spearman's rho

	Average2	Outlier1	Outlier2
Average1	.7	.3	.1
Average2		.3	.5
Outlier1			.4

Comments:

	Average	Outlier
Generality - decrease rating	1	2
Frequency - seems like it was an isolated event	2	2
% of peers - are there any consequences when this occurs?		1
% of peers - decreased rating	1	
Severity - sounds like an overreaction		1
Severity - didn't act out, just acted angry		1
Severity - maybe he just had a bad day	1	
Severity - a "power mope"; a severe response	1	
Duration - 1 week, behavior not an issue yet	2	1
Would rate lower now		1
Wouldn't refer to Technical Assistance Team		1

PBD 65

Spearman's rho

	Average2	Outlier
Average1	.8	-.6
Average2		-.6

Comments:

	Average	Outlier
Age? High school student - may be puberty	1	1
A situational behavior?	1	1
Severity - not that severe		1
Severity - mild, challenge to authority	1	
Generality - no generality, makes it mild	2	1
Frequency - low, less serious	2	1
% of peers - makes it less severe	1	
% of peers - not a real common problem	1	
Duration - short, decreased rating	1	

VITA

SHEILA S. GIERE

Home address:

90 N. 400 E., Apt. W
Logan, Utah 84321
(801) 752-0198

Work address:
Technology
Division, DCHP
VVT 7628
Logan, Utah 84322-6800
(801) 750-3734

EDUCATIONAL BACKGROUND

Currently enrolled in the Combined Professional-Scientific Psychology Doctoral program, Utah State University. Emphasis in Clinical Psychology. Anticipated completion date: Summer 1990

Degree	Date	Institution	Areas of Concentration
M.S.	1986	St. Cloud State St. Cloud, MN	Applied Behavior Analysis
B.A.	1980	College of St. Benedict St. Joseph, MN	Psychology
A.A.	1978	Willmar Comm. College Willmar, MN	Psychology

EMPLOYMENT EXPERIENCE

9/86-Present   Graduate Assistant, Technology Division, Developmental Center for Handicapped Persons, Utah State University. Responsibilities include conducting research activities, writing grants, undergraduate and inservice instruction, writing artificial intelligence expert systems, developing instructional materials. Have provided technical assistance to programs receiving funding from the Utah Developmental Disabilities Council. Assisted in conducting Program Administrative Reviews for the Utah State Office of Education, monitoring compliance with Public Law 94-142 and the Utah rules and regulations for classifying students for special education services.

6/86-8/86	Internship in applied behavior analysis. Human Services Support Network, St. Paul, MN. Served children and adults with developmental disabilities, mental illness, and behavior problems in community-based settings. Parent training, behavior management and educational programming.
9/85-5/86	Graduate Assistantship, St. Cloud State University. Assistant to the Dean of the College of Education. Conducted follow-up studies of teacher education graduates. Assisted in organizing data, report writing, and other preparations for NCATE and Minnesota Board of Teaching reviews.
9/85-3/86	Teaching Assistantship, St. Cloud State University. Psy 434/534 Applied Behavior Analysis I, and Psy 435/535 Applied Behavior Analysis II.
6/85-8/85	Field Study in organizational behavior management. Participated in program development and quality assurance activities at the Minnesota Learning Center, Brainerd, Minnesota.
9/84-5/85	Graduate Assistantship, St. Cloud State University. Teaching assistant for Psy 115 Personalized System of Instruction Introductory Psychology and Psy 419/519 Scientific and Professional Ethics. Assisted with Behavior Analysis Program development.
3/81-9/84	Bar-None Intensive Treatment Center, Anoka, MN. Direct care of autistic and autistic-like behaviorally disordered children. Major responsibilities included writing and implementing educational and behavioral treatment programs, teaching self-care, independent living, recreation, and communication skills.
6/84-7/84	Minneapolis Children's Hospital Program for Autistic and Other Exceptional Children. Teacher-therapist, summer school.

11/83-6/84	Ramsey County Community Social Services. In-home intervention program, which included parent training and behavioral interventions.

10/81-9/84	Bar-None Respite Care Program. In-home respite care provider for families with handicapped children.

6/79-8/79 1/80-12/80	Research Assistant, Psychology Department, St. John's University, Collegeville, MN. Duties included organizing data, computer processing of data, interviewing subjects, facilitating group discussions.

PROFESSIONAL AFFILIATIONS, HONORS, AWARDS

Association for Behavior Analysis

Student Member, American Psychological Association

Council for Exceptional Children

Phi Kappa Phi

Irene Dunn Memorial Student Award, Region VIII American Association on Mental Deficiency, 9/29/86

PRESENTATIONS

Giere, S., & Baer, R. (1988, December) Computer-based preservice and inservice training in special education classification. Paper presented at the Council for Exceptional Children/Technology and Media Division national conference, Reno, Nevada.

Baer, R., Ferrara, J.M., & Giere, S. (1988, March). Behavior Consultant: A prototype expert system for behavior management. Paper presented at the Council for Exceptional Children national conference, Washington, D.C.

Giere, S., Prater, M.A., Baer, R., Thornburg, M., Ferrara, J., & Althouse, B. (1987, May). 'Behavior Consultant': An expert system for classroom behavior management. Poster presented at the Association for Behavior Analysis, Nashville, Tennessee.

Rudrud, E., & Giere, S. (1987, May). Functional cooking skills. Paper presented at the Ninth Annual National 89-313 Provider's Conference, Rapid City, South Dakota.

Giere, S., Baer, R., & Prater, M.A. (1987, April). 'Behavior Consultant': A prototype computerized expert system for providing teacher advice on managing childrens' behavior. Paper presented at Teaching behavior disordered students: A cooperative venture, sponsored by the Utah Council for Children with Behavior Disorders (a division of Utah CEC), Salt Lake City, Utah.

Giere, S., Rudrud, E., and McKay, S. (1986, November). Functional grocery shopping skills training program. Poster presented at the Association for Persons with Severe Handicaps, San Francisco, California.

Zaiser, A., and Giere, S. (1985, May). A geographical analysis of behavior analysis. Poster presented at the Association for Behavior Analysis, Columbus, Ohio.

Giere, S. (1985, April). Increasing positive peer interaction through group contingent reinforcement. Poster presented at the Minnesota Association for Behavior Analysis, Minneapolis, Minnesota.

Zaiser, A., and Giere, S. (1985, April). Behavioral geography. Poster presented at the Minnesota Association for Behavior Analysis, Minneapolis, Minnesota.

Jensen, M., Giere, S., and Graham, P. (1985, April). Women in MnABA. Poster presented at the Minnesota Association for Behavior Analysis, Minneapolis, Minnesota.

ARTICLES SUBMITTED

Giere, S., Baer, R., Ferrara, J.M., Prater, M.A., & Thornburg, M. Expert systems as an adjunct in providing behavior management training for regular and special education personnel. Submitted to the Journal of Artificial Intelligence in Education.

Giere, S., Rudrud, E.H., & McKay, S. Functional shopping and meal preparation skills. Submitted to the Australia and New Zealand Journal of Developmental Disabilities.

<u>PRODUCTS</u>

Ferrara, J.M., Williams, D., & Giere, S. (1987). <u>CLASS.IH</u> [Computer Program]. Artificial Intelligence Research and Development Unit, Developmental Center for Handicapped Persons, Utah State University. Logan, Utah.

Giere, S., Baer, R., Ferrara, J., Elwell, C., & Althouse, B. (1988). <u>IH.Trainer: A systematic approach for classification of students with intellectual handicaps.</u> Technology Division, Developmental Center for Handicapped Persons, Utah State University. Logan, Utah.

Baer, R., Ferrara, J.M., Giere, S.S., & Serna, R. (1987). <u>Behavior Consultant</u> [Computer Program]. Technology Division, Developmental Center for Handicapped Persons, Utah State University. Logan, Utah.

<u>GRANT FUNDED</u>

Hofmeister, A., Giere, S., & Lubke, M. (1988). <u>Determining the effectiveness of cooperative programming and remediation with group-based videodisc math and science programs in mainstreamed classrooms.</u> U.S. Department of Education--Office of Special Education and Rehabilitative Services. Utah State University.