PRACTITIONER AGREEMENT ON PROBLEM IDENTIFICATION

IN CONSULTATION

by

PATRICIA A. BRIX

B.Sc., The City College of New York, 1974
M.A., Cardinal Stritch College, 1985

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Educational Psychology
and Special Education)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November, 1994

© Patricia A. Brix, 1994
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Educational Psychology/Special Education

The University of British Columbia
Vancouver, Canada

Date December 19, 1994
This was a descriptive study of the agreements reached by learning assistants and classroom teachers when identifying a student's problem(s) during a consultative problem identification interview. The behavioral consultation research literature suggested that problem identification was a critical component of the problem solving process (Bergan & Tombari, 1976) however, the reliability of information gathered during the consultation interviews required further investigation (White & Edelstein, 1991). This study addressed the issue of reliability of the problem identification interview in consultation by examining interrater and interparticipant agreements as to the priority, nature and number of problems identified during the interview.

Nine learning assistance teachers conducted problem identification interviews with each of four classroom teachers from their individual schools regarding students who the teacher identified as difficult to teach. Participants rated their problem identification interviews with an evaluative rating scale of interview helpfulness, and levels of problem identification and shared understanding in their interview dyad. Post consultation interviews with each participant revealed the levels to which each identified the presenting problems in priority by nature and number. Results reported the level to which each interview dyad (N=36) agreed upon the problem(s) identified. Two raters gave independent ratings to the level of shared understanding of the problem(s) identified by the participants as well as to the priority, number and nature of the
problem(s). Participant-rater agreements were determined for the same variables.

The results reported a moderate level of agreement (Kappa = .66) between the participants as to the nature of the highest priority problem. A moderate level of agreement was determined between Rater 2 and the learning assistance teachers (K=.67) and the classroom teachers (K=.78) regarding the nature of the highest priority problem as well.

The implication of these findings suggested that the dynamic process of problem identification is reliable. However, the process may result in lack of complete agreement between participants until the process results in problem descriptions which are specific enough to allow problem solution to be attempted. A replication of this study is needed to further validate these results. Further research is warranted in order to confirm the level at which problem identification is completed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>vii</td>
</tr>
<tr>
<td><strong>CHAPTER ONE: INTRODUCTION TO THE PROBLEM</strong></td>
<td>1</td>
</tr>
<tr>
<td>Prereferral Intervention</td>
<td>3</td>
</tr>
<tr>
<td>Behavioral Assessment</td>
<td>4</td>
</tr>
<tr>
<td>Consultation</td>
<td>5</td>
</tr>
<tr>
<td>Interviews</td>
<td>7</td>
</tr>
<tr>
<td>The Purpose of This Study</td>
<td>11</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>12</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>13</td>
</tr>
<tr>
<td>Scope and Limitations</td>
<td>15</td>
</tr>
<tr>
<td>Summary</td>
<td>16</td>
</tr>
<tr>
<td><strong>CHAPTER TWO: REVIEW OF THE LITERATURE</strong></td>
<td>18</td>
</tr>
<tr>
<td>Consultative Problem Solving</td>
<td>18</td>
</tr>
<tr>
<td>The Initial Interview</td>
<td>19</td>
</tr>
<tr>
<td>Problem Identification Interview</td>
<td>23</td>
</tr>
<tr>
<td>Empirical Support for Problem Identification</td>
<td>25</td>
</tr>
<tr>
<td>Clinical Behavior Therapy Interview</td>
<td>28</td>
</tr>
<tr>
<td>Interviewer Training</td>
<td>34</td>
</tr>
<tr>
<td>Consultee Problem Identification</td>
<td>37</td>
</tr>
<tr>
<td>Current Practices</td>
<td>40</td>
</tr>
<tr>
<td>Summary of the Problem</td>
<td>41</td>
</tr>
<tr>
<td><strong>CHAPTER THREE: DESIGN AND METHODOLOGY</strong></td>
<td>43</td>
</tr>
<tr>
<td>Nature of the Study</td>
<td>43</td>
</tr>
<tr>
<td>Procedure</td>
<td>45</td>
</tr>
<tr>
<td>Participants</td>
<td>45</td>
</tr>
<tr>
<td>Setting</td>
<td>46</td>
</tr>
<tr>
<td>Assignment of Participants:</td>
<td>46</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>48</td>
</tr>
<tr>
<td>Summary</td>
<td>56</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS, CONTINUED

<table>
<thead>
<tr>
<th>CHAPTER FOUR: RESULTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>57</td>
</tr>
<tr>
<td>Participants</td>
<td>57</td>
</tr>
<tr>
<td>Classroom Teachers/ Consultees</td>
<td>58</td>
</tr>
<tr>
<td>Prior Experience with Consultation</td>
<td>59</td>
</tr>
<tr>
<td>Summary</td>
<td>62</td>
</tr>
<tr>
<td>Psychometrics</td>
<td>62</td>
</tr>
<tr>
<td>Participant Interview Evaluation Measure: Part I</td>
<td>62</td>
</tr>
<tr>
<td>Interview Helpfulness</td>
<td>65</td>
</tr>
<tr>
<td>Interview Problem Identification</td>
<td>66</td>
</tr>
<tr>
<td>Shared Understanding Between Participants</td>
<td>68</td>
</tr>
<tr>
<td>Post Consultation Interview Agreement</td>
<td>69</td>
</tr>
<tr>
<td>Interparticipant Agreement</td>
<td>71</td>
</tr>
<tr>
<td>Interview Helpfulness</td>
<td>71</td>
</tr>
<tr>
<td>Problem Identification</td>
<td>72</td>
</tr>
<tr>
<td>Shared Understanding</td>
<td>72</td>
</tr>
<tr>
<td>Participant - Rater Agreement</td>
<td>73</td>
</tr>
<tr>
<td>Participant Interview Evaluation Measure: Part II</td>
<td>76</td>
</tr>
<tr>
<td>Problem Identification by Number, Nature, and Priority</td>
<td>76</td>
</tr>
<tr>
<td>Number of Problems</td>
<td>76</td>
</tr>
<tr>
<td>Nature of the Problem</td>
<td>79</td>
</tr>
<tr>
<td>Reliability of Ratings of the Nature of the Problem</td>
<td>81</td>
</tr>
<tr>
<td>Other Problems Missed or Ignored</td>
<td>83</td>
</tr>
<tr>
<td>Summary</td>
<td>84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER FIVE: SUMMARY AND CONCLUSIONS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interview Evaluation Measure</td>
<td>88</td>
</tr>
<tr>
<td>Part I:</td>
<td>88</td>
</tr>
<tr>
<td>Interparticipant and Interrater Agreement</td>
<td>89</td>
</tr>
<tr>
<td>Interview Evaluation Measure</td>
<td>91</td>
</tr>
<tr>
<td>Part II: Number of Problems</td>
<td>91</td>
</tr>
<tr>
<td>Nature of the Problems</td>
<td>92</td>
</tr>
<tr>
<td>Implications</td>
<td>93</td>
</tr>
<tr>
<td>Limitations</td>
<td>95</td>
</tr>
<tr>
<td>Future Directions</td>
<td>96</td>
</tr>
</tbody>
</table>

REFERENCES .......... 97
APPENDICES .......... 105
LIST OF TABLES

Table 4.1: Participant Demographic Information ........................................ 58
Table 4.2: Participant Demographic Information ........................................ 59
Table 4.3: Participant Familiarity with Consultation ................................. 60
Table 4.4: Participant Familiarity with Consultation .................................. 61
Table 4.5: Means (and Standard Deviation) of Learning Assistance and Classroom Teacher Ratings of Interview Characteristics Across Schools .................................................. 64
Table 4.6: Distribution of Learning Assistants' and Classroom Teachers' Ratings of Interview Helpfulness ................................................................. 66
Table 4.7: Distribution of Learning Assistants' and Classroom Teachers' Ratings of Problem Identification ................................................................. 67
Table 4.8: Distribution of Learning Assistants' and Classroom Teachers' Ratings of Shared Understanding ............................................................... 68
Table 4.9: Interrater Agreement ................................................................. 69
Table 4.10: Participant - Rater Agreement .................................................. 74
Table 4.11: Participant Interview Evaluation: Part II Mean Number of Problems Reported by Participants in Individual Schools ..................... 77
Table 4.12: Problem Identification: Number of Problems Identified By Participants ................................................................. 77
Table 4.13: Interrater Agreement on the Number of Identified Problems Per Interview ................................................................. 79
Table 4.14: Problem Identification: Nature of the Problem and Ranking Identified by Participants ................................................................. 80
Table 4.15: Interrater Agreement On the Nature of the Highest Priority Problem ................................................................. 82
Table 4.16: Comparison Agreement Between Participants and Rater 2 .......... 82
Table 4.17: Other Problems Unreported During The Interview .................. 83
ACKNOWLEDGEMENTS

Deus Caritas Est

An undertaking of this kind can not and is not done alone. I am grateful for the direct and indirect contributions made by so many colleagues, friends, and former students to this work. Notable contributions were made by the following persons to whom I must express my greatest appreciation:

It is to Dr. William McKee, for his patience and tireless effort in supervising this project from inception to completion, that I owe a tremendous debt of gratitude. Thanks, Bill, for your willingness to share your wisdom, time, and expertise in support of my emergent skills as a researcher.

My gratitude also extends to the members of my committee, Dr. Nand Kishor and Dr. Jon Shapiro for their timely assistance and encouragement.

An additional word of thanks goes to the members of the EPSE faculty, staff and graduate students for their support with special thanks to Dr. William Reynolds, the PRTC staff and Leona Spencer for their assistance along the way.

I wish to acknowledge the contribution of the Learning Assistance Teachers and Classroom Teachers from the Vancouver Catholic Schools for their willingness to risk participation in research by sharing their consultative skills to further the field of educational psychology.

• • • • •

I joyfully owe a continuing debt of gratitude to the Sisters of Charity of St. Vincent de Paul (Halifax) for their willingness to support and share in my dream as it unfolds within our communal commitment to making the love of God visible in the world today. A special acknowledgement goes to Sister Joan Butler in appreciation for her visible support and inspiration.

Long distance thanks extend to the Eastern connection whose love and support continue to sustain me: my parents, my siblings and their families, and especially Hobo:

"Belay on.....Climbing....."
CHAPTER ONE: INTRODUCTION TO THE PROBLEM

The past two decades have brought about a growing number of changes in the field of special education regarding the classification and placement of special needs students. A reflection of these changes is mirrored in the evolving roles of special education service providers, and in particular, that of the school psychologist. Traditionally the major role of school psychologists has been the identification and classification of special needs students through standardized assessment procedures. Known as the "gatekeepers," school psychologists are responsible for identifying students in need of special educational placement and students who can return to the regular education stream (Will, 1988). Reschly (1988) reported that up to two-thirds of the school psychologist's time was spent in determining student eligibility for special education. The growing influx of students through the referral/classification/placement process led to questioning of the efficiency and effectiveness of traditional assessment practices. Will (1986) contended that this practice resulted in the development of two parallel systems of education in which services had become fragmented and thus reduced opportunities for communication and collaboration of resources. Another concern with this practice had been the resulting denial of services to "at-risk" students who do not fall within the established educational policy boundaries of "special needs" but for whom the availability of some psychological or educational interventions would mean the difference between a successful rather than unsuccessful school experience.
The US Department of Education responded to this growing concern by establishing the Regular Education Initiative (Will, 1988) which proposed the expansion of the regular education system to include all students. School psychologists need not abandon their assessment practices but take a proactive stance by using intervention and prevention techniques to assist classroom teachers in their role of educating all students, particularly those with special learning needs. Prior to the Regular Education Initiative, the National Association of School Psychologists issued a position statement challenging the inadequacy of assessment measurement techniques used for categorical placement of students and supporting assistance with student program development (Kratochwill & Sheridan, 1990). These proposed changes suggested the need for a new paradigm of service delivery for all students requiring educational and psychological services in schools.

The need for alternatives to current practices was further supported by empirical research of assessment methods. Lilly (1988) reviewed critiques of the special education and regular education approaches to traditional assessment and classification practices, reporting that such practices did not offer enough empirical evidence to support continued use. Reschly (1988) stated concern with the psychometric adequacy of commonly used assessment instruments. Rosenfield and Reynolds (1990) cited examples of the lack of treatment utility with current assessment practices. Standardized assessment results were not easily transformed into practical interventions. Another important consideration of the referral/assessment/placement outcome was its orientation to funding and
eligibility issues. Inappropriate diagnoses and mislabelling were not uncommon occurrences, particularly with a large number of referrals.

Despite expressed skepticism with traditional assessment methodology, it continues to be utilized. Alternative assessment methods are increasingly used but they need further research to establish their effectiveness as supplemental or replacement procedures for traditional ones. Suggested skills to supplement traditional assessment practices have included: more structured approaches to gathering client information from direct observation of behaviors and their surrounding environmental events; increasing the practitioner’s knowledge of behavioral change principles and instructional design; and interpersonal relationship skills such as those inherent to consultation methodology (Reschly, 1988, p.468). These skills typify the methodology of prereferral intervention, behavioral assessment, and consultation, which are among the more popular alternative assessment practices. They support the role of school psychologists as problem-solvers, decision-makers and enablers of other professionals to do the same. Each of the practices share a similar approach to problem solving and receive the same criticisms as well.

Prereferral Intervention

A focal shift from standardized assessment practice to an intervention-oriented one has received more recent attention and support (Will, 1988). A prereferral intervention approach identifies problems and intervening strategies which will assist the student to remain in an appropriate environment either in place of or prior to formal referral. This approach potentially reduces the number
of referrals, classifications and placements into special education. Support for this model was demonstrated in the Carter and Sugai (1989) nationwide survey of US State Departments of Education. Nearly 75% of the individual states' policies support some form of a prereferral intervention model. The proposed Special Education Guidelines (1994) in British Columbia encouraged teachers to use this approach prior to consulting with resource personnel. As popular as it seemed, this model was also reported by Carter & Sugai (1989) as lacking sufficient empirical support to demonstrate its effectiveness in keeping students in the mainstream of education. They encouraged its continued use, however, and suggested alternative practices to supplement those reported in the survey. Kratochwill and Sheridan (1990) favorably reviewed several models of prereferral intervention but also noted that empirical support was lacking. Practitioners were encouraged to continue to use this approach for its beneficial aspects while on-going research was suggested.

Behavioral Assessment

A similar alternative service delivery model is found in behavioral assessment. Practitioners utilize similar skills in gathering information about the client and interfering problems using a functional analysis approach (O’Neill, Horner, Albin, Storey & Sprague, 1990). This process involves targeting a problem behavior, identifying its surrounding events and introducing an intervention intended to alter the behavior to a desired state. Behavioral assessment utilizes standard assessment techniques of information gathering such as questionnaires, checklists, self-report measures and standardized
instruments in order to provide a complete contextual picture of the targeted behavior. Effective practitioners characteristically can apply constructs from social learning theory as a basis for determining solutions to problem behaviors or problematic solutions (Alessi, 1989). The interventions developed during this process may be refined throughout assuring feedback on effectiveness to practitioners. It is the methodology of behavioral assessment which supports the practice of prereferral intervention, intervention, and prevention program development. This methodology has helped to refocus the role of the school psychologist into a stronger problem solving model in which problems rather than students are identified.

Consultation

A third alternative service delivery model, which incorporates aspects of both prereferral intervention and behavioral assessment, is consultation. Consultation methodology is inherent to the practice of assessment and intervention problem solving models. It is an interpersonal influence model which requires dialogue for change to occur between participants. The practice of consultation in a school setting has been given a new and longer look by professionals seeking to re-focus traditional assessment procedures into greater collaborative problem solving activities. The major goal of consultation is that of solving current problems and increasing the ability of its practitioners to solve future problems (Gutkin & Curtis, 1990). Consultation provides a framework for enhancing traditional assessment methods by providing a structure for voluntary communication in which problem solving can occur among professionals. It is
practiced in many occupations. Of the ten models of consultation identified in school settings, the model of behavioral consultation has been noted to be the most widely researched and popular with educators (Medway, 1979; West & Idol, 1987). Consultation methodology underlies the alternative service delivery models suggested for use in the regular education initiative. Consultation has generated much attention in professional journals and research literature in recent years, especially in its utilization in the field of school psychology.

Consultation methodology contains the basic approaches suggested by most alternative service delivery models. It has provided the framework for this study of an essential component of its process: problem identification.

Witt and Elliott (1983) observed that as a total unit, the process of consultation has demonstrated its effectiveness, particularly with regard to its efficiency of professional time and energy. However, in order for consultation to be understood and utilized to its fullest, it is necessary to address each component part separately. From a behavioral consultative perspective, Witt and Elliott (1983) discussed the importance of the initial interview phase of consultation, the entry level into its process of problem solving. The initial interview provides the context for problem identification to occur in consultation. As an integral part of assessment practices, its effectiveness is worthy of further consideration. Identifying a problem for solution is a critical component because as Lazarus (1973) noted, a faulty diagnosis can interfere with assistance intended for participants. In a consultative framework, this suggested that difficulties with
the interview process, especially in problem identification, could impede the progress of consultation.

**Interviews**

The typical way to gather referral information about a student’s problem is through an interview. Initially this interview may be held between classroom teachers and school psychologists to identify and operationalize the problem(s) in order to proceed with a problem solving process. A behavioral interview focuses on systematic targeting of behaviors along with identification of antecedents and consequences to allow the assessor to conduct a functional analysis of the undesired behavior. An assumption made in this process is that the chances are greater that problem solving will proceed to a successful completion once the problem behavior is identified or targeted (Bergan & Tombari, 1976). This has placed strong emphasis on the interview process facilitating problem identification.

Appropriate and effective use of an interview was a universal concern for its practitioners. Since the interview itself constitutes an assessment practice, albeit an indirect method, it is subject to the same psychometric requirements of reliability and validity as are all standardized assessment procedures. Although not all researchers agree that behavioral assessment methodology need be evaluated by the same traditional psychometric requirements, as with other assessment procedures, there needs to be some application of empirical analysis by which a successful assessment outcome may be identified and supported.
Gresham and Davis (1988) reported inadequate empirical support for reliability of behavioral interviews. The studies they cited in support of their claim typically focused on interrater reliability to establish reliability for the interview as an assessment procedure. Bergan and Tombari (1975; 1976) conducted early research on the reliability of the behavioral consultation interview demonstrated as part of a sequence of problem solving activities. Their findings suggested that when a problem was identified it would lead to resolution. In their analysis of 806 consultation cases problem identification accounted for nearly 41% of the variance of specific consultant variables present at this initial phase. In turn, problem identification contributed significantly to plan implementation leading to problem solution. This finding led to the specification of verbalizations made during consultation interviews which would lead to problem identification. A coding system to analyze verbalizations, known as the Consultation Analysis Record (CAR) was devised and reliability was determined by comparing independent ratings of verbalizations. Interrater reliability of verbalizations suggested for problem identification was $r = .92$.

Follow-up studies conducted by Bergan and Tombari (1976), Brown, Kratochwill and Bergan (1982) and others supported this practice of establishing reliability of independent ratings of verbalizations in the process based on their schema of an interview format.

The findings from Bergan and Tombari (1976) contributed to the assumption that problem identification must occur in order for the remainder of the problem solving process to proceed. An interesting finding is that problem
identification only occurred in 43% of the cases and the remaining 57% did not have a problem identified for various reasons. A major reason cited for 44% of the referrals not continuing was that a referral for testing or staff evaluation had occurred, the rest were quite variable but not unusual given the circumstances. Bergan and Tombari (1976) concluded that consultative problem solving inevitably led to problem resolution and those unresolved or uncompleted cases had not identified a problem at the referral stage. Follow-up studies typically included consultant training to ensure that problem identification in consultation would not be overlooked. An issue which remained was the reliability of problem identification itself as determined beyond the simple acknowledgement of whether or not it had occurred.

The issue of reliability has been addressed primarily as interrater reliability. Haynes and Jensen (1979) noted that interview reliability is not often determined which could lead to questionable validity for information gathered during the interview. This may be troublesome when identified problems are targeted for intervention. Studies reporting interobserver and interrater agreement on the interview information gathered is noticeably sparse. Gresham (1984) reviewed a number of studies which attempted to investigate interview reliability through interrater reliability. He reported that interrater reliability can be studied through the coding of verbalizations by independent raters which demonstrates good reliability as in the coding system developed by Bergan and Tombari. This method addresses the ability of the raters to agree on coding for the interview verbalizations and may not necessarily address the nature of the
interview content. Further in-depth study would be required to examine the level of their agreement or to validate the nature of the problem on which they are agreeing.

Interrater reliability can also be studied as agreement between the interviewers and between the raters as was attempted by Hay, Hay, Angle and Nelson (1979) which reported the lack of interviewer generality with regard to the specific problems addressed in the interviews. Further investigation needed to be done in order to support the reliability of information gleaned from the interview for the development of interventions.

This criticism of problem identification in behavioral interviews was also addressed in the clinical behavior therapy literature. An important finding in a study conducted by Hay, Hay, Angle and Nelson (1979) was that interviewers of the same clients were able to agree on the number of problem areas identified but not on the nature of the problems identified, although interrater agreement as to the problem areas identified for each interviewer ranged from .90 to .75. Since the purpose of the interview was to target (identify) problem behaviors for which a functional analysis for problem solving could occur, this finding did not support problem identification agreement among the interviewers. Wilson and Evans (1983) presented case studies to 118 professionals and found a mean percentage agreement of 38.6% for the highest priority problem area. A later study conducted by Felton and Nelson (1984) found low inter-assessor agreement for hypothesized controlling variables following the specification of the problem behavior to the interviewers prior to conducting the interview.
Recent criticism of interview use comes from White and Edelstein (1991) and Nezu and Nezu (1993) who conclude there remains a lack of empirical evidence to support the accuracy and reliability claims of the behavioral assessment interview. The underlying principles of psychometric assessment practices require that an instrument first be reliable in order to be valid and assure the practitioner of appropriate results (Anastasi, 1988). Research findings from several approaches raise the question as to whether or not the initial interview is a reliable way of identifying the nature of problems, particularly in selecting them as target behaviors for planning appropriate interventions. Also, problem identification or the selection of target behaviors for intervention which occurs within this process may not be accomplished in a way that would validate an identified problem to solve. A potential way in which to approach the reliability issue would be to conceptualize the reliability of the interview process as consistent observations of problem representations by both participants in the problem identification interview. From a consultative perspective, agreement between participants could be sought as to the problems identified which will be targeted for intervention.

The Purpose of This Study

The purpose of this study was to describe current practices of the initial interview as a means of identifying the nature of a problem or target behaviors within the practice of consultation and to evaluate its effectiveness in terms of interparticipant and interrater agreements on the nature and number of the problems identified. This was based on the assumption that an interview
conducted by a reasonably skilled interviewer would likely lead to problem identification and agreement between the participants as to what had been discussed in their interview regarding the nature of the highest priority problem. This study attempted to respond to the reported needs of researchers to establish greater empirical support for use of problem identification in the initial interview as a reliable problem solving technique. Independent ratings for consultative participants' agreement were obtained from the participants themselves. Interrater agreements were determined from two raters independently rating the participants' responses to problem identification and agreement measures. This study proposed to increase support for current practices of consultation models used in school situations. It attempted to demonstrate the reliability of the process through agreement on problems identified by participants and two independent raters. In summary, this study attempted to answer the following question:

Do practitioners of consultation agree on the problem identified following an initial interview for problem identification in the process of consultation?

Significance of the Study

This study added to current knowledge of the reliability of the initial interview for identifying the nature of a student’s (client) problems. The consulting participants’ agreements on the interview outcome were compared with agreements made by independent evaluators (raters) of the interview results.
This study attempted to respond to identified research needs of evaluating components of the consultation process leading to successful outcomes rather than viewing it solely as a unitary process. It obtained further knowledge of consultation efficacy as a problem solving process used by consulting teachers in naturalistic settings.

**Definition of Terms**

**Behavioral Assessment Interview** - an interview during which the necessary information to engage in the assessment process is elicited using behavioral descriptors in operational and measurable terms.

**Behavioral Consultation Interview** - an interview between a consultant and consultee regarding a client rather than an interview directly between a consultant and client or therapist and client. This interview is characteristically similar to a behavioral assessment interview.

**Client** - a student with whom the teacher/consultee is experiencing difficulty or finds difficult to teach.

**Consultant** - a school psychologist or educator with some training and/or experience in conducting a problem solving process in a consultative manner. The consultant usually has direct contact with the consultee and occasionally, if appropriate, with the client. In this study the learning assistance teacher from each school assumed the role of the consultant in conducting the consultation interviews and therefore the terms "consultant, consulting teacher, and learning assistant or assistance teacher" are used interchangeably.
Consultation - an indirect service delivery model of problem solving which occurs through collaborative dialogue between a consultant and consultee regarding a client with whom the consultee is experiencing difficulty. Participants engage in the process voluntarily and consultees anticipate improvement in their present problem-solving skills as well as in handling future situations.

Consultee - a classroom teacher who is experiencing difficulty with a student and voluntarily seeks assistance with this student through the consultative process. The terms "consultee and classroom teacher" are used interchangeably.

Functional Analysis - the process of behavioral assessment in which the target behaviors are identified along with their intervening variables (antecedents and consequences) which are impacting upon these behaviors.

Interrater Agreement - ability of two independent raters, following a training session by the researcher, to obtain similar results when scoring transcribed protocols.

Intervention - the process of manipulating the intervening variables around target behaviors as a way of bringing about positive behavioral change.

Interview - the verbal interchange or dialogue which occurs between two or more persons as a medium for obtaining relevant information.

Participant Agreement - established by agreement of the consultant and the consultee as to the nature of the client's (student's) problem(s) as stated in independent post-consultation interviews following the consultation interview process of problem identification.
Prereferral Intervention - the practice of identifying problems and conducting a functional analysis from which situational solutions (interventions) are devised and attempted either in place of or prior to formal referrals for assessment.

Problem - an identified discrepancy between a demonstrated behavior and its desired state.

Problem Identification - the beginning of the consultative problem solving process which involves the consultant and consultee naming and agreeing upon the target behaviors exhibited by the client to be in need of change.

Reliability - consistency of results over time, place, and persons. The ability of a psychometric instrument to produce the same results again and again thereby reducing error variance (Anastasi, 1988).

Target behaviors - those behaviors exhibited by the client which the consultant and consultee identify as interfering with the client’s ability to be successful in the classroom.

Scope and Limitations

This study examined participant agreement on the nature of students’ problems following the initial interview, a critical component of the problem identification phase in consultation. The consultation model of problem solving was approached primarily from a behavioral perspective, identified in research literature since 1970. Volunteer learning assistants and classroom teachers from nine schools in the Catholic Archdiocese of Vancouver participated by conducting initial interviews for problem identification. Post-consultation interviews were conducted by the researcher. Post-consultation interview
transcriptions were coded for participant agreement and rated independently by two raters to establish interrater agreement as to the nature of the problem(s) identified. Results reflected current consultation practices used in the Catholic schools in Vancouver.

**Summary**

Alternative models of assessment challenged the role of school psychology to include ways of providing services to more students following the regular education initiative. Consultation has been one way in which the problem solving aspects of assessment are emphasized. An interview is the context in which problem solving is initiated and the importance of problem identification is paramount to successful completion of the process. Studies conducted in several disciplines as to the effectiveness of defining a problem in an interview have demonstrated that raters of the interview content generally agree on what was discussed but participant agreement on problem areas is considerably lower. This could lead to gathering questionable data in an assessment interview and also to the reliability of the problem identified as well. A new study was proposed to investigate problem identification within an interview context by assessing the level of agreement by participants as to the nature of the highest priority problem.

This study proposed to describe the practice of problem identification in the initial interview currently used by learning assistance teachers in consultation with classroom teachers. The agreement on a student's highest priority problem(s) following their consultation interview will be determined.
This study addressed the following question:

Do practitioners of consultation agree on the problem identified following an initial interview for problem identification in the process of consultation?
CHAPTER TWO: REVIEW OF THE LITERATURE

The purpose of this chapter was to review the literature with respect to problem identification as it occurs in the initial interview of consultation. Research literature from several assessment methodologies were reviewed with regard to their use of an interview as an instrument for conducting problem identification.

Consultative Problem Solving

Approaching the practice of school psychology from a consultative framework allows the psychologist to function in a problem solving mode. The process of problem solving is well described from a behavioral perspective, particularly in behaviorally oriented literature (e.g. Kazdin, 1985; Kratochwill, 1985; Martens, 1993). Consultation, notably behavioral consultation, has been identified as an important way of increasing the problem solving skills of both consultants and consultees. The process which the consultant and the consultee undertake to problem solve is represented in a series of stages which suggests that problem resolution follows sequential movement through designated steps. The first step is known as problem identification in which consultant and consultee voluntarily engage in a dialogue to define and clarify problems faced by the consultee in relationship to a client. The problem analysis phase requires further discussion of the problem and a functional analysis of its environmental contexts followed by the formulation and implementation of plans or interventions to solve it. These interventions are then assessed to determine their
effectiveness and finally evaluated for the attainment of a positive outcome with the possibility of redesigning them if the desired outcome is not achieved (Bergan & Kratochwill, 1990; Gutkin & Curtis, 1990; Idol, 1990; Polsgrove & McNeil, 1989; West & Idol, 1987). A fundamental aspect of this approach is that it is goal-directed in which the participants may structure the process in light of specified goals to which they anticipate achievable ends (Bergan & Kratochwill, 1990).

An advantage of viewing the consultation process as a problem solving activity is that the consultant is able "...to operationalize the conceptual basis of the process" (Sloves, Docherty & Schneider, 1979, p. 30). That is, consultation participants actively address a problem so that it will be "operationalized," reframed, represented or clarified in a structure which will allow it to be examined systematically within the environment. Participants are able to determine and agree upon the presence and function of the problem. This enables a process for successful resolution to occur. In order to clarify consultative problem solving as an assessment practice, it is important to examine the stage at which the problem is identified, the place where "operationalization" begins, the initial interview.

The Initial Interview

The interview is a method used in all forms of assessment. It is a component of the assessment procedure utilized as a basis for referral. An initial interview is the first formal conversation or dialogue which occurs for the purpose of setting the goals of assessment. The interview is more than a
nonspecific cursory meeting between consultant and consultee to gather minimally reliable information about the client (Gresham, 1984). It can assist participants in determining a direction in which to proceed. In behavioral consultation, the initial interview between a consultant and consultee is used for targeting behaviors which are deemed "problems." It allows the collection of information necessary for determining what outcomes the consultee desires as a function of the process as well as assisting to evaluate these outcomes (Haynes & Jensen, 1979).

The consultation interview is not unlike interviews conducted for the purpose of clinical diagnoses. In addition to gathering background information about a client from a variety of sources, it may include a direct interview with the client. The clinical interview is used to assist in forming a preliminary diagnosis as to the client’s presenting problems. The suggested interview format for this situation is a structured one during which the interviewer asks designated questions, typically following a standard set of procedures and guidelines (Nuttall & Ivey, 1986). As an assessment instrument, the interview is required to demonstrate adequate psychometric properties to assure the clinician and client that reliable and valid information is gathered for diagnostic purposes. The interview would then be subjected to the requirements of reliability and validity in order to ensure that the information gathered will lead to a successful outcome (Gable, Friend, Laycock & Hendrickson, 1990).

In clinical practice, the diagnostic interview is critical to the assessment process. Nuttal and Ivey (1986) reported that early studies which investigated the
reliability and validity issues around the interview process found it to be an unreliable measure. Some reasons cited for its questionable reliability were lack of clarity and precision in operationalizing diagnoses and a need for better trained clinicians. On-going solutions were found by refining the operational definitions needed for characterizing psychiatric conditions and using the kappa statistic for equalizing base rates for different diagnoses (Nuttal & Ivey, 1986, p. 114).

Investigating the reliability and validity of psychiatric clinical interview data was part of a comprehensive research project developed by Graham and Rutter (1968). They investigated the reliability of gathering information about a child’s psychiatric disorder based on the parent interview. The authors reported a semi-structured interview format was used in which mothers of 268 children suspected of having psychiatric disorders were given an opportunity to elicit spontaneously symptoms observed in their children. 119 children from this group were identified by their mothers as having disorders characterized as neurotic, antisocial, mixed or other based on their behavior for the previous year. This diagnosis was supported, in part, on mothers’ responses when queried whether they believed their child to have a problem and its manifestation by degree of severity and in comparison with other children. The interviews were rated independently by the authors using a 4 point scale designed to rate the presence or absence of symptoms rather than the mothers’ responses. Interrater agreement was determined by a second rating given to 80 of the interviews schedules for symptoms and overall severity of the disorder
and diagnosis. A correlation of .81 for overall diagnostic agreement was achieved between the two raters.

A second interview was conducted with thirty-six of the mothers which asked for symptoms exhibited during the same time period as the first interview. Overall diagnostic agreement of ratings between the two interviews was .64. An interesting finding reported between the first and second interview was that the correlation coefficient for the ratings given the parental perception of childhood disturbance was .43 indicating that a parent was as likely to state she perceived her child as disturbed in the first as in the second interview. The researchers suggested that parental inconsistency in identifying their children's problems would require consideration in further research activities. Other difficulties which the researchers encountered were rating individual symptoms because of overlap of many symptoms by researcher descriptions. Also, parental descriptions of symptomatic behaviors did not always conceptually match researcher descriptions. For example, the parental description of "overactive behavior" was characteristic of three distinct areas described by the authors. An additional threat to reliability was interviewers who did not adhere to the standardized administration of the interview resulting in data which was useless to the process.

The researchers concluded that following a structured interview resulted in reliable overall psychiatric disorder diagnoses and individual symptoms were rated highly reliable when expressed in specific behavioral terms. They encouraged continued use of a semi-structured interview as a way of obtaining
sufficient information regarding the client. Parental inconsistency in identifying symptoms across interviews was a concern for future research.

Gathering reliable and useful information is critical to the interviewer in establishing a description of the problem(s) which needs to be addressed. This suggested that interviews could be more effective when structured to gather information and that the information may be used effectively when presented in behavioral terms. This approach is similar to the problem identification interviews which behavioral consultants conduct as a way of representing a problem for solution.

**Problem Identification Interview**

The early work in establishing behavioral consultation as a structure for a problem solving process comes primarily from John Bergan and his associates (e.g. Bergan, 1970; Bergan & Tombari, 1975, 1976; Bergan & Kratochwill, 1990). Bergan (1977) provided a comprehensive framework in which problem solving with regard to a client (or student) could occur as part of a structured interview process between a consultant and consultee. The initial interview begins with identification or targeting of the problem(s) and moves systematically to resolution(s).

Several goals are accomplished by the participants as they move through the initial interview structure for problem identification. Kratochwill (1985) reported that in the targeting of problem behaviors, the consultant assists the consultee in describing and identifying the problem(s) of concern. Evans (1985) described the goal of problem identification as involving information gathering,
sorting and interpreting in light of behavioral principles so as to create a true representation of the problematic situation. Martens (1993) stressed the importance of defining the behavior in observable terms, obtaining estimates of how often it occurs, under what condition, and the beginning of on-going data collection for use in evaluating treatment effectiveness. As a key component in the process, baseline data is used as a measure of current performance level. Data to substantiate problem identification can come from several sources. Standardized test results, work samples, and observations are among those useful for data collection. Witt and Elliott (1983) concur with the need for data collection at this stage as one of the important components in the initial interview for problem identification. Inclusion of all components which researchers suggest are essential to problem identification would be unwieldy were it not for the structured interview and standardized checklists devised by Bergan and Kratchowill.

Bergan and Kratochwill (1990, p. 72) offered a structured approach for the initial interview for problem identification by suggesting the following steps:

1. Establish objectives.
2. Establish measures for performance objectives.
3. Establish and implement data collection.
4. Display data.
5. Define the problem by establishing the discrepancy between current performance, as reflected in the data collected, and the desired performance, as indicated in the performance objectives.
Martens (1993) added another step, that of setting up a second interview time indicative of the participant's desire to continue with the process. In order to formalize this approach, Kratochwill and Bergan (1990) developed the Problem Identification Checklist which serves as an interviewer guide through the problem identification interview.

Kratochwill (1985) reported that problem identification is completed when the target behaviors have been established and a treatment goal set by the consultant and consultee. Once the behaviors have been targeted it is understood by practitioners that by identifying the problem in terms that are mutually agreeable to both participants, the problem solving process is anticipated to have a successful outcome.

**Empirical Support for Problem Identification**

Gresham (1984) reported that empirical support for using a behavioral interview as a means for identifying a client's problem(s) has been demonstrated primarily as interrater reliability. Application of interrater reliability in establishing the importance of problem identification leading to successful problem resolution comes from the work of Bergan and Tombari (1975, 1976). They developed a coding system for consultation verbalizations known as the Consultation Analysis Record (CAR). Raters are trained to code interview verbalizations in message units for source (speaker), verbal process (interactions), interview content (behaviors), and control (actions directed by verbalizations). Interviews are coded four times for the presence of data which correspond with the subcategories of each of the four areas. The reliability of the interview is
determined by the consistency of the raters’ coding for specific verbalizations in the appropriate categories. Bergan and Tombari (1975) reported 96% interrater agreement in assigning verbalizations to the specific units of observation.

The CAR coding system was used to establish the verbalizations necessary to accomplish the specific goal for each stage of behavioral consultation. The criteria for establishing problem identification were derived from the behavior specification utterances which specify the behaviors, the settings where they occur, and procedures for measuring observational specifications. Bergan and Tombari (1975, p. 220) stated that in addition to discussing behavior, conditions, and measurement procedures, consultation participants would have to agree as to what has been discussed in each of these areas. Summarization and validational verbalizations would need to be present as an indication of the necessary agreement. The remaining phases of consultation also have specific verbalizations which must occur and be rated in order to validate manifestation of the specific stage in the consultation process.

Bergan and Tombari (1976) demonstrated the effectiveness of their procedure in a study of consultant skill and efficiency. This study involved training 11 psychologists to use a four stage model of problem solving to conduct a total of 806 consultation cases with classroom teachers during the course of a school year. Effectiveness measures were determined by the interviewers’ general efficiency, skill in applying psychological principles, and interviewing skills. The occurrence of each phase of the problem solving process was noted on the case reporting form which indicated from the interviewer whether or not
a phase had occurred. Success was measured by an indication that the goal set in the problem identification phase had been achieved.

Problem identification and problem analysis interviews were used to determine interview skills. This was done using the CAR format. Interrater reliability was established for the content, process and control categories by coding verbalizations from the audio-taped interview. Agreement between two coders on verbalizations from the problem identification and problem analysis interviews was reported with a Scott coefficient of .88 and .92 for control, .87 and .90 for content and 1.00 and 1.00 for process. Three multiple regression analyses were performed on all of the data. A significant main finding was the impact of problem identification on the rest of the process. Over 40% of the variance in the occurrence of problem identification \( (R = .637) \) was accounted for by interview-effectiveness and consultant skills as measured from the CAR. In turn, problem identification accounted for almost 59% of the variance in plan implementation, the next suggested phase in the process. These findings underlined the importance of problem identification and consultant interview skill in bringing this about in order for the rest of the consultation process to continue.

It was interesting to note that the actual identification of the problem was determined by the interviewer responding to the researchers’ query as to its occurrence. Only 43% of the 806 cases received problem identification interviews which then contributed significantly to the regression equation. This suggested that problem identification itself may have been perceived to be important
although fewer than half of the consultants engaged in it. The strongest alternative to the use of the problem identification interview was a testing referral. Other reasons cited for not identifying clients' problems varied but were not extraordinary to a school situation. Bergan and Tombari (1976) reported that the alternatives to problem identification varied predictably with consultant skill and efficiency variables so that future research in this area focused on training consultants as effective problem solvers.

**Clinical Behavior Therapy Interview**

Establishing adequate reliability for interviews is also addressed in the clinical behavior therapy literature. An interrater reliability approach was undertaken in Hay, Hay, Angle, and Nelson's (1979) investigation of the reliability of problem identification in behavioral interviews. Their approach differed from Bergan and Tombari's (1976) by investigating agreement between interviewers rather than rating assigned verbalizations. Hay et al. (1979) accomplished this by attempting to ascertain the number of specific problem areas identified across four interviewers of four clients and by agreement among interviewers for the specific problems identified for each client. An additional purpose to their study was to identify the sources of variance between the interviewers which could potentially lower their agreement. Four clients were interviewed by each of four interviewers for the purposes of conducting a problem identification interview. Following the interview, each interviewer recorded a verbal summary of what had transpired. The content of the interview transcriptions and verbal summaries were coded and rated for problem areas
identified directly or items queried by the interviewers. The criteria for establishing the problem areas came from the Cautela and Upper behavioral coding system. This system listed 25 life areas in which problems could occur with specific items listed for each area to be endorsed by the rater. Problem areas were identified when the name of an area or item within an area was named during the interview and the client indicated the frequency, duration, or intensity of interference during the interview. Raters also coded those areas in which the interviewer asked questions. Interrater agreement on the presence of the problem was necessary for it to be included in the analysis.

Interrater agreements were calculated with the exact percentage agreement method. Mean interrater agreement scores for problem areas identified from the interview transcriptions was .90 and for the verbal summaries .83. The agreement score calculated from the transcript for items identified as problem areas when queried was .87. Agreement scores for areas and items queried during the interview was .85 and .75, respectively.

When considering the number of problem areas identified, clients differed significantly on interview transcripts, $F(3,9)=4.17, p<.05$ and on the verbal summaries, $F(3,9)=10.63, p<.01$. The interviewers were not found to differ significantly in either reporting the number of problems or the problem areas identified. This suggested the possibility of generalizing across interviewers in terms of the overall numbers of problem areas identified during the interview.

Inter-interviewer agreement reported a different finding. Agreement between interviewers with respect to the specific area of problem identification
resulted in a mean agreement score of .55 calculated from the interview transcriptions. For items considered to be problems the mean agreement score per problem area was .40. The mean agreement score between interviewers for problem identification based on the verbal summaries was .48. Further analyses revealed significant differences between the interviewer with regard to the number of problem areas they investigated during the interview, $F(3,9) = 4.02$, $p < .05$. Agreement scores calculated on the type of questions used resulted in a mean score of .62 for broad problems areas and .29 for specific items within the problem area. By comparison, the mean agreement score of the clients' consistency in responding was .86. Further comparisons were done of the verbal summaries to the interview transcriptions which suggested that an average of 28% of the information discussed in the interview was lost in the summary.

The results of this study revealed that interviewers did not identify significantly different numbers of problem areas. Inter-interviewer agreement results report a different finding in which low to marginal agreement was found for the specific areas discussed in the interviews. Consideration of differing interviewer skills, training and bias were cited as potential reasons for the obtained results. A caution was also made as to the use of a recording device for information gathering to avoid potential loss of data. Suggestions made for future problem identification interviews were that using a more structured interview format and conducting a functional analysis as part of the interview be considered.
The findings from Hay et al. (1979) suggested that interviewers, or actual participants, do not share agreement based on the outcomes of their interview which was evaluated by independent raters using a different coding scheme. The findings from Bergan and Tombari (1975) suggested that high levels of interrater reliability came from agreement between raters on interview verbalizations. These findings are of concern because they suggest that interview participants may engage in a process without a clear indication as to the nature of a problem they are seeking to address.

The implications of these results and their comparisons have an impact on the purpose of the behavioral interview and its use as an assessment instrument. It suggested that if participants do not agree on the problem areas discussed there would be a negative impact on targeting a behavior and choosing an appropriate intervention for solution. This would suggest that further study of agreement between participants and raters may clarify how problems are identified at this early stage in problem resolution.

Wilson and Evans (1983) conducted an initial investigation of the reliability of the process of target behavior selection. Their study addressed the question of how treatment goals were formulated in the initial stages of assessment. One hundred eighteen respondents completed questionnaires comprised of three case studies of children exhibiting severe problems of varying complexities. Participants responded to open-ended questions regarding their impression of the child's identified problem, whether treatment was warranted, an indication of treatment goals and, stated treatment targets in ranked order. Results were
reported as interrater agreements to the responses with two independent raters and a third who was blind to the purpose of the study.

The finding of this study which impacted upon the issue of reliability of problem identification was that the overall percentage agreement of 38.6% reflected moderate agreement among the participants as to their general impression of the nature of the highest priority problem presented. This was accomplished by determining the mean of the percentages of agreement of each participant on their selection of the highest priority problem targeted for resolution for each case. This finding was considered low for intersubject agreement in choosing the priority target behaviors. Among considerations given to future research was that a more thorough investigation of factors which influence agreement be made. This was an analogue study which used an open response questionnaire. It is uncertain whether the findings can generalize to naturalistic settings. It does however, lend support to the Hay et al. (1979) finding of low agreement between participants regarding the identification of the nature of problems.

A study which addressed reliability through agreement between assessors was conducted by Felton & Nelson (1984). Their purpose was to examine assessor agreement with regard to identifying hypothesized controlling variables in conducting functional analyses and designing treatment plans. Six assessors were randomly assigned to one of two groups which were categorized as interview only or interview with role-play and questionnaires. All were required to conduct a comprehensive behavioral interview on the same three "clients."
Problem definitions were supplied in order to focus attention on identifying controlling variables. Following the interviews, each assessor's list of controlling variables and treatment proposals was compared with the other two from their group and interassessor agreement was determined using a percentage of agreement formula. Interrater agreement, by two graduate psychology students' independent judgements, using the same formula was determined to be .90 and followed a correction procedure for disagreements. Decisions were made in two-steps: first, by initial independent agreements between judges and then their discussed agreements following from their disagreements.

Scores for mean inter-assessor agreement across clients, variables and treatment plans was .41 for the interview-only group and .40 for the group using multiple devices. Individual variable agreements were reported for the two groups ranging from a mild agreement of .21 and .24 for organism variables (past history, physiological problems) to greater agreement on treatment proposals of .59 and .62. The researchers reported that the major factor they found to have affected low agreement among assessors was the differences found in the specific questions asked by each of the assessors.

These findings also support those of Hay et al. (1979). They suggest low reliability when comparing agreement between participants rather than agreements between the raters. Considerations put forth by the researchers were to seek ways of improving agreement among assessors, in particular through training of interviewing skills. An additional consideration was suggested as to the importance of agreement in targeting behaviors. Were participants'
agreement demonstrated not to be of importance, there would be no concerns as to deriving interventions based on their support for establishing treatment integrity.

**Interviewer Training**

Follow up on the findings of Bergan and Tombari (1975, 1976) and Hay et al. (1979) has been demonstrated by the number of studies which approach the problem from a training perspective. When considering the findings of the Hay et al. (1979) study Brown, Kratochwill, and Bergan (1982) addressed the issue of training interviewers with a specific set of skills designed to reduce interviewer variability in questioning. This study introduced a standard format for training in interviewing skills designed to elicit problem definitions. A secondary purpose was to test a standardized interviewing skills training package.

Four graduate students were trained in the model demonstrated on three students who role-played clients. Three observers were trained to use the Problem Identification Checklist and rate consultants' verbalizations accordingly. Interobserver agreements were calculated from observers one and two, observer three's agreements (calibrating observer) were calculated over one third of all the sessions across all phases. Interobserver agreements were calculated for each of the four categories and subcategories of verbalizations. The range over four trainees for the behavior category was .77 to .98; behavior setting range was .82 to .86, the observation category was .98 to 1.00 and summarization category ranged from .78 to 1.00. The findings support Bergan and Tombari’s (1975) claim of reliability for problem identification when determined by interrater reliability,
and demonstrate that training can be effective in improving interviewer skills. Brown et al. (1982) concluded that the issues of agreements across interviewers surfaced by Hay et al. (1979) could be accounted for in the standardization of interviews thus preventing the loss of relevant data. Brown et al. (1982) continued to address the reliability question in terms of interviewer verbalizations assessed by raters on their problem identification measure. They did not address the issue of agreement among interviewers (or interviewees) as to the content of their discussions.

A similar study was conducted by Duley, Cancelli, Kratochwill, Bergan and Meredith (1983) which addressed the training issues of interview skills as they generalized to a motivational analysis interview (identification of reinforcements) using the training package described in Brown et al. (1982). This study also responded to the lack of a functional analytic approach in Hay et al. (1979). Interviewing was conducted in an analogue situation. Observers received training to code for verbalizations. The results obtained were the probability of inter-rater agreement across baseline (.95), posttraining (.96) and generalization (.88) phases of the study and analyzed with Bergan’s quasi-equiprobability model. A social validational component was added and assessed using three expert judges and giving interviews an additional rating to support analysis of the skill development by training, interjudge agreements and correlation between ratings and target skills. In the final analysis 88% of the variance was shared between the percentage of skill improvement demonstrated and social validation ratings across the participants and interviews. This study continued to
support the findings of Brown et al. (1982) on the importance of skill training for interviewers. It demonstrated that generalization of skills over time was possible in an analogue situation. Although it was an attempt to respond to an element missing from the Hay et al. (1979) study, the functional analytic approach, it did not directly address the issue regarding agreement of participants as to the nature of the problems discussed in their problem identification interview.

Further support for training and generalization of skills to other disciplines was demonstrated in the study conducted by Keane, Black, Collins & Vinson (1982). Clinical pharmacy students received training in behavioral interviewing skills in order to conduct interviews with hospital patients. The purpose of using a behavioral approach was to demonstrate the effectiveness of gathering information using functional analysis regarding patient compliance with medication. Students were randomly assigned to three treatment groups which consisted of two levels of training and a control group. Interview skills were taught and assessed with regard to content and style. Content referred to patients’ medical histories and style to interviewers’ use of specific skills. Reliability was determined by interrater agreements using the percentage agreement formula for content area occurrences. Pearson correlations were used for style components. The results supported the group which received behavioral rehearsal training and practice to improve their interview skills. Comprehensive training enabled them to increase significantly their use of open-ended questioning as determined by a repeated measures ANOVA of significant groups X time interaction, F(2,32)=9.13, p<.01). Further analysis at the generalization
phase revealed a significant reduction in the use of close-ended questions, 
F(2,31) = 4.02, p < .05) and the amount of interviewer speaking time, F(2,31) = 3.70, 
p < .05). In conclusion, Keane et al. (1982) suggested that training was essential 
to effective interviewing, had potentially good generalizability to other related 
fields, and when divided into content and style areas may have more interactive 
influence then first thought. Keane et al. (1982) reported that Hay et al.’s (1979) 
results may be reflected in the variability of content and style. This study 
suggested that emphasis may be placed on the interview content and interviewer 
style. Following training interviewers were able to generate more content 
material but this was only assessed by interrater agreement and not by 
agreements with the participants which is what the Hay et al. (1979) study 
suggested.

Consultee Problem Identification

A somewhat different focus was offered by the following two studies. 
They approached consultation problem identification from the perspective of the 
consultee. Curtis and Watson (1980) assessed changes in problem clarification 
skills of consultees following consultation interviews with high-skilled and low-
skilled consultants. Twenty-four classroom teachers were assigned randomly to 
one of eight consultants. Consultants were assessed for their skill levels 
following initial neutral interviews. Further skill training was given to already 
highly skilled consultants to maximize the differences. Following three 
consultation sessions, the consultees working with the highly skilled consultants 
were found to have improved problem clarification skills. Results were based on
the verbalizations scored using Bergan’s (1977) Consultation Analysis Record and a Problem Identification Checklist designed to assess consultees’ problem clarification skills. The areas of problem identification which did not reach significance were the index of content relevancy and percentage of factual utterances. The results offered further support to the importance of problem identification skill training but suggested, where improvement was not significant, that some skills are already part of a teacher’s repertoire without formal training. These skills, such as the component of data collection for problem solving, did not demonstrate any improvement following consultation training. This study marked a shift in focus from the consultant’s problem solving to that of the consultee in response to the consultant’s training. The researchers attempted to describe the consultees’ ability to identify or clarify problems of actual students in the problem identification phase of consultation. The results suggested that consultees’ skills required further consideration particularly when attempting to clarify a problem and their input at this level may be a critical factor in identifying their students’ problems. This study was also important as it involved actual consultants with classroom teachers discussing their attending students.

The primary goal of Cleven and Gutkin’s (1988) study was to increase consultee’s skills in problem solving using a cognitive modeling approach. Their procedure involved training university students to write problem definition statements after viewing a demonstration of various levels of interviewing on a series of videos. Reliability was established by ratings made of 10 pilot cases. Of
interest in this study is the use of 6 raters for interrater reliability who were blind to the nature of the study and the group placement of the participant. Dependent measures of Behavioralness, Goal, and Process scales were each rated with a three-point criteria from the Problem Definition Description Questionnaire. Behavioralness required inclusion of multiple behavioral examples stated in concrete, observable, and measurable terms. The Goal scale rated the presence and clarity of a goal statement. The Process Scale required that participants develop a concrete behavioral problem definition, components must be prioritized with a goal statement included. Levels of scoring criteria were given for responses which included all components with varying degrees of clarity. Percentage agreement was determined by checking the raters’ scores against the researcher’s. Reliability coefficients were reported for the scales of Behavioralness (.98), Goal (.92) and Process (.92) Additional reliability checks were done by having the each evaluator rate five additional protocols and their ratings were compared with the original. Subsequent percentage agreements were Behavioralness = .94, Goal scale = .95, and Process scale = .92.

The results were supported with qualitatively better behavioral problem descriptions written by those who viewed the cognitive model of problem solving. This meant that their problem identification statements contained specific relevant behavioral examples in observable and measurable terms and that their statements were goal-directed. This study endorsed cognitive modelling as an effective training mode for consultees in the consultative process. It suggested that improved skills may also contribute to problem
identification in problem solving. The training of skills directed toward the consultee suggests that effective problem identification could reflect agreement between participants given that they were able to define the problems in terms which would be useful to both when initiating the process.

**Current Practices**

Current literature speculated on the continued use of a problem solving approach in behavioral assessment/consultation. White and Edelstein (1991) noted that research support for behavioral assessment interviews was limited. Their concern was addressed in support for the investigatory interview which closely resembled the format followed in behavioral assessment and consultation. They suggested the use of accuracy as a determinant for interview reliability. They also endorsed the practice of training interviewers in the skills needed to establish inter-interviewer agreement.

Nezu and Nezu (1993) suggested a problem-solving sequence as a potential solution to the on-going concern of identifying and selecting target behavior on which to focus intervention activities. They described an approach which recognizes multiple causalities as a back drop for choosing more appropriate target behaviors. Their goal in target selection was future-oriented for effective functioning. Included in their approach is the use of the SORCK model for assistance in identifying target behaviors and their relational variables. This approach and model await empirical support. This model considered the dynamic dimension of the interview process and acknowledged difficulties when considering the contextual situation in which problem solving occurs.
Summary of the Problem

The purpose of this study was to address an issue of empirical support for the interview as an assessment instrument. Problem identification established in the initial interview was reported to be a critical component in reaching a solution. Reliability for this process has been demonstrated in the behavioral assessment/consultation literature using the standardized format introduced by Bergan (1977) and Bergan & Kratchowill (1990). Interrater reliability was established for the structured interview process by comparison (agreement) of ratings of participant verbalizations made during the interview using a standardized coding scheme. The clinical behavior therapy literature represented by Hay et al. (1979) reported strong interrater agreement on the problems discussed in an interview but could not establish inter-interviewer agreement on the nature of problem areas discussed in an initial interview situation.

Bergan and Tombari’s (1975) findings regarding the importance of problem identification have led to the development of structured interviews and coding of verbalizations to support problem identification. This approach has helped to establish interview reliability based on interrater scoring. Variability, a threat to reliability, was reduced by training raters who were able to code transcriptions with near-perfect accuracy to meet the requirements of the structured format. Strong support for training interviewer skills as another way of reducing variability was demonstrated in the literature. Training was also recommended for consultees contributing additional support for this approach.
Subsequent studies reporting inter-interviewer or inter-participant agreement were not able to demonstrate strong agreements. Trained raters achieved strong interrater agreements for these interviews, however, the participants themselves did not seem to agree on the content areas of their interview, in fact, this approach was rarely considered in the literature. This suggested that participants may not reach an agreement on the nature of the problem identified in the interview. Consequently the problem may not be identified during the interview and, in turn, potentially interfere with the problem solving process.

The present study attempted a response to the reliability issue by assessing the level of agreement of participants as a way of establishing support for successful problem identification in the initial interview.
CHAPTER THREE: DESIGN AND METHODOLOGY

Nature of the Study

This study was designed to focus attention on the initial interview component of behavioral consultation. The study assessed the reliability of problem identification in the initial consultation interview by obtaining a measure of agreement from the participating classroom teacher and school-based consultant (learning assistance teacher) as to the nature of the problem they discussed. Interrater agreement was also assessed using two independent ratings of the interview agreement results to determine if the participants agreed upon the nature of the problem considered to be the highest in priority for the student discussed. The study followed a descriptive nonexperimental design in which participants were asked to engage voluntarily in a consultative interview about a student whom the teacher would characterize as difficult to teach. In order to facilitate generalization of the results of this study to consultation practice in a naturalistic school environment, the study was conducted in the participants' school and a real student's problems were discussed. Participants were asked to conduct the interview in the same manner in which they usually address problematic situations regarding their students. The researcher videotaped the interview.

Immediately following the consultation interview, participants were asked independently for a verbal description of the nature of the problem discussed in the interview and to indicate the order of priority given to each one, if more
than one was discussed. Each participant was asked to provide a rating of the extent to which the interview was helpful in identifying the student’s problem(s), the extent to which the identified students’ problems were adequately identified, and the extent to which the interview was successful in achieving a shared understanding of the problem associated with the identified student. Participants were asked to complete a demographic informational questionnaire and problem identification evaluative checklist to aid in the evaluation of the interview process. The identification evaluation measure used was an adaptation of one used in a program evaluation model reported by Knoff (1982a, 1982b) and was similar in style to the parent interview conducted by Graham and Rutter (1968).

The tape recorded interview of the participants’ verbal response to the researcher’s question was transcribed for an evaluation of the representation of the problem and for rating agreement between the consultant and consultee as to the nature of the student’s problem. The researcher and a trained research assistant coded each transcription for a level of agreement between the learning assistance teacher and classroom teacher on their independent description of the problem(s) identified during their problem identification interview. The two raters also provided ratings using the same evaluation form as the participants. This procedure of data collection and analysis was similar to that followed by Hay et al. (1979), Felton and Nelson (1984), and suggested by Nezu and Nezu (1993).
Procedure

Participants

The participants in this study were 45 educators currently employed within the separate school system, the Vancouver Archdiocese Catholic Schools, a member of the Federation of Independent Schools Association of British Columbia. There are currently 45 schools operating in the Archdiocese. Recruitment was solicited by an explanatory letter sent to each principal inviting voluntary participation of individual school’s consulting teacher/learning assistance teacher and 4 classroom teachers. Additional support in seeking voluntary participation was offered by the Office of the Superintendent of the Vancouver Catholic Schools. Agreement to participate in the study was formalized by each participant’s signing a letter of informed consent. Participants’ signatures indicated their willingness to take part, having been informed of the purpose of the study, that all information shared during the research study would be held in confidence, and that they were free to withdraw from the study at any time. During the period of the consulting teachers’ participation, researcher substitution was offered, but not required, for any students scheduled to be seen by the consulting teacher at the time of the interview.

Nine volunteer consulting teachers (learning assistance teachers) representing nine schools situated in the Catholic Archdiocese of Vancouver participated. Their participation required that they conduct problem identification interviews and respond to questions regarding the process and outcomes of each
consultation interview. Consulting teachers conducted interviews with four classroom teachers from their schools. Each consultation interview was focused on a student who was then experiencing difficulties in the participating teacher's classroom. While the classroom teacher participated in the consultation interview, the researcher provided a certified substitute teacher through a prearrangement with the individual school's principal and learning assistance teacher.

Setting

All interviews and data were collected at the participants' individual schools. Typically these interviews were conducted in the learning assistant's room, the usual setting for such contacts, and a location offering the required level of privacy. The video camera was set up by the researchers in close proximity to the participants during the interview, but the researcher was not present during the actual interview. The interview usually lasted between 25 and 35 minutes, not including the amount of time need for prior instructions to be given and post-consultation interviews and written measures to be completed.

Assignment of Participants:

Consultants and Consultees

The consultants (learning assistance teachers) requested 4 classroom teachers (consultees) employed at their school to participate voluntarily in an initial consultation interview to discuss one student (client) in each teacher's class. The identified student was one whom the teacher described as difficult to teach and about whom the participants had had no previous formal discussions.
Prior to each interview, the researcher met with both participants simultaneously to review the procedures. At this time an overview was given as to what would be expected by their participation in the study. Each participant received a copy of all the measures he/she was expected to complete. Participants were directed to complete the demographic questionnaire before the start of the interview and the evaluation questionnaire at the end of the interview, when alternating with the post-consultation interview. Directions were read to the participants and clarified by inviting participants to conduct the interview just as they would typically do so. (A copy of the directions is included in Appendix #1.) They were encouraged to focus on describing what about the student was currently causing him/her to be difficult for the classroom teacher to teach.

Once the participants had completed the demographic questionnaire and indicated they were ready to begin, the researcher turned on the video camera and left the room. At the conclusion of the interview, the researcher was summoned back into the room. The post-consultation interview was conducted typically with the classroom teacher first while the learning assistant teacher completed the evaluation form in another room. Once the first interviewee completed the post-consultation interview, he/she left the room to complete the evaluation measure. The learning assistant returned and was administered the post-consultation interview at that time. Completed evaluation measures were returned to the researcher soon after concluding the post consultation interview.
Instrumentation

Measures:

1. Demographic Questionnaire

An information questionnaire was administered to the learning assistance teacher and to the classroom teacher. The information requested was the participants' age (range), sex, level of educational degree, number of years teaching, current title, and prior knowledge of, or training in, consultation practices. This information was used for descriptive purposes. The participants' responses are reported in percentages, categorized as learning assistance teachers (consultants) and classroom teachers (consultees). (A copy of the participant demographic information form is included in appendix #2.)

2. Problem Identification Interview Evaluation Scale

This measure required the consultation participants to evaluate their interview session in terms of consultation process efficacy, adequacy of problem identification, and shared understanding of the nature of the problem. Responses were endorsed by strength using a four-point Likert-type scale in a format similar to that reported by Graham and Rutter (1968).

This measure was intended to be completed independently by the participants following their consultation interview. In the majority of cases, the classroom teachers completed their evaluation scale after the post-consultation interview with the researcher. The learning assistance teacher completed the evaluation scale in another location while the classroom teacher was being interviewed by the researcher.
The first section of the written evaluation was divided into two parts. Part I restated to the participants the main goal of a problem identification interview which was to formulate a comprehensive description of the major problem or problems. Participants were then asked to respond to three questions regarding focus of the interview as a useful assessment measure of problem identification. The participants were to endorse the level at which they found the interview to succeed in reaching its goal for the particular variable stated. A four point Likert-type scale was used for each question. The first point on the scale was to be endorsed if the topic in question did not occur at all, or was not addressed at all. The second point was endorsed if it was somewhat addressed, the third point required that it be mostly addressed, and finally, the fourth point was interpreted as completely addressed.

**Interview Helpfulness**

The first question requested participants to rate the level at which they found the interview to be helpful in identifying the student's problem. This question was selected to be an indication of the usefulness of the interview, a measure of consumer satisfaction. The ratings offered for endorsement ranged from a "One" (Not At All Helpful) to a "Four" (Completely Helpful).

**Problem Identification**

The second question required an outcome evaluation of the extent to which the participants believed they had adequately identified the student's problem(s) as a result of the interview. This question evaluated the extent to which problem identification was perceived to have occurred by the participants.
Responses were rated at four points ranging from "One" (Not At All Identified) to "Four" (Completely Identified).

Shared Understanding/Agreement

The third question addressed the main goal of the study which was to identify the participants' agreement as to the nature of the problem(s) discussed in the interview. The question required the participant to rate what he/she believed to be the level of shared understanding with his/her interview partner. The four point responses for this question ranged from "One" (Not At All Understood) to "Four" (Completely Understood).

Problem Areas in Priority

Part II of the evaluation required participants to recall up to three of the problems discussed in the interview beginning with the problem they considered to rank the highest priority in their discussion. These rankings in priority were similar in nature to those used by Cleven and Gutkin (1988) in their Problem Definition Questionnaire. The problems cited were to be categorized primarily as either academic or social/emotional/behavioral. General descriptors were given for both categories in a checklist format in which participants were asked to endorse how they perceived the problems to be interfering. Three areas of descriptors were offered for problems which were academic in nature: content area deficit, production deficit, specific skill deficit. The descriptors offered for the social/emotional/behavioral - identified problems were: social skill deficit, behavioral excesses, behavioral deficits, and personality variables. Brief examples were cited with each descriptor to assist in matching the problem descriptor with
the problem presented during the interview. Three separate sheets, indicating the rank of the problem to be reported, were provided, one for each problem. Interview participants completed their evaluations after their interview session.

**Additional Problems**

A final section of the evaluation requested the interview participants to identify further any other problems which they believed to be present with the student but not mentioned during the consultation session. Those who endorsed the presence of unreported problems were asked to specify if the problems were missed or ignored during their interview session. (A copy of the entire interview evaluation scale is provided as appendix #3.)

**Analysis of the Evaluation**

In order to assess the levels of agreement reached between the participants, the kappa statistic was used to determine overall agreement between the learning assistance teachers and classroom teachers on their response ratings for each of the three areas. Kappa is the statistic most often used to calculate agreement in preference to the percentage agreement formula (agreements/agreements + disagreements x 100) typically used in behavioral assessment. The use of Cohen’s Kappa was suggested by Lee and Suen (1984) as it corrects for the possibility of agreements by chance which the percentage agreement cannot do and therefore may report inflated results. Kappa statistics reported in this study were calculated using Cohen’s formula as suggested in Suen and Ary (1989) and Lee and Suen (1984). Interpretation of the level of acceptability for agreement percentages was offered by House et al. (1981) as
those generally accepted by researchers in behavioral sciences. Average agreement at or above 70% (.70) is considered necessary for agreements, above 80% (.80) is adequate and above 90% (.90) is good.

Further analyses were conducted using chi-square analyses of group responses by considering the association between the participants and their responses. Significant results were reported with alpha set at < .10 rather than using the more conventional standard of < .05 so as to equalize the balance between Type I and Type II errors. Suggested by Cascio and Zedeck (1983), this is a way of increasing power for a small sample size. Contingency coefficients are reported as an indication of the strength of the relationship between the raters and their status of ratings.

3. Agreement Indicator: Post-Consultation Interview

Following the consultation each participant was interviewed by the researcher or an assistant. In this interview each participant was to describe what they believed to be the nature of the student’s problem. The participants were asked to summarize the content of their consultation interview and present the problems discussed in order of priority. Although the post-consultation interviews were conducted independently, the classroom teacher was typically interviewed first, immediately following the consultation session. Participants completed the remaining evaluation forms while waiting to be interviewed by the researcher or immediately after their post-consultation interview with the researcher.
Responses from each participant were tape recorded and later transcribed. Two raters reviewed the written transcripts to code the post consultation interviews for mutual agreement between respondents' descriptions of identified problems, particularly the problem which was given the highest priority. In those cases where respondents identified multiple problems, the level of participant agreement was coded for the problem described by the participant as of highest priority.

Independent Interview Evaluator:

Raters

Two raters, the researcher and another rater, independently examined the transcriptions of the participants' responses collected at the end of the interview for agreement as to the nature of the identified problem(s). In addition to an overall rating of the post consultation interview, the raters completed an evaluation form identical to the second part of the form completed by the participants identifying the number and nature of the problems recalled by the interview participants. Both raters were knowledgeable of the goals and procedures of the interview format in the consultation process. Approximately two hours of training was provided by the researcher as to procedures for identifying targeted problems and scoring for agreement between consultant and consultee through the use of pilot transcriptions. Training consisted of viewing a sample interview video tape, listening to the audio-taped responses to the pilot participants' individual post-consultation interview responses and rating sample interview transcripts of the interview dialogue between the researcher and the
participants. Individual responses to the pilot transcriptions were compared to assure agreement between the raters prior to their individual rating of collected data.

**Agreement Ratings**

The ratings of agreement were based on the learning assistants’ and classroom teachers’ description of the student’s problems described in their independent post-consultation interview. Two trained raters independently coded/rated the level of agreement reached by the learning assistant and classroom teacher. The rating was accomplished in a two step process.

The first step was a rating of the individual post-consultation interviews for the nature of the problem with subsequent descriptions, the number of problem areas identified and the priority given when more than one problem was identified. Each of the 72 post-consultation interviews were selected in random order for this evaluation. Interview tapes with the accompanying transcription were rated independently by the two raters by completion of the problem identification forms, with descriptors, for up to three problems in priority. These were the same problem identification evaluation forms completed by the participants after their consultation session.

Use of the audio-tape for the post consultation interview and rating of the transcriptions was based on the description provided by Hay et al. (1979) of their assessment for problem identification in an initial interview. The written problem identification measure was similar to the measure of consultation evaluation
developed by Knoff (1982a, 1982b) for a retrospective assessment of problem identification in a clinical interview.

The second step in the rater evaluation required the two raters, independently, listen to and read the transcript of both participants' post consultation interviews about the same student. The raters evaluated the level of agreement between the descriptions provided by the two participants. Each rater coded the interview dyads on a four point scale of "One" (Complete Disagreement), "Two" (Mostly Disagree), "Three" (Mostly Agree), and "Four" (Complete Agreement). Complete agreement indicated that both participants described the same problem with the same descriptors in the position of highest priority problem in their post-consultation interview. If other problems were discussed, the priority given them by both participants also matched. (See appendix #5 for an example of a transcript which was rated as having complete agreement.) A rating of mostly agree received a level three rating based upon participant agreement of the nature of the problem with the highest priority but with dissimilar descriptions or lack of descriptions to support their choice of the problem. Typically the first and second priority problems were reversed in their endorsements. Also, other indicated problems did not reflect the same order of priority. (See appendix #6 for a transcript which was rated as "mostly agreed.") A rating of mostly disagree, which received a level 2 rating, was described as different problems receiving the highest priority endorsement and any remaining problems did not share the same priority level. Any descriptors for the problems were completely dissimilar. A rating of complete disagreement which received a
level 1 rating was described as completely dissimilar problems receiving the highest priority position and no indication of agreement in the nature, number, priority ranking or descriptions was evident. (Examples of level 1 and 2 ratings were not available from the transcripts since no interview dyad received those endorsements from the raters.)

Summary

The purpose of this chapter was to describe the nature of this study of participant agreement on problem identification in the initial interview process of consultation. The study described the current practice of problem identification within the initial interview of consultation used by participants from nine Catholic Schools. Consultants and consultees reported their shared understanding (interparticipant agreement) of the nature of a student’s problems following their consultation interview and completion of evaluative measures. Raters independently rated the post consultation interviews for participant agreement on the nature and number of the problems identified.
CHAPTER FOUR: RESULTS

This chapter presents the results of the study conducted on participant agreement between the consultant (learning assistance teacher) and consultee (classroom teacher) regarding the nature of school-based problems of a client (student) who was described as difficult to teach.

Demographics

Participants

Learning Assistants/Consultants

The participants (n=45) of this study were nine learning assistance and thirty-six classroom teachers from nine schools of whom volunteer participation had been requested. The learning assistance teachers conducted individual problem identification interviews with each of four classroom teachers from their schools.

Table 4.1 summarizes the demographic information obtained from the consultant/learning assistance teachers. All nine learning assistants were female, most in the age range of 31 - 40. Six of the learning assistants held Bachelor of Education degrees and two held Master’s degrees. Their mean number of years in the field of education was 11.55 (SD=4.79). The learning assistants served students from primary to secondary grades.
### Table 4.1: Participant Demographic Information

**Learning Assistance Teachers/Consultants**

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>%age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Assistance Teachers</td>
<td>09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>09</td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under 30</td>
<td>01</td>
<td></td>
<td></td>
<td>11.1</td>
</tr>
<tr>
<td>31 - 40</td>
<td>05</td>
<td></td>
<td></td>
<td>55.5</td>
</tr>
<tr>
<td>41 - 50</td>
<td>03</td>
<td></td>
<td></td>
<td>33.3</td>
</tr>
<tr>
<td>Level of Educational Training (N=8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BEd</td>
<td>06</td>
<td></td>
<td></td>
<td>75.0</td>
</tr>
<tr>
<td>MA</td>
<td>02</td>
<td></td>
<td></td>
<td>25.0</td>
</tr>
<tr>
<td>Years in Education</td>
<td></td>
<td>11.55</td>
<td>4.79</td>
<td></td>
</tr>
<tr>
<td>Level of Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>02</td>
<td></td>
<td></td>
<td>22.2</td>
</tr>
<tr>
<td>Secondary</td>
<td>01</td>
<td></td>
<td></td>
<td>11.1</td>
</tr>
<tr>
<td>Primary/Intermediate</td>
<td>05</td>
<td></td>
<td></td>
<td>55.5</td>
</tr>
<tr>
<td>Intermediate/Secondary</td>
<td>01</td>
<td></td>
<td></td>
<td>11.1</td>
</tr>
</tbody>
</table>

**Classroom Teachers/Consultees**

Thirty-six classroom teachers participated as consultees and responded to the demographic questionnaires as summarized in Table 4.2. One half of the classroom teachers were under the age of 30; their mean number of years in the education field was 9.27 (SD=7.73). Twenty-three teachers (63.8%) held Bachelor of Education degrees. Primary through secondary level students were served by the classroom teachers who participated. Teachers of primary level students represented 44.4% (n=16) of the classroom teachers. Teachers of intermediate students comprised 25% (n=9) and 19.4% (n=7) taught on the secondary level.
Table 4.2: Participant Demographic Information
Classroom Teachers/Consultees

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom Teachers</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>31</td>
<td></td>
<td></td>
<td>86.1</td>
</tr>
<tr>
<td>Male</td>
<td>05</td>
<td></td>
<td></td>
<td>13.8</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under 30</td>
<td>18</td>
<td></td>
<td></td>
<td>50.0</td>
</tr>
<tr>
<td>31 - 40</td>
<td>07</td>
<td></td>
<td></td>
<td>19.4</td>
</tr>
<tr>
<td>41 - 50</td>
<td>08</td>
<td></td>
<td></td>
<td>22.2</td>
</tr>
<tr>
<td>Over 50</td>
<td>03</td>
<td></td>
<td></td>
<td>8.3</td>
</tr>
<tr>
<td>Level of Educational Training</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BA</td>
<td>06</td>
<td></td>
<td></td>
<td>16.6</td>
</tr>
<tr>
<td>BEd</td>
<td>23</td>
<td></td>
<td></td>
<td>63.8</td>
</tr>
<tr>
<td>BA + credits</td>
<td>06</td>
<td></td>
<td></td>
<td>16.6</td>
</tr>
<tr>
<td>MA</td>
<td>01</td>
<td></td>
<td></td>
<td>2.7</td>
</tr>
<tr>
<td>Years in Education</td>
<td></td>
<td>9.27</td>
<td>7.73</td>
<td></td>
</tr>
<tr>
<td>Level of Students</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>16</td>
<td></td>
<td></td>
<td>44.4</td>
</tr>
<tr>
<td>Intermediate</td>
<td>09</td>
<td></td>
<td></td>
<td>25.0</td>
</tr>
<tr>
<td>Secondary</td>
<td>07</td>
<td></td>
<td></td>
<td>19.4</td>
</tr>
<tr>
<td>Primary/Intermediate</td>
<td>03</td>
<td></td>
<td></td>
<td>8.3</td>
</tr>
<tr>
<td>Intermediate/Secondary</td>
<td>01</td>
<td></td>
<td></td>
<td>2.7</td>
</tr>
</tbody>
</table>

Participants with teaching assignments on combined levels comprised 11% (n=4).

Prior Experience with Consultation

Participants were questioned as to their previous experience, training in and rating of consultation as part of their occupational function. The learning assistance teachers’ responses are reported in Table 4.3. Of the eight
Table 4.3: Participant Familiarity with Consultation

Learning Assistance Teachers (N=9)

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal Training in Consultation Practices (N=8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>02</td>
<td>25.0</td>
</tr>
<tr>
<td>No</td>
<td>06</td>
<td>75.0</td>
</tr>
<tr>
<td>Prior Experience with Formal Consultation Practices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>02</td>
<td>22.2</td>
</tr>
<tr>
<td>Some</td>
<td>02</td>
<td>22.2</td>
</tr>
<tr>
<td>Frequent</td>
<td>05</td>
<td>55.5</td>
</tr>
<tr>
<td>Rating of Prior Consultation Experience (N=7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>06</td>
<td>85.7</td>
</tr>
<tr>
<td>Neutral</td>
<td>01</td>
<td>14.2</td>
</tr>
</tbody>
</table>

respondents, two (25%) had been formally trained in consultation through workshops and/or university courses, the remaining six (75%) had no previous formal training. (Two reported informally that they had gotten information on consultation from colleagues and knew about it through professional development seminars but did not feel their experience warranted formal endorsement.) Seven (77.7%) of the nine learning assistants had prior experience in practicing consultation at their schools and five (55.5%) reported this as their usual approach to problem solving with colleagues in their schools. Two of the learning assistance teachers had not used consultation in a formal manner prior to their experience in this study.

In rating their prior consultation experiences on a three-point scale: positive, neutral, and negative, six learning assistants (85.7) endorsed a positive
rating and one (14.2%) endorsed the neutral rating. No one endorsed the
negative rating.

Table 4.4 reports the results of the classroom teachers responses to queries
regarding their prior experience with consultation. Seven classroom teachers
(19.4%) had received some formal training in consultation practices either
through workshops or professional development seminars. Twenty-nine
classroom teachers (80.5%) reported no formal training. Prior experience with the
practice of consultation was reported by twenty seven classroom teachers
(74.9%), ten of whom (27.7%) reported frequent experiences with this method of
practice.

Table 4.4: Participant Familiarity with Consultation

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>N</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal Training in Consultation Practices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>07</td>
<td>9.4</td>
</tr>
<tr>
<td>No</td>
<td>29</td>
<td>80.5</td>
</tr>
<tr>
<td>Prior Experience with Formal Consultation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>09</td>
<td>25.0</td>
</tr>
<tr>
<td>Some</td>
<td>17</td>
<td>47.2</td>
</tr>
<tr>
<td>Frequent</td>
<td>10</td>
<td>27.7</td>
</tr>
<tr>
<td>Rating of Prior Consultation Experience (N=33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>22</td>
<td>66.6</td>
</tr>
<tr>
<td>Neutral</td>
<td>11</td>
<td>33.3</td>
</tr>
</tbody>
</table>

On the three point scale rating the prior experiences as positive, neutral or
negative, 33 classroom teachers responded in total, two-thirds of the respondents
endorsed the positive rating and one-third gave it a neutral endorsement. No one endorsed the negative rating.

Summary

The study sample included 9 learning assistance teachers and 36 classroom teachers whose combined mean of years in the field of education was 10.41 years (SD = 1.61). Seventy-five percent of the learning assistance teachers, who conducted consultation interviews as consultants, had not received formal training in consultation practices. Only half of the learning assistants used consultation as a problem solving approach. Two of the learning assistants had never used consultation prior to participating in the study. Eighty percent of the classroom teachers, who participated as consultees, had not received any training in consultation and approximately 25% did not use consultation.

The participants were relatively untrained in the use of consultation. Few participants reported frequent use of this format for problem solving but those who did use consultative problem solving reported moderately positive past experiences. This was not an "expert" sample of consultation practitioners.

Psychometrics

Participant Interview Evaluation Measure: Part I

Both members of the Learning Assistant-Classroom Teacher interview dyad completed the three item Interview Evaluation: Participants scale immediately following their consultation interview or post-consultation interview with the researcher. (A copy of the scale is included in Appendix #3.) This measure required the participants to respond using a four point, Likert-type
scale, to each of three questions regarding three aspects of their consultation interview. Responses to the three questions reported the extent to which the process was helpful, how thoroughly the problem was identified, and the level of understanding of the problem shared between the participants. Four rating levels were given for each of the three questions and scoring of the levels was similar for each question. A "1" response meant no occurrence or a complete disagreement with the topic in question, "2" represented a choice of somewhat of an occurrence of the topic or moderate disagreement, "3" was a stronger endorsement for the occurrence of the behavior and moderate agreement, and "4" was the strongest endorsement for the presence of an occurrence and/or complete agreement.

Table 4.5 reports the means and standard deviations of the participant responses to the three interview evaluation questions by individual schools. The overall mean rating of learning assistants on interview helpfulness was 3.22 with a standard deviation of 0.59 and the classroom teachers' mean rating was 3.14 with a standard deviation of 0.80. Responses across schools showed considerable variability within a range of 3.00 - 3.75 for learning assistants and 2.75 - 3.75 for classroom teachers. One third of the schools showed exact mean and standard deviation score agreement between dyad participants. Four of the learning assistance teachers gave the same level three rating to each of her four interviews while the classroom teachers who participated in these interviews varied in their rating of interview helpfulness.
Table 4.5: Means (and Standard Deviation) of Learning Assistance and Classroom Teacher Ratings of Interview Characteristics Across Schools

<table>
<thead>
<tr>
<th>School</th>
<th>Participant</th>
<th>Helpfulness</th>
<th>Identification</th>
<th>Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LA</td>
<td>3.00(0.82)</td>
<td>3.00(0.82)</td>
<td>2.50(0.58)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>3.00(0.82)</td>
<td>3.25(0.50)</td>
<td>3.25(0.50)</td>
</tr>
<tr>
<td>2</td>
<td>LA</td>
<td>3.50(0.57)</td>
<td>2.50(0.58)</td>
<td>3.00(0.82)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>3.50(0.57)</td>
<td>3.50(0.58)</td>
<td>3.25(0.96)</td>
</tr>
<tr>
<td>3</td>
<td>LA</td>
<td>3.00(0.00)</td>
<td>3.00(0.00)</td>
<td>3.00(0.00)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>3.25(0.50)</td>
<td>2.75(0.50)</td>
<td>3.50(0.58)</td>
</tr>
<tr>
<td>4</td>
<td>LA</td>
<td>3.25(0.50)</td>
<td>3.00(0.82)</td>
<td>3.50(0.58)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>3.75(0.50)</td>
<td>3.00(0.82)</td>
<td>3.75(0.50)</td>
</tr>
<tr>
<td>5</td>
<td>LA</td>
<td>3.75(0.50)</td>
<td>3.25(0.50)</td>
<td>3.25(0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>3.75(0.50)</td>
<td>3.25(0.96)</td>
<td>3.75(0.50)</td>
</tr>
<tr>
<td>6</td>
<td>LA</td>
<td>3.00(0.00)</td>
<td>3.00(0.00)</td>
<td>2.75(0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>3.00(0.82)</td>
<td>2.75(0.95)</td>
<td>3.25(0.96)</td>
</tr>
<tr>
<td>7</td>
<td>LA</td>
<td>3.50(0.57)</td>
<td>3.00(0.00)</td>
<td>3.25(0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.75(0.50)</td>
<td>3.25(0.50)</td>
<td>3.50(0.58)</td>
</tr>
<tr>
<td>8</td>
<td>LA</td>
<td>3.00(0.00)</td>
<td>3.50(0.58)</td>
<td>3.00(0.00)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.50(1.00)</td>
<td>3.00(0.82)</td>
<td>3.25(0.50)</td>
</tr>
<tr>
<td>9</td>
<td>LA</td>
<td>3.00(0.00)</td>
<td>3.00(0.82)</td>
<td>3.50(1.00)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.75(0.96)</td>
<td>3.00(0.00)</td>
<td>3.25(0.50)</td>
</tr>
</tbody>
</table>

| Overall | LA          | 3.22(0.59)  | 3.02(0.56)     | 3.08(0.60)   |
|         | CT          | 3.14(0.80)  | 3.08(0.65)     | 3.42(0.60)   |

The overall mean of problem identification ratings by learning assistants was 3.02 with a standard deviation of 0.56 and by the classroom teachers was 3.08 with a standard deviation of 0.65. Participants from one school reported the
same rating between their mean and standard deviation scores for this area. The range of scores for learning assistants (2.50-3.50) showed considerable variability and as did the narrower range of scores for classroom teachers (2.75-3.50).

Overall mean rating of participant shared understanding (agreement) by learning assistants was 3.08 with a standard deviation of 0.60 and by the classroom teacher was 3.42 with a standard deviation of 0.60. This showed the greatest difference between the participants' ratings of the three areas. Learning assistants' ratings ranged between 2.50-3.50 and classroom teachers' range was between 3.25-3.75 with great variability indicated by the standard deviations. The results indicated heterogeneity between individual participants' ratings and across schools.

The learning assistants were constant for each school and, consequently, their ratings across each area were more homogeneous than those of the classroom teachers. One learning assistant gave each interview the same rating across all areas. Three learning assistants gave the same ratings to interview helpfulness and problem identification but then varied on shared understanding. 

**Interview Helpfulness**

A summary of the consultation participants' ratings of the degree of their satisfaction with the helpfulness of the interview in defining students' problems is presented in Table 4.6.

The majority of ratings by the Learning Assistants were ratings of "mostly helpful" with a smaller percentage indicating their interviews were "completely
Table 4.6: Distribution of Learning Assistants' and Classroom Teachers' Ratings of Interview Helpfulness

<table>
<thead>
<tr>
<th>Participant</th>
<th>Percent of Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>LAT</td>
<td>0</td>
</tr>
<tr>
<td>CT</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: Ratings of helpfulness were on a 4-point scale where "1" = Not At All Helpful, "2" = Somewhat Helpful, "3" = Mostly Helpful, and "4" = Completely Helpful.

Eight percent of the Learning Assistants' responses assigned a level two rating, "somewhat helpful", to their interview with the classroom teacher. The classroom teachers' ratings were more evenly distributed. Twenty-five percent of the classroom teachers rated their interview with a level two rating, 36% rated it with a level three and 39% gave their interview a level four rating. Although a substantial proportion of classroom teachers found their interview to be completely helpful (39%), 25% of the teachers rated their interviews as only somewhat helpful. The learning assistants' responses were generally more favorable as to the helpfulness of the interview in defining students' problems.

Interview Problem Identification

A summary of the learning assistants' and classroom teachers' evaluative ratings of the degree to which they identified student problems during the interview is presented in Table 4.7.
Table 4.7: Distribution of Learning Assistants’ and Classroom Teachers’ Ratings of Problem Identification

<table>
<thead>
<tr>
<th>Participants</th>
<th>Percent of Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>LAT</td>
<td>0</td>
</tr>
<tr>
<td>CT</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: Ratings of Problem Identification were on a 4-point scale where "1" = Not At All Identified, "2" = Somewhat Identified, "3" = Mostly Identified, and "4" = Completely Identified.

The majority of the Learning Assistants’ responses (70%) endorsed the interview as having mostly identified the problem(s) discussed. The remaining 30% of endorsements were almost evenly divided between "somewhat" identified and "completely" identified. Over half of the classroom teacher endorsements (58%) assigned ratings of mostly identified. The remaining endorsements were divided between somewhat identified and completely identified. One quarter of the classroom teachers’ responses endorsed complete identification of the problem. The distribution of percentages suggests that the learning assistants and classroom teachers endorsed the same levels of agreement as to the degree of problem identification reached in their interview by their choice of agreement levels but not in the strength of their endorsements at each level. They differed in the strength of their endorsements.
Shared Understanding Between Participants

A summary of the learning assistance teachers' and classroom teachers' rating of the degree of their shared understanding of student problems between consultation participants are presented in Table 4.8.

Table 4.8: Distribution of Learning Assistants' and Classroom Teachers' Ratings of Shared Understanding

<table>
<thead>
<tr>
<th>Participants</th>
<th>Percent of Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAT</td>
<td>1 14 61 25</td>
</tr>
<tr>
<td>CT</td>
<td>0 6 47 47</td>
</tr>
</tbody>
</table>

Note: Ratings of Shared Understanding were on a 4-point scale where "1" = Not At All Understood, "2" = Somewhat Understood, "3" = Mostly Understood and "4" = Completely Understood.

More than half of the learning assistants' ratings of shared understanding from their interview were designated as "mostly understood." Twenty-five percent of endorsements assigned ratings of "complete understanding" and the remainder were designated as "somewhat understood." Ninety-four percent of the classroom teachers' ratings were evenly divided between "mostly and completely understood." Only a small percentage endorsed a "somewhat understood" rating. The distribution of percentages suggests the learning assistants' perspective had greater variability and reflected more disagreement than the classroom teachers.
Post Consultation Interview Agreement

The first level of analysis of these ratings was to investigate the level of interrater agreement to establish the interrater reliability of ratings between the raters. Interrater agreement was determined by comparing ratings assigned by two independent raters to each pair of post consultation interviews. (A copy of the rating scale is provided as Appendix #4.) All ratings fell between levels three and four, "mostly agree" and "complete agreement, respectively. Distribution of ratings are summarized in Table 4.9.

Table 4.9: Interrater Agreement

<table>
<thead>
<tr>
<th>Interrater Overall Agreement</th>
<th>Kappa</th>
<th>Percent of Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Rater 1</td>
<td>0.78</td>
<td>0</td>
</tr>
<tr>
<td>Rater 2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The raters demonstrated an interrater percentage agreement of .89, calculated by the percentage agreement formula (number of agreements/ number of disagreements + agreements). In order to correct for chance agreement, Cohen's Kappa statistic was calculated for the interrater agreement as well. The level of agreement based on the analyses of the post-consultation interviews tape transcriptions and accompanying tapes was $K = .78$, which is considered average and necessary results for overall agreement.
All of Rater 1’s and Rater 2’s responses were between levels three and four indicating that they rated all interviews as agreements. Rater 1’s percentage of ratings of the interview agreements were almost evenly divided between a "mostly agree" (53%) and "complete agreement" (47%). Rater 2’s percentage ratings placed a stronger endorsement on mostly agreeing (64%) than on complete agreement (36%). Neither rater assigned ratings of level one or two indicative of a disagreement between interview participants.

When compared with the consultation participants’ percent of ratings of shared understanding/agreement on the problem(s) identified in the interview the results indicated that the participants assigned more disagreement than the raters. The learning assistants and classroom teachers assigned to 14% and 6% of the interviews respectively, ratings of "somewhat agreed" to their shared understanding of the interview. This indicated some disagreement on how they perceived their dyad’s shared understanding of the problem(s) identified in the interview. The raters independently reported that the post consultation interview reflected two levels of agreement (either "mostly" or "completely") between the participants. The strongest overall level of support was a rating of "mostly understood" given by the learning assistants and raters to the shared understanding reached as a result of the interview. The classroom teachers evenly divided their ratings between mostly and completely shared understanding. There appears to be general agreement among participants and raters that the interview participants were mostly in agreement as to the
problems discussed but given the diversity of rating assignments a general rather than specific conclusion is drawn.

For the purpose of subsequent analyses involving the independent raters, the ratings provided by Rater 2 were used.

Interparticipant Agreement

The second level of analysis investigating agreement between the participants’ ratings of the effectiveness of interview helpfulness, problem identification and shared understanding applies the Kappa statistic for chance-corrected percentage agreement and Chi-Square analyses on their response agreements.

Interview Helpfulness

In order to explore similarities between the learning assistants’ and classroom teachers’ responses as to how helpful they found the interview to facilitate problem identification, a Kappa statistic of agreement between the learning assistance teachers’ and the classroom teachers’ responses was calculated at $K= .26$, a low level of agreement. This reflected the variation of responses between learning assistants and classroom teachers as to the level of interview helpfulness. When the participants’ responses were analyzed as a group, the chi-square calculation was significant, $(X^2 (2) = 5.674, P< .10)$ indicating that an association could be made between the participants and their responses. That is, the pattern of responses for learning assistants was significantly different from that of classroom teachers. The contingency coefficient of .27 indicates the presence of a modest relationship between the
raters (participants) and the status of their ratings. The learning assistants found the interview to be more helpful in identifying student problems than the classroom teachers.

**Problem Identification**

In order to investigate further the level of problem identification which occurred between the participants during the interview, Kappa was calculated for participant agreement at $K = -0.06$, indicating substantial disagreement between the learning assistants and classroom teachers in their rating the extent to which problems were identified in their interview. When the participants' responses were analyzed as a group, the chi-square calculation was not significant ($X^2 (2) = 1.039$, $p = 0.595$), no association could be made between the ratings and the status of the raters. This indicates that their disagreement ratings, as noted in the Kappa result, are not related to their participant status.

**Shared Understanding**

Further investigation of the shared understanding of interview participants on the problem identified in the interview resulted in a Kappa calculated at $K = 0.05$ indicating weak agreement between the learning assistance teachers and classroom. This suggests that the level of shared understanding was viewed differently by the participants. This supported the variability between ratings given by the participants to their shared understanding or agreement reached from their interview.

In order to investigate the similarity of ratings of shared understanding provided by the learning assistance teachers and classroom teachers as a group,
chi-square analysis was conducted on their ratings. This analysis yielded a significant chi-square ($X^2 (2) = 5.42, p < .10$) indicating differences in the pattern of response ratings given by learning assistance teachers from those of the classroom teachers. A contingency coefficient of .26 indicated a modest relationship between the ratings and the status of the raters. As a group, the learning assistants viewed their shared understanding or agreement reached during the interview differently from the classroom teachers.

**Participant - Rater Agreement**

Further investigation of the shared understanding or agreement between participants was done by comparisons made between ratings of participants' agreement on problem identification with the second rater's overall agreement ratings of participant shared understanding expressed in the post consultation interview. The results of the comparisons are summarized in Table 4.10. Rater 2's endorsements of the level of agreement between participants were distributed between level three ("mostly agree") and level four ("complete agreement"). Sixty-four percent of the ratings given by Rater 2 and 64% of the ratings given by the learning assistants were level 3 ratings of "mostly agree". The actual number of interviews on which the learning assistants and Rater 2 agreed upon for a level 3 rating was 15. However, the learning assistants endorsed 14% at a level two rating, "somewhat understood", a level response not endorsed by the rater. The rater endorsed complete agreement (level 4) with 36% of the ratings compared with 22% for the learning assistants. When rating levels for individual interviews were compared with learning assistant ratings, a percentage
Table 4.10: Participant - Rater Agreement

<table>
<thead>
<tr>
<th>Participant Agreement with Rater 2</th>
<th>Kappa</th>
<th>Percent of Ratings Response Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>LAT</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>CT</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>

Note: Ratings of Agreement were on a 4-point scale where "1" = No Agreement, "2" = Somewhat Agreed, "3" Mostly Agreed and "4" = Complete Agreement

agreement of .56 was reached and when corrected for chance agreement, the Kappa statistic reported for their agreement was K=.14, indicating little agreement between the learning assistants and Rater 2 on specific interviews.

Comparisons between the classroom teachers and Rater 2 percentages revealed greater overall agreement in the distribution of their endorsements between level three and level four. Classroom teachers were evenly divided in their ratings of levels three and four of participant shared understanding, whereas rater two gave stronger endorsement to level three, "mostly agree." Classroom teachers assigned 6% of the ratings to level two, "somewhat understood", a response level not used by rater 2. The comparison of the classroom teachers' ratings with rater 2's ratings using Kappa indicated little agreement (K=.17) between rating sources for individual interviews. They
appeared to rate their shared understanding/agreement of the interview differently which resulted in greater variability in ratings. Although comparisons between the learning assistants and second rater and the classroom teachers and second rater appear similar, the low Kappa in each case suggests that they are using different criteria to evaluate their shared understanding. These findings suggest that the ratings of post consultation interview agreement do not accurately predict participants' ratings of their shared understanding of problems identified in the consultation interview.

Further chi-square analysis of interview shared understanding was conducted on the interparticipant agreement with Rater 2. This analysis yielded a significant chi-square ($X^2 (4)=9.78, p<.10$) indicating that there was a significant difference between the participants and rater in the association of the raters with the status of their ratings. The participants (learning assistants and classroom teachers) were more likely to distribute their ratings of understanding across three levels of responses to include an indication of a "somewhat understood" interview situation. Rater 2's distribution was between levels three and four, indicative of stronger agreement between the participants. A contingency coefficient of .29 indicated a modest relationship between the ratings and the status of the raters. This finding supports the participants' and rater's use of different criteria from which to evaluate their shared understanding of the problems identified during the interview. It suggests that the interview participants were more likely to identify potential disagreements than the rater.
Participant Interview Evaluation Measure: Part II

Problem Identification by Number, Nature, and Priority

The second part of this measure was a checklist which allowed participants to identify up to three student problems they may have discussed during the interview. The first part of the checklist was aimed at broad problem identification as to the nature of the student's problem and further descriptors could be endorsed to assist in clarifying some general characteristics. Participants were asked to decide whether the problem discussed during the interview was academic or social/emotional/behavioral. Next, they were to endorse one or more of the descriptors provided for greater clarification. In addition, participants were asked to indicate the problems in order of highest priority, if more than one problem was selected.

Number of Problems

The results of the mean number of problems reported by participants within individual schools and followed by the overall mean are reported in Table 4.11. The learning assistants and classroom teachers reported approximately the same number of problems (M=2.00 and 2.08, respectively) given the opportunity to identify up to three problems per student. The participants of one school shared the same mean and standard deviation. Participants from the remaining eight schools varied in their responses as noted in the distribution of their responses by percentages.

Table 4.12 summarizes the number of problems by percentage of occurrence which were identified by the participants. Participants were provided
Table 4.11: Participant Interview Evaluation: Part II Mean Number of Problems Reported by Participants in Individual Schools

<table>
<thead>
<tr>
<th>School</th>
<th>Participant</th>
<th>Problem Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LAT</td>
<td>1.25 (0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>1.75 (0.50)</td>
</tr>
<tr>
<td>2</td>
<td>LAT</td>
<td>1.75 (0.96)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.00 (0.82)</td>
</tr>
<tr>
<td>3</td>
<td>LAT</td>
<td>1.75 (0.96)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>1.50 (0.58)</td>
</tr>
<tr>
<td>4</td>
<td>LAT</td>
<td>2.00 (0.82)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.00 (0.82)</td>
</tr>
<tr>
<td>5</td>
<td>LAT</td>
<td>2.25 (0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.00 (0.00)</td>
</tr>
<tr>
<td>6</td>
<td>LAT</td>
<td>2.75 (0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>1.75 (0.50)</td>
</tr>
<tr>
<td>7</td>
<td>LAT</td>
<td>1.75 (0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.25 (0.50)</td>
</tr>
<tr>
<td>8</td>
<td>LAT</td>
<td>2.75 (0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.50 (0.60)</td>
</tr>
<tr>
<td>9</td>
<td>LAT</td>
<td>1.75 (0.50)</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>2.25 (0.50)</td>
</tr>
</tbody>
</table>

Overall LAT 2.00 (0.75)  CT 2.08 (0.64)

Table 4.12: Problem Identification: Number of Problems Identified By Participants

<table>
<thead>
<tr>
<th></th>
<th>One Problem</th>
<th>Two Problems</th>
<th>Three Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N  Percent</td>
<td>N  Percent</td>
<td>N  Percent</td>
</tr>
<tr>
<td>LAT</td>
<td>10 28</td>
<td>16 44</td>
<td>10 28</td>
</tr>
<tr>
<td>CT</td>
<td>6 17</td>
<td>21 58</td>
<td>9 25</td>
</tr>
</tbody>
</table>

Kappa = .10
the opportunity of identifying up to three problem areas. Approximately half of
the students discussed in the interview by either participant were identified with
at least two problems. The classroom teachers’ distribution of the number of
problems was more heterogeneous than the learning assistants’ distribution.
Two or more problems were reported more often than a single problem
according to the distribution of percentages. The percentage of agreement
between the participants was .53 and Kappa statistic was calculated on the
corrected for chance agreement of the participants as to the number of problems
identified. Results of K=.10 reflects the diversity of the distribution indicative of
low agreement.

Table 4.13 summarizes the interrater agreement on the number of
identified problems per interview and compares the number identified by the
second evaluator (Rater 2) with the participants’ responses. The Kappa for
agreement between Rater 1 and Rater 2 on the learning assistants’ responses was
0.50 indicating moderate agreement. Interrater agreement on the classroom
teachers’ responses using Kappa was 0.31, indicating mild agreement. The low
Kappa scores suggest the measure’s insufficient sensitivity in detecting many
variations in problem descriptions and inadequate criteria for description.

Table 4.13 also summarizes Rater 2’s distribution of the numbers of
problems identified as compared with the participants’ distribution. Using a
corrected for chance agreement Kappa, moderately low agreement (K=.28) was
found in comparing the learning assistants’ distribution with Rater 2. Their
strongest endorsements for two and three problems with a close agreement in
Table 4.13: Interrater Agreement on the Number of Identified Problems Per Interview

<table>
<thead>
<tr>
<th></th>
<th>One Problem</th>
<th></th>
<th>Two Problems</th>
<th></th>
<th>Three Problems</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Percent</td>
<td>N</td>
<td>Percent</td>
<td>N</td>
<td>Percent</td>
</tr>
<tr>
<td>LAT:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R1</td>
<td>10</td>
<td>28</td>
<td>13</td>
<td>36</td>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td>R2</td>
<td>7</td>
<td>19</td>
<td>15</td>
<td>42</td>
<td>14</td>
<td>39</td>
</tr>
<tr>
<td>Kappa = .50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R1</td>
<td>6</td>
<td>17</td>
<td>16</td>
<td>44</td>
<td>14</td>
<td>39</td>
</tr>
<tr>
<td>R2</td>
<td>4</td>
<td>11</td>
<td>17</td>
<td>47</td>
<td>15</td>
<td>42</td>
</tr>
<tr>
<td>Kappa = .31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 2 with the Learning Assistants:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAT</td>
<td>10</td>
<td>28</td>
<td>16</td>
<td>44</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>R2</td>
<td>7</td>
<td>19</td>
<td>15</td>
<td>42</td>
<td>14</td>
<td>39</td>
</tr>
<tr>
<td>Kappa = .28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 2 with the Classroom Teachers:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>6</td>
<td>17</td>
<td>21</td>
<td>58</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>R2</td>
<td>3</td>
<td>8</td>
<td>18</td>
<td>50</td>
<td>15</td>
<td>42</td>
</tr>
<tr>
<td>Kappa = .10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Identifying two problems in approximately half of the interviews. Comparisons with the classroom teachers revealed weak agreement (K = .10) with Rater 2’s endorsements. The greatest diversity of responses occurred between the rater and classroom teachers as to agreement on the number of problem identified. Meaningful interpretation of these results is reduced by the low interrater agreements reported.

**Nature of the Problem**

Results of the identification of the nature of the highest and second highest problems in priority is reported in Table 4.14. (Endorsements for a third problem were considerably lower and not included in the analysis.) Percentages
Table 4.14: Problem Identification: Nature of the Problem and Ranking Identified by Participants

HIGHEST PRIORITY PROBLEM:

<table>
<thead>
<tr>
<th>Participant</th>
<th>Academic</th>
<th></th>
<th>Soc/Emot/Beh</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Percent</td>
<td></td>
<td>N Percent</td>
</tr>
<tr>
<td>LAT</td>
<td>15</td>
<td>42</td>
<td>21</td>
</tr>
<tr>
<td>CT</td>
<td>17</td>
<td>47</td>
<td>19</td>
</tr>
</tbody>
</table>

SECOND PRIORITY PROBLEM:

<table>
<thead>
<tr>
<th>Participant</th>
<th>Academic</th>
<th></th>
<th>Soc/Emot/Beh</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Percent</td>
<td></td>
<td>N Percent</td>
</tr>
<tr>
<td>LAT</td>
<td>26</td>
<td>62</td>
<td>10</td>
</tr>
<tr>
<td>CT</td>
<td>19</td>
<td>60</td>
<td>12</td>
</tr>
</tbody>
</table>

reported close to an even split between academic and social/emotional/behavioral problems between both participants favoring social/emotional/behavioral descriptions over academic descriptions for the highest priority problem identified. A Kappa statistic to confirm participant agreement of the highest priority problem was calculated at $K = .66$, indicative of moderately strong agreement. This result indicates that the participants agreed at least 66% of the time when asked to decide whether the highest priority problem was academic or social/emotional/behavioral. (The percentage of their actual agreement was determined to be .83). This finding suggests that participants are somewhat more likely than not to agree on a general description of the highest priority identified problem.
Percentages reported for the second priority problem indicated somewhat stronger endorsement for academic over social/emotional/behavioral problems for the learning assistants and the classroom teachers.

All learning assistants assigned a second problem to all students discussed in the interview and 31 of the 36 classroom teachers reported a second problem for the same student.

**Reliability of Ratings of the Nature of the Problem**

Interrater reliability was determined as described earlier by the two raters using the participants’ post consultation interviews as the source of information for the ratings. The ratings of each participants’ post consultation interview were considered separately from their dyad’s post consultation interview in order to reduce potential influence of the other’s student problem representation. The interrater agreements on the participants’ responses to the nature of the problems identified are summarized in Table 4.15.

Interrater agreement reliability of ratings was determined at a Kappa of .61 when comparing the nature of the problems identified by the learning assistants and .66 when comparing those identified by the classroom teachers. Although these scores reflect moderate agreement on the type of problem given highest priority by the participants, the agreements were not sufficiently strong enough from which to draw meaningful conclusions. These findings also suggest that caution be used in further analyses of the results.

Further analyses were conducted using the participants’ ratings of the highest priority problem in comparison with the second rater’s ratings of the
Table 4.15: Interrater Agreement On the Nature of the Highest Priority Problem

<table>
<thead>
<tr>
<th>Academic</th>
<th>Soc/Emot/Beh</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Percent</td>
<td>N Percent</td>
</tr>
</tbody>
</table>

Learning Assistants:

(N=35)  
R1  11 31  
R2  18 51  

Kappa = .61

Classroom Teachers:

(N=36)  
R1  13 36  
R2  17 47  

Kappa = .66

participants' post consultation interviews. The results are summarized in Table 4.16. Moderately strong agreement was found between Rater 2 and the learning assistants (K=.67) and Rater 2 with the classroom teachers (K=.78). Since the interrater agreement was only of moderate strength on the participants' ratings of the nature of the highest priority problem, the comparative data of the

Table 4.16: Comparison Agreement Between Participants and Rater 2

<table>
<thead>
<tr>
<th>Academic</th>
<th>Soc/Emot/Beh</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Percent</td>
<td>N Percent</td>
</tr>
</tbody>
</table>

Rater 2 & LAT  
LAT 15 42 2158  
R2 19 53 1747  
Kappa = .67

Rater 2 & CT  
CT 17 47 1953  
R2 17 47 1953  
Kappa = .78
participants with the rater, although strong in appearance, is interpreted with caution.

An interesting finding in these comparisons is the strong level of agreement between the classroom teachers and second rater. A possible explanation for this is that they used similar criteria for identifying a problem as academic verses social/emotional/behavioral problems. The low interrater reliability established for this measure requires caution in interpreting this finding. However, this suggests there may be potential for stronger agreement than the results indicated which could depend upon the way in which problem identification is operationalized by using more stringent criteria.

Other Problems Missed or Ignored

The final stage of the checklist evaluation form completed by the participants asked them to report if any other problems were known about the student which may have been missed or ignored by either participant during the interview. Responses were presented as percentages in Table 4.17.

Table 4.17: Other Problems Unreported During The Interview

<table>
<thead>
<tr>
<th>Problem Description</th>
<th>Percentage</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Participant</td>
<td>N</td>
<td>Yes</td>
</tr>
<tr>
<td>Other Problems</td>
<td>LAT</td>
<td>35</td>
<td>22.8</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>36</td>
<td>19.4</td>
</tr>
<tr>
<td>Problems Missed</td>
<td>LAT</td>
<td>25</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>27</td>
<td>03.7</td>
</tr>
<tr>
<td>Problems Ignored</td>
<td>LAT</td>
<td>26</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>27</td>
<td>0.0</td>
</tr>
</tbody>
</table>
A reduced number of responses were made to these final questions regarding the content of the interview by the respondents. Nearly twenty-five percent of the participants' responses indicated other problems did exist for the students in question. Subsequent responses to the remaining questions were inconsistent and difficult to detect a pattern for analysis.

**Summary**

Analyses were conducted on the data collected from the nine learning assistance teachers and thirty-six classroom teachers who participated in this study of problem identification in consultation. Demographic information regarding educational background and prior experience with consultation in a school setting was collected initially from the participants. Approximately 75% of the participants from each group reported some experience with consultation prior to their involvement in this study. Nearly all participants lacked formal training and extensive practice therefore this was considered a nonexpert sample of consultation practitioners.

Learning assistants were asked to conduct problem identification interviews with four classroom teachers from their schools regarding real attending students. Following their interviews, each participant was interviewed separately for a summary statement as to what problem were identified during their interview. Transcriptions of these post-consultation interviews were transcribed and rated independently for agreement between the participants as to the nature, number, and priority of the problem(s) discussed. Interrater
agreement on the participant dyad's overall agreement or shared understanding of the interview content was calculated with Cohen’s Kappa at 0.78.

Interparticipant agreement was investigated using a participant evaluation of the interview on variables of interview helpfulness, problem identification, and shared understanding of the interview content. Significant differences were determined regarding the association of ratings given by learning assistants and classroom teachers on interview helpfulness and shared understanding. The second rater’s evaluations were compared with the learning assistants’ and classroom teachers’ responses but resulting Kappa statistics yielded an indication of weak agreement, not strong enough to make any conclusive predictions regarding the participants’ agreement.

Interparticipant agreement was also investigated for the number and nature of the problems reported in the post consultation interview and second part of the evaluation form. Participant agreement on the number of problems reported was calculated with a Kappa which was too low to establish strong agreement. Participant agreement on the nature of the problem was also calculated with Kappa (K=.66) which indicated moderate agreement.

Interrater agreements were too low for comparisons with the participants in the areas of nature and number of reported problems. This reflected the measure’s insufficiency in detecting the nature and number of the problems discussed. In light of this caution is taken in interpreting the agreement between the participants since the raters were unable to demonstrate an acceptable level of agreement using Kappa.
The majority of the interview evaluations indicated that further problems were present with the student and not discussed but there was no consistent pattern from which to evaluate the responses adequately.
CHAPTER FIVE: SUMMARY AND CONCLUSIONS

The purpose of this study was to describe practitioner agreement on problems identified following an initial interview for problem identification in the process of consultation. The study was undertaken in response to criticism of problem identification in behavioral consultation. Studies of problem identification reported in behavioral consultation literature lacked generalizability of agreement across interviewers as to the nature of the problems discussed in the interview (Gresham & Davis, 1988).

The current study described the levels of agreement between consultant and consultee (learning assistants and classroom teachers) when engaged in identifying a client's (student) problem during an initial interview of consultation. Agreement or reliability of problem identification was operationalized in the study in several ways. The primary measure of reliability of problem identification was rating participant agreement on post consultation descriptions of problems discussed during the consultation interview. Two trained raters reviewed the consultation participants' descriptions of the problems they identified during their interview.

Interparticipant agreement was determined by comparing the learning assistants' and classroom teachers' responses to the interview evaluation measure which evaluated the interview and the identified problems discussed by nature and number. Further determination of the reliability of problem identification was done by comparing participants' ratings with the second rater
for overall agreement on problem identification, problem descriptions and the number of problems identified.

**Interview Evaluation Measure**

**Part I:**

Following their post consultation interviews, participants were asked to complete an evaluation of their interview in terms of helpfulness, level of problem identification and shared understanding within their dyad. There was considerable variability in the responses from both groups when examining the mean scores. The classroom teachers' responses were more heterogeneous than those of the learning assistants. It is noted that the endorsements of the learning assistants, who participated in four interviews to the classroom teachers' one, were more consistent across their individual ratings for helpfulness, less so for problem identification and the least consistent in rating the level of understanding of identified problems shared by the participants. This suggests some validity for the utility of the interview process as an initiator of the representation of a student's problem for which greater agreement between participants for more complete identification may still be sought.

The variability of responses for interview helpfulness is reflected in the distribution of the endorsement percentages. The classroom teachers generally gave lower helpfulness ratings to the interview than did the learning assistants. This suggests that learning assistants found the interview to be more helpful in identifying students' problems.
The second question in the evaluation measure specifically addressed problem identification and the percentages of endorsement ratings showed that participants disagreed upon the degree to which the problem had been identified. Learning assistants generally supported a lesser degree of problem identification than did classroom teachers. This suggests that the learning assistants were less likely to identify the presenting student's problem in the same way as the classroom teachers.

Endorsements of participants' shared understanding of the problem discussed during the interview reflected some disagreement between the participants. Classroom teachers generally gave more support to levels of agreement and learning assistants' responses suggested support for some disagreement or a less complete agreement on what was discussed during their consultation interview.

The findings of this study suggest that learning assistants found the interview helpful in initiating an emerging identification of the problem while acknowledging that their understanding of the problem was not in full agreement with the classroom teachers.

Interparticipant and Interrater Agreement

Reliability of problem identification was measured by interparticipant and Interrater agreements using the Kappa statistic for comparisons of chance corrected agreements and Chi-Square analyses for comparisons of the associations between the raters and their ratings on the three item, four point rating scale. This second level of analysis of interview helpfulness, problem
identification, and shared understanding revealed low levels of agreements between the participants on their ratings of the helpfulness of the interview (K = .26) and shared understanding of the problems discussed (K = .05). Significant differences between the learning assistants and classroom teachers suggested that their ratings of the interview could be predicted from their status as raters. The learning assistants and classroom teachers disagreed on the level to which the problem was identified in the interview. However, there was no significant difference between the participants on their problem identification rating levels. The distribution pattern of their responses appeared similar. Both groups of participants gave their strongest endorsement to problem identification at a 'mostly' but not 'completely' identified level. By comparison, more support was given to mostly identified problem identification by the learning assistants. Conversely, the classroom teachers gave slightly more support to the problem as 'completely' identified.

Interrater agreement was established on the raters' evaluation of the post consultation interview. A four point scale was used by the raters to evaluate the participant post consultation interview statement for a comprehensive description of the problem. Raters evaluated the overall agreement between the participants based on their statements about the highest priority problems discussed. The Kappa statistic yielded a .78 which represents an acceptable level of agreement, indicating that the agreements occurred more frequently than would be expected by chance.
Additional support for the reliability of agreement on a shared understanding of the problem was attempted in order to confirm interrater and interparticipant agreements. As a further measure of assessing the reliability of these agreements, the second rater’s evaluations of overall interview agreement were compared with those of the learning assistants and classroom teachers. Rater 2’s endorsements were divided between two levels of overall agreement signifying agreement between participants. The participants, however, indicated some disagreement by using a partial rating closer to disagreement than to agreement. The Kappa scores calculated indicated that the agreements of ratings were slightly more frequent than could be expected by chance.

These results suggest that the participants were using a criteria for determining shared understanding that was different from the measure’s specifications. Further investigation of ways to improve upon how to uncover participant agreement is warranted.

Interview Evaluation Measure

Part II: Number of Problems

Following their post consultation interview, the participants were asked to respond to a problem identification measure which provided a global description of the nature of the problem and the number of problems specified for the student. Problems were identified as either academic or social/emotional/behavioral. A list of general descriptors was provided for each category which the participants could endorse as further clarification of the problem(s). Participants were given the opportunity to identify and describe up
to three problems in order of priority for each student. Most participants assigned at least two problems to the students discussed during the interview.

**Nature of the Problems**

The highest priority problem was designated as the problem which participants specified that the most attention was given in the interview. Descriptions of highest priority problems included its identification as having the greatest discrepancy between the expected performance (behavior) and actual performance and, typically, a trigger of other specified problems. In order to establish reliability of agreement of the highest priority problems, interparticipant and interrater agreements were calculated. Interparticipant agreement on the highest priority problem was calculated at $K = .66$, indicating moderately strong agreement between the learning assistants and classroom teachers on the nature of the highest priority problem. Interrater agreement on the nature of the highest priority problem resulted in $K = .61$ for the learning assistants and $K = .66$ for the classroom teachers. These are moderately strong agreements which fall below acceptable agreement levels suggesting that caution be used in interpreting the findings.

The interparticipant agreement determined by Kappa is of interest as it reports that the learning assistants and classroom teachers agreed on the description of the highest priority problem as either academic or social/emotional/behavioral at least 66% of the time. This implies that there is a basic level of agreement on the problems identified which can occur in the problem identification interview. This suggests that the general criteria used to
identify the problems were insufficient to facilitate agreement beyond a basic level.

Participants were also asked if any problems were ignored or missed during their interview. Most agreed that not all problems were discussed but there was no detectable pattern in their responses which could be evaluated further in this way.

The research findings of Hay et al. (1979) reported that the only acceptable agreement across interviewers was on the number of problems identified for each client. This was not the case for these findings, however. The interview evaluation measure imposed a limitation of three possible responses and the structure for responding did not seem to capture an adequate description to discern how many problems could be accurately identified. Hay et al. used a version of Cautela and Upper’s checklist of 25 problems, each with a specific list of symptoms/descriptors sensitive to more subtle differences in the problem descriptions offered. A comprehensive instrument similar in nature would potentially have made the comparisons easier to discern.

Implications

The findings of this study revealed agreement was present regarding the problem(s) identified by participants during their interview and reported in the post-consultation interview. It is noted, however, that the level of inter-participant agreement was mild and moderately supported by the interrater agreement. This suggests that a replication of the study be done in order to
validate these findings as support for the reliability of problem identification by participant agreement.

The findings of this study suggest that the participants did not have strong agreement on the outcome of their problem identification interview. There was, however, an indication that the participants moderately agreed on the nature of the problem discussed which may partially account for the variability in the participants' responses to their perception of shared understanding and interview helpfulness. Support for this finding comes from Bergan and Tombari (1976) who reported that problem identification occurred in only 43% of the consultation cases. Their successful cases provided support for problem identification as a critical element needed for problem solution.

The participants from this study did not give full endorsement to complete problem identification or shared understanding. The variability of responses were more indicative of an emerging agreement which in time and with more feedback could potentially lead to agreement in identification. The participants were able to agree at least 66% of the time as to what was the nature of the highest priority problem. This suggests the unfinished quality of problem identification. The lack of agreement between the participants as to the level of problem identification reached in the interview suggests that problem identification is in itself a dynamic process of stages which can better be accomplished when moving to a more complete stage, characterized in part by agreement between participants.
The concept of agreement may also need to be further clarified. Few of the participants chose complete or absolute agreement as a characteristic of their problem identification. This suggests that if problem identification is possible without absolute agreement, a level of acceptability would have to be determined at which point sufficient enough identification is made for problem solution to occur.

**Limitations**

The findings suggested that several limitations of the study interfered somewhat with the effects of a complete description of the reliability of problem identification. The evaluation instrument used to aid in identification and clarification of the problems was problematic in that it provided limited support in articulating the problems. Use of an established measure similar to the Cautela and Upper instrument used in the Hay et al. study may offer more reliable problem representation which can assist the participants and raters to stronger agreement and a more comprehensive description of presenting problems. Inadequate and the descriptors were too vague to assist either the participants or raters in articulating the problem.

Another issue raised is that of training interviewers. Empirical support for Bergan and Tombari (1975) and Hay et al. (1979) typically included a training component to ensure that problem identification would occur. Interviewing skills, particularly in eliciting appropriate verbalizations for problem identification were the focus of several follow-up studies (Brown et al. 1982; Kratochwill et al. 1989). It is noted from the participants' demographic
information that only two of the learning assistants received formal training in consultation, although many more of them used it. Two learning assistants were also using a consultation model for the first time in this study. The different styles of procedures in problem identification may have confounded the process, particularly if they were unable to agree on an identified problem during their interview. The number of participants was small, a larger sample, including a greater variety of experiences would be needed to draw any further conclusions.

Future Directions

Further investigation of the problem identification component of consultation is still needed. The literature suggests that interviewer training and the standardization of the interview format will lead to problem identification (Brown et al. 1982). A further descriptive study of the interview process itself would prove helpful in attempting to establish causal relationships between the problem identification interview and its successful outcome.

Consideration may be given to Nezu and Nezu's (1993) suggested model for targeting problem behaviors. The model offers a multidimensional approach to problem solving which considers a wider variety of intervening variables. This model is still theoretical in nature and may be unwieldy to attempt in a school setting. It does, however, challenge the problem solvers to reconceptualize their practices and find new perspectives. Greater emphasis on the dynamic nature of problem identification in the use of a consultative problem solving process suggests future research consideration be given to this area.
REFERENCES


Levin (Eds.), *Handbook of mental health consultation* (pp. 3-28). Maryland: NIMH.


APPENDICES
APPENDIX #1

DIRECTIONS FOR CONSULTATION PARTICIPANTS

Prior to the Start of the Interview:

You are asked to conduct an interview for the next thirty minutes about a student whom you have identified as difficult to teach. Please conduct your interview in the same way in which you would begin to discuss a student with problems at your school in order to reach a solution. At the end of the interview the researcher will ask you to state what you have identified as the main difficulties this student presents in teaching.

Post Consultation Interview Question/Statement:

You have just met with a colleague in a consultation session about _____________. Please recall the main problem or problems identified during this consultation interview about _____________. Being as specific as you can, please summarize what problem or problems were identified during this interview. If more than one problem was identified be sure to tell which problem was the most important or given the highest priority.
APPENDIX #2

Identification Number: ____________________

PARTICIPANT DEMOGRAPHIC QUESTIONNAIRE

1. Age:
   20 - 30 ____  41 - 50 ____
   31 - 40 ____  Over 50 ____

2. Sex ________

3. Highest degree level of educational training received:
   ____________________

4. Total number of years in teaching or other educational roles: _________

5. Occupational Title: (If more than one title is applicable, please indicate by an approximate number of years in each area.)

<table>
<thead>
<tr>
<th>Title</th>
<th>Currently Held</th>
<th>Formerly Held</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>Administrator</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>Special Educator</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>Teaching Assistant</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>Learning Assistant</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>Counsellor</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>School Psychologist</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>Consulting Teacher</td>
<td>________</td>
<td>________</td>
</tr>
<tr>
<td>Specialist (specify area):</td>
<td>________</td>
<td>________</td>
</tr>
</tbody>
</table>

6. Indicate the level of students which you are now teaching/ or with whom you have the most daily contact.

   Primary   Intermediate   Secondary
Identification Number: __________________________

7. Have you had any prior experience with the consultation process?
   None  Some  Frequent

8. How would you rate your prior experience with consultation?
   Positive  Neutral  Negative

9. Have you had any formal training in consultation?
   Yes  No

10. If so, how did you receive this training?
    University Course  ______
    Pro-D In-Service  ______
    Workshop  ______
APPENDIX #3

Identification Number: ____________________________

INTERVIEW EVALUATION: PARTICIPANTS

The goal of a Problem Identification Interview in consultation is to formulate a comprehensive description of a major problem or problems which may then be targeted for the development of an intervention.

1. Was this interview helpful in identifying the student's problem?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not At All Helpful</td>
<td>Somewhat Helpful</td>
<td>Mostly Helpful</td>
<td>Completely Helpful</td>
<td></td>
</tr>
</tbody>
</table>

2. As a result of this interview to what extent was/were the student's problems adequately identified?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not At All Identified</td>
<td>Somewhat Identified</td>
<td>Mostly Identified</td>
<td>Completely Identified</td>
<td></td>
</tr>
</tbody>
</table>

3. As a result of this interview to what extent do you feel that you and the other participant have a shared understanding of the major problem(s)?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not At All Understood</td>
<td>Somewhat Understood</td>
<td>Mostly Understood</td>
<td>Completely Understood</td>
<td></td>
</tr>
</tbody>
</table>
4. Use the outline below to describe the student's major problem(s) identified in your interview.

(Please check only those items discussed in your interview. If more than one major problem was identified, use the separate sheets provided for each problem, up to three.)

HIGHEST PRIORITY PROBLEM

A. First decide, was the identified problem primarily: (Choose one)

ACADEMIC ___ OR SOCIAL/EMOTIONAL/BEHAVIORAL ___

(If the problem was primarily Academic, complete section B only.
If the problem was primarily SOCIAL/EMOTIONAL/BEHAVIORAL,
complete section C only.)

B. How would you describe the problem(s) identified as ACADEMIC:

Content Area Deficit ____
(ie. General weakness noted in an area of study)

Production Deficit ____
(ie. Incompletion of assignments, tasks, etc.)

Specific Skill Deficit ____
(ie. Noted lack of skills in a specific area)

C. How would you describe the problem(s) identified in the areas of SOCIAL/EMOTIONAL/BEHAVIORAL:

Social Skill Deficit ____
(ie. Interpersonal relationship difficulties)

Behavioral Excesses ____
(ie. Frequent responses to stimuli)

Behavioral Deficits ____
(ie. Infrequent responses to stimuli)

Personality Variables ____
(ie. Behaviors characteristic of negative self-evaluation)
SECOND PROBLEM IN PRIORITY

A. Then decide, was the second identified problem primarily: (Choose one)

ACADEMIC ___ OR SOCIAL/ EMOTIONAL/ BEHAVIORAL ___

(If the problem was primarily Academic, complete section B only.
If the problem was primarily SOCIAL/EMOTIONAL/BEHAVIORAL,
complete section C only.)

B. How would you describe the problem(s) identified as
ACADEMIC:

Content Area Deficit
(ie. General weakness noted in an area of study)

Production Deficit
(ie. Incompletion of assignments, tasks, etc.)

Specific Skill Deficit
(ie. Noted lack of skills in a specific area)

C. How would you describe the problem(s) identified in the areas of SOCIAL/EMOTIONAL/BEHAVIORAL:

Social Skill Deficit
(ie. Interpersonal relationship difficulties)

Behavioral Excesses
(ie. Frequent responses to stimuli)

Behavioral Deficits
(ie. Infrequent responses to stimuli)

Personality Variables
(ie. Behaviors characteristic of negative self-evaluation)
THIRD PROBLEM IN PRIORITY

A. Then decide, was the third identified problem primarily:
   (Choose one)

   ACADEMIC ___ OR SOCIAL/EMOTIONAL/BEHAVIORAL ___

   (If the problem was primarily Academic, complete section B only.
    If the problem was primarily SOCIAL/EMOTIONAL/BEHAVIORAL,
    complete section C only.)

B. How would you describe the problem(s) identified as ACADEMIC:

   Content Area Deficit ______
   (ie. General weakness noted in an area of study)

   Production Deficit ______
   (ie. Incompletion of assignments, tasks, etc.)

   Specific Skill Deficit ______
   (ie. Noted lack of skills in a specific area)

C. How would you describe the problem(s) identified in the areas of SOCIAL/EMOTIONAL/BEHAVIORAL:

   Social Skill Deficit ______
   (ie. Interpersonal relationship difficulties)

   Behavioral Excesses ______
   (ie. Frequent responses to stimuli)

   Behavioral Deficits ______
   (ie. Infrequent responses to stimuli)

   Personality Variables ______
   (ie. Behaviors characteristic of negative self-evaluation)
5. Are there other problems for this student which are as or more important than those identified but did not surface during the consultation interview?

Yes ____  No ____

Were these problems missed?

Yes ____  No ____

Were these problems ignored?

Yes ____  No ____
RATERS' EVALUATION OF PARTICIPANT AGREEMENT
OF IDENTIFIED PROBLEMS

PART 1

The goal of a Problem Identification Interview in consultation is to formulate a comprehensive description of a major problem or problems which may then be targeted for the development of an intervention.

Based on your reading of this post-consultation interview transcript, rate the extent of agreement between the consultant and consultee on the student problem(s) identified in the post-consultation interview.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Disagreement</td>
<td>Mostly Disagree</td>
<td>Mostly Agree</td>
<td>Complete Agreement</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX #5

EXAMPLE OF AN INTERVIEW WHICH RECEIVED AN INTERRATER AGREEMENT RATING OF COMPLETE AGREEMENT ("4").

POST CONSULTATION INTERVIEW TRANSCRIPTS
Learning Assistance Teacher

Okay, one of the problems that we talked about first was that ... um ...
(Student) has difficulty understanding any kind of abstract questions ... um ...
involving problem solving, use of logic. And despite, uh, giving her some help on a one-to-one, either the teacher or the teacher assistant (TA), uh, she didn’t seem to be able to understand any better. Um, so giving concrete examples, um, she seemed to kind of, uh, shut down and not be able to take in what was being taught. And then we came to the second problem which is an aspect of it, is that, um ... the real problem seems to be the emotion, her emotional response to not understanding and then receiving help. So, she seems to be nervous ... um ... not able to take in the, sort of adapted instruction and she, she tends to kind of shut down, and nod and say, "Yes, I’ve understood it." When, in fact, the next minute when you ask her a question it’s obvious she has not understood it.

Query: How does she say she’s shut down. I mean, what ...

LAT: Just kind of a glazed look ... um, nodding ... smiling ...

Query: Like this ....

LAT: (agrees)

Query: Okay, what did you agree was the highest priority problem?
LAT: We felt that dealing with the emotional response to receiving help needed to be dealt with first. So we thought that we needed to find some ways to help (Student) feel more comfortable ... um ... asking for help, or asking questions when she does not understand and then receiving the help.

Query: Alright?

LAT: Okay?

Query: Anything else?

LAT: And we also felt that there was an aspect ... um ... we might want to look at ...

Query: So, this would be the third thing ...

LAT: Well, it might be part of the emotional response. (Student) has an older brother who is a high achiever and ... um ... (Student) feels very inadequate compared to him. And the mother also feels very ... um ... guilty and ... and concerned about (Student) because she’s quite different for her brother and how she learns and her achievement. So that we thought that might be taking a look at the family interactions and what’s happening at home might help us to ... to ... to, uh, find out a way of dealing with the emotional response. But she seems to be really self-conscious and have ... have some self-esteem problems around dealing with the fact that she ... she doesn’t learn easily and her brother does.
The main problem with (Student) is ... um ... there are two main problems, two problems, but to get to the second one you need to tackle the first one and that is ... um ... her emotional state when confronted with difficulty. She is, um, quite nervous, and very insecure, and, um, the problem that, that I come up with is how to approach her. I feel that sometimes I do intimidate her and, as well, um, just, uh, reading into her ... her facial expressions I find it very difficult. She just comes across with this ... with a blank look that’s ... that seems very hard for ... for me to read. And sometimes she agrees that she does understand something when ... when, uh, when the next question is ... is ... to ... to quiz her on that. She ... she very well doesn’t understand. So to get through the nervousness and that’s what we need to tackle first. And then dealing with, once we’ve tackled that coming up with ways that, she’s very good with her rote memorization, but when it comes to problem solving ... or ... um ... abstract thinking, things like that, she finds it very difficult to put things into categories and ... and problem solve. So that’s ...

Query: So that’d be the second ...

CT: ... the second ..

Query: Okay, so you need to get the main one, is that ...

CT: That’s right.

Query: ... the emotional.

CT: That’s right, get through that and then we can deal with her learning and difficulties that ... she ... that she has there.
Query: Okay ... will you have her next year ... likely? Or ...

CT: No ... no. Probably not ... probably not.

Query: But this is still something that's good to be addressed.

CT: That's right, yup. I ... I taught grade 8 and 9 this year and next year I just have the grade 8's 'cause we're starting a new program so unfortunately I won't get to see some of the kids that I had this year. So ...

Query: Any other problems that come up in your discussion, with (LAT) or ...?

CT: No, I think that's it in a nutshell, that's about it, I guess?!

Query: Okay ...
APPENDIX #6

EXAMPLE OF AN INTERVIEW WHICH RECEIVED AN INTERRATING AGREEMENT RATING OF MOSTLY AGREED ("3")

POST CONSULTATION INTERVIEW TRANSCRIPTION

Classroom Teacher

For the student that we just discussed, I would say that ... um ... the main, I would say the main concern that I have for him is the sense of responsibility for his own learning and I think that that’s really at the root of most of the other concerns. That’s ... um ... once he gets responsible for his own learning I think that that’ll improve. He’ll be more organized, once he’s more organized he’ll have a better time at studying. Once he starts studying a little bit more, it’s going to help him in the classroom. And ... um ... He’s actually ... I enjoy him as a student. He can be challenging but yet he’s refreshing because I can see that he’s got it there. He just hasta, sort of, scratch the surface, and ... um ... and continue to build his confidence because he’s very sports-wise, he’s excellent and, and he enjoys sports and does well in it.

And socially with his peers he’s popular and the kids like him. I think a lot of the time the academics get in the way and because he doesn’t realize that he might not be on the same level as the rest of his friends that a lot of the time that gets him down and rather than work on that he just figures, "Well, I’m popular, maybe I’ll just be funnier or make a joke about it." And ... I think that he’s gotta get over that fear factor? And a lot of times he has to really feel that competence and trust with his teacher. Um ... because, you know, he’s said that and, and I think that once he feels comfortable and he realizes, okay, well, y’know she’s there to help me, she’s not going to be upset with me because I didn’t get the homework done, um, that he feels better about it and he can say, "Well, (Teacher), I had a problem with THIS, in the assignment and then he knows that ... um ... that I’ll say, "Okay, well, let’s extend the deadline. What
could be a conceivable deadline for you? Two days?" "Oh, I can do it in two days." And a lot of times, if he's got the idea that ... um ... you're not there to catch him up on a mistake that you're there to assist him, he'll meet you. Um ... he'll meet you halfway and he'll work well for you. And that's ... that's the big thing that I wanna see for him for the end of this year and next year, is that he doesn't lose that competence that he's gained in the year because I can just see just from September to now ... um ... how much improvement that he's made. Part of it could be because, you know. HE'S 13!!! You know, there's that sense of maturity that's being built in as well. But, uh, yeah, that's ... that's my goal for him, is just to make sure that ... um ... he realizes that it's his learning, that he has to be more responsible for it and not depend on, you know, mom to pick up the homework or someone to tell him what the assignment is last minute. Or to finish it up in the morning. And I think that that will get him on the road to more small successes.

POST CONSULTATION INTERVIEW TRANSCRIPTION

Learning Assistance Teacher

Okay, uh, there's a few problems that were identified — one, was that, um ... oh yeah, he (Student), um, is still having trouble accepting responsibility for his own learning and um ... and that, um, he doesn't respond well to being placed in a situation where he needs support and he is not getting support from his family in that area. They both feel that it's ... it's more derogatory and will not help him, they don't want to stigmatize him.

And, um, also another problem that he was having is ... is in general, is getting his ideas down on paper. He seems to be much more of a verbal learner and, uh, I think the main problem, again it's ... it's a bit of a stand-off, there is ... is the problem of him accepting, accepting ... um ... help and not feeling ... um ... and not letting it affect his self-esteem. And also getting information down on ... on paper were the two, uh, major problems that we're having, that we found.