The Global Assessment of Relationship Functioning as a Global Assessment of Individual, Couple, and Family Functioning

Ben Lowell Ashcraft

Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/etd

Part of the Social and Behavioral Sciences Commons

Recommended Citation
https://digitalcommons.usu.edu/etd/2662
THE GLOBAL ASSESSMENT OF RELATIONSHIP FUNCTIONING AS A GLOBAL ASSESSMENT OF INDIVIDUAL, COUPLE, AND FAMILY FUNCTIONING

by

Ben Lowell Ashcraft

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

MASTER OF SCIENCE

in

Family and Human Development
ABSTRACT

The Global Assessment of Relationship Functioning as a
Global Assessment of Individual, Couple,
and Family Functioning

by

Ben L. Ashcraft, Master of Science
Utah State University, 1997

Major Professor: Dr. Scot Allgood
Department: Family and Human Development

Relationship issues underlie most problems that people bring to therapy.

Global relationship assessment has great potential in helping therapists account
systemically for global issues in individual, couple, and family functioning. The
purpose of this thesis was to assess the concurrent validity of the Global Assessment
of Relationship Functioning (GARF) with standardized self-report measures on
individual, couple, and family levels of relationship functioning and to assess the
GARF's discriminant validity in distinguishing between clinical and nonclinical
groups. It was hypothesized that GARF scores would correlate with scores obtained
on the self-report assessments and that a statistically significant difference would be
found between clinical and nonclinical volunteer groups on both types of assessments.
Data were collected from 27 individuals, 30 couples, and 14 families, of which half were currently receiving therapy and the other half were nonclinical volunteer participants from Utah State University. Both groups were given the same self-report assessments and were each assigned GARF scores by the interviewing therapist. The Outcome Questionnaire 45.2 (OQ-45.2) was used to assess individual and family self-reported relationship functioning. The Dyadic Adjustment Scale (DAS) was used to assess couple self-reported relationship functioning.

The data were analyzed using correlational analysis and independent t tests to compare GARF scores to scores obtained on the self-report assessment measures. The GARF was found to be significantly correlated with the self-report assessment scores of volunteer individuals and couples, supporting the concurrent validity for two of the six groups. In addition, for couples and families, the GARF and the self-report assessments showed statistically significant differences between clinical and volunteer groups in the same direction, supporting its discriminant validity. These findings are limited due to the small group size and the inconsistency of the results across all three groups. Possible explanations for the results are discussed along with implications for using the GARF as a measure of global assessment in therapy. Based on these findings, there is limited evidence that the GARF is a valid measure in its current computer rater form.
ACKNOWLEDGMENTS

I would like to thank Dr. Scot Allgood for his continued support, encouragement, and advisement as I worked with him in completing this project. He has truly given more than was required by his position and I will always be grateful for his dedicated time and effort. I would also like to thank my committee members, Brent C. Miller and Thomas R. Lee, for their numerous contributions to this thesis. I appreciate their willingness to participate on my committee.

I give special thanks to all my family members, particularly my parents, my wife, and two sons, for their patience, support, and love as I worked to complete this thesis for my master's degree. I love and appreciate all of you. I could not have done it without you.

Ben Lowell Ashcraft
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>COPYRIGHT</td>
<td>ii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Theoretical Position</td>
<td>6</td>
</tr>
<tr>
<td>Purpose of Study</td>
<td>8</td>
</tr>
<tr>
<td>II. REVIEW OF THE LITERATURE</td>
<td>10</td>
</tr>
<tr>
<td>Types of Assessment</td>
<td>10</td>
</tr>
<tr>
<td>Global Assessment Devices</td>
<td>18</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>21</td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>22</td>
</tr>
<tr>
<td>Design</td>
<td>22</td>
</tr>
<tr>
<td>Respondents</td>
<td>22</td>
</tr>
<tr>
<td>Instrumentation</td>
<td>25</td>
</tr>
<tr>
<td>Procedures</td>
<td>33</td>
</tr>
<tr>
<td>IV. RESULTS</td>
<td>36</td>
</tr>
<tr>
<td>V. DISCUSSION</td>
<td>44</td>
</tr>
<tr>
<td>Correlations</td>
<td>44</td>
</tr>
<tr>
<td>t Tests</td>
<td>50</td>
</tr>
<tr>
<td>Limitations</td>
<td>57</td>
</tr>
</tbody>
</table>
Implications for Therapy ........................................... 58
Recommendations ....................................................... 60

REFERENCES .................................................................... 63

APPENDIX ..................................................................... 68
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Summary of Respondent Characteristics</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>Summary of Correlations Between GARF and Self-Report Assessments</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>Summary of Independent t Tests for Individuals, Couples, and Families on Self-Report Assessments and GARF</td>
<td>41</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Four research approaches</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

There are various reasons why people come to therapy, including finances, sexual dysfunctions, depression, eating disorders, and various forms of abuse, to name just a few. However, some theorists suggest that underlying many therapeutic concerns are unresolved relationship issues (Becvar & Becvar, 1996; Kerr, 1981; Nichols & Schwartz, 1991). Consistent with this assumption, relationship problems could be considered the underlying factor (primary problem) that results in other problem symptoms (secondary problems) created from the dysfunctions within the relationship.

A study done by Horowitz (1979) helps support this assumption. Researchers interviewed clients, which resulted in a list of 192 problem statements. Of the 192 statements, 146 were classified as interpersonally related. Horowitz (1979) stated, however, that a client symptom list alone is unable to reveal the underlying dimensions of a problem, or the reasons a particular symptom exists. Horowitz found that each symptom on the check list could be classified into three interpersonal related dimensions, which are: (1) the degree of psychological involvement between the subject and the other person (the person to whom the problem statement is directed), (2) the nature of their involvement (e.g., friendly, hostile), and (3) their intention to influence, change, or control the other person (Horowitz, 1979). Thus, for whatever reason a person may seek therapy, the therapist has to account for the relationships the person is involved in.
In order to properly treat a client’s interpersonal or relational dysfunctions, clinicians and researchers need to be able to appropriately identify what these dysfunctions are and how they interfere with symptom resolution. Because functioning within a relationship can be affected by many factors, it is important to have valid ways of assessing relationships. Assessment of how well a relationship is functioning is particularly important in three different areas, including: (1) the individual’s relationships with non familial or unrelated others, (2) couple relationships, and (3) family relationships. Through assessment of relationship functioning within these three areas, researchers and clinicians can better understand the mental functioning within the relational context and the relationship satisfaction that members feel.

There are many relationship assessment devices such as the Outcome Questionnaire 45.2 (OQ-45.2; Lambert et al., 1996), Dyadic Adjustment Scale (DAS; Spanier, 1976), and Family Assessment Device (FAD; Epstein, Baldwin, & Bishop, 1983) that separately examine individual, couple, and family functioning in terms of mental health or relationship satisfaction. However, assessment of functioning in all three areas with one global assessment measure is something that several researchers and clinicians have felt could be useful in terms of efficiency and effectiveness (Julien, Markman, & Lindahl, 1989; Krokoff, Gottman, & Hass, 1989; Weiss & Tolman, 1990).
One such assessment device is the Global Assessment of Relationship Functioning (GARF; American Psychiatric Association, Diagnostic and Statistical Manuel of Mental Disorders, 1994), which was developed to tap into the overall functioning of relationships in the individual, couple, and family areas. The GARF is found in the Diagnostic and Statistical Manual of Mental Disorders: Fourth Edition (DSM-IV, 1994) and is used by therapists and clinicians to give an overall assessment of a relationship’s functioning from an observer’s perspective. The observer can rate the clients on an interval scale from 1 to 100 regarding relationship functioning, with higher scores representing better functioning.

The DSM-IV is a diagnostic manual that sets forth criteria for diagnosing various mental health disorders in the clinical field (DSM-IV, 1994). The DSM-IV is a revised edition of the DSM-III and the DSM-III-R, which both had enormous influence on the mental health field and were used by a variety of clinicians and researchers from various orientations (Denton, 1989). These diagnostic and statistical manuals of mental disorders have been important for the fact that they have helped develop a common language within the mental health field for diagnosing clients with various disorders of functioning (Denton, 1989). Without this manual, there is no set standard in the field regarding what may be considered functional or dysfunctional. In addition, the diagnostic system used in the DSM manuals has become accepted by most third-party reimbursers, who use this classification system as criteria for reimbursement of services (Benson, Long, & Sporakowski, 1992).
The DSM-IV, as well as previous editions including the DSM-III and the DSM-III-R, uses a multiaxial classification system, which has been considered one of the most significant advances of the DSM (Denton, 1990). The diagnostic criteria are separated into five different axes including: (I) clinical disorders, (II) personality disorders and mental retardation, (III) general medical conditions, (IV) psychosocial and environmental problems, and (V) global assessment of functioning (DSM-IV, 1994). The GARF is used on Axis V, and can be used individually or in conjunction with a variety of other global assessment measures, including the Global Assessment of Functioning (GAF), Social and Occupational Functioning Assessment Scale (SOFAS), and the Defensive Functioning Scale (DFS; DSM-IV, 1994). The fact that global assessment is included in this manual as one of the five multiaxial classifications of functioning (DSM-IV, 1994) emphasizes the important role global assessment plays in understanding mental health disorders.

Of the various ways of assessing relationship functioning, global assessment is the easiest way to gain an overall understanding of the entire relationship process. Global assessment is designed to give informative results about overall functioning with a limited amount of time and effort (Weiss & Tolman, 1990). Furthermore, global assessment assists in gaining an overall understanding of the relationship process by combining various individual factors of functioning into a global scale that looks at all individual relevant parts. Because global assessment has this unique
capability, it can be useful for both researchers and clinicians who are seeking to understand global processes of functioning.

However, some researchers may feel that global assessment is not useful because it is difficult to treat global functioning. Although treatment of global process may not be possible, without global assessment, clinicians and researchers could lose sight of the whole picture. By focusing solely on individual pieces, we have only limited understanding of each piece’s effect on overall functioning. By using global assessment, a clearer understanding of individual component’s effects on overall process can be obtained.

Because relationships are assumed to be a combination of several different dimensions (Lambert et al., 1996), global assessment can be useful in understanding overall relationship functioning through it’s ability to combine all relevant parts. According to Lambert et al. (1996), individual and family relationship functioning consists of “how the person feels inside (subjective discomfort), how they are getting along with significant others (interpersonal relationships), and how they are doing in important life tasks such as work and school (social role performance)” (p. 1). This has been clarified by other researchers to include perceived understanding (Cahn, 1989), a sense of friendship, enjoyment of being together, sharing various activities (Lauer, Lauer, & Kerr, 1990), and positive usage of humor (Carroll, 1990; Krokoff, 1991; McBrien, 1993), all of which contribute to a satisfying and well-functioning relationship. The GARF, using all of the above concepts, defines a functional
relationship as one with agreed upon patterns or routines, flexibility, occasional conflicts that are adequately resolved, shared understanding, and an optimistic atmosphere where feelings can be freely felt and shared between members (DSM-IV, 1994). These various components combine systemically together, with each individually influencing the overall functioning of a relationship, and mutually forming a global process of relationship functioning.

Theoretical Position

Overall functioning of individuals, marital couples, and families relationships can be particularly important because of the systemic dimension that exists in all types of relationships. One well known and accepted theory of relational interaction is general systems theory. Because systems theory has become so broad, including many different theories and philosophies, this discussion will speak specifically to cybernetic systems theory, which focuses on relationship systems as ones that are governed through feedback (Guttman, 1991). For purposes of simplification, cybernetic systems will be referred to as systems theory throughout this discussion.

Systems theory implies that human interaction has a systemic influence on all people involved in a relationship (Becvar & Becvar, 1996; Kerr, 1981; Nichols & Schwartz, 1991). Systems theory focuses on describing relationships and various patterns of interaction (Becvar & Becvar, 1996). For example, systems theorists would conclude that individual functioning affects, and is affected by interactions with
others, that marital functioning is determined through the systemic interaction process between both partners, and that family functioning is systemically influenced by all of the family members. Systems theorists conclude that problems are reflections of a disturbance in the relationship system (Kerr, 1981). If any part of a relationship, as defined by the GARF and the OQ-45.2, is dysfunctional, it will likely influence and contribute to dysfunction in other parts of the relationship system. If the rules or roles of the relationship are not clearly agreed upon, this could carry a negative effect into the relationship atmosphere, making it closed and impersonal. Dysfunctions in the relationship and interaction processes can lead to mental health difficulties, which lower relationship satisfaction levels and may move clients to desire therapy.

The above are the assumptions under which many marriage and family therapists work (Becvar & Becvar, 1996). Therefore, when an individual, couple, or family comes into therapy, one assumption is that their presenting problems are related to their relationship functioning, whether it be individual, marital, or family. Becvar and Becvar (1996) noted that family therapy might be more appropriately termed relationship therapy under the assumptions of systems theory because the focus of therapy is on the relationship processes that are contributing to problematic functioning. They stated that therapy, in this context, could “work at the individual level, couple level, the extended family level, the neighborhood level, or the societal level, and indeed many family therapists do that” (Becvar & Becvar, 1996, p. 13). These systems theorists specifically noted that dysfunctions are developed through
our relationships with others (Becvar & Becvar, 1996), which implies that individuals, couples, or families could be experiencing lower mental health levels as a result of their past or current relationship functioning.

Relationship functioning as defined by the GARF (DSM-IV, 1994) and the OQ-45.2 (Lambert et al., 1996) have systemic dimensions. Both the GARF and the OQ-45.2 mention the relationship’s affective state and its interpersonal aspects as important parts of relationship functioning. Clients who present to therapy with sexual difficulties, work problems, or discipline concerns may be experiencing these as a result of unresolved interpersonal conflicts, unmet relationship needs, or a relationship atmosphere where members are unable to share feelings freely. If individuals, couples, or families are experiencing subjective discomfort, or difficulties getting along with others or accomplishing important life tasks (Lambert et al., 1996), this may create dysfunction in a number of different areas of life.

Purpose of Study

The purpose of this study is to determine if the Global Assessment of Relational Functioning Scale (DSM-IV, 1994) is a valid measure to assess relationship functioning of individual, couple, and family relationships on a global level. This was accomplished by comparing GARF scores to scores received on the OQ-45.2 and the DAS self-report measures, which are well accepted and tap into the same constructs. The hypotheses are first, that concurrent validity for the GARF will
be supported through a positive correlation on measures of relationship functioning in the individual, couple, and family relationship areas; and second, discriminant validity will be supported through lower mental health functioning scores on the GARF and the self-report measures for the clinical groups than a volunteer counterpart. The results should establish whether the GARF is able to validly assess differences in relationship functioning in individuals, couples, and families and whether the GARF can distinguish between problematic and good relationship functioning.
CHAPTER II

REVIEW OF THE LITERATURE

In order to understand why the establishment of concurrent validity on the GARF is important, it is important to understand factors relating to global assessment of relationship functioning. The following discussion will address global assessment versus other types of assessment techniques, outline various global assessments and their contributions to the field, and end with a discussion on the usage of the GARF.

Types of Assessment

Assessment of relationship satisfaction and mental health can be a key component in understanding individual, couple, or family functioning. Olson (1976) described a 2x2 model of assessment with frame of reference and types of data as the key dimensions. From the reporter’s frame of reference (see Figure 1), both insider and outsider assessments are possible. Insider assessment is a report on functioning that is from the perspective of someone who is inside the system such as a spouse or a family member. An outsider assessment would be a report from someone who is external to the designated system (e.g., marriage relationship, family relationship) such as a therapist or neighbor. Because each of these perspectives could vary from the other, it is recommended that both be used in research and treatment (Olson, 1976).
On the other dimension are the types of research data that can be collected, which are subjective and objective (Olson, 1976). Subjective data are based upon an insider’s or outsider’s reflections or personal experience. Objective data consist of behavioral reports that are based on concrete evidence that was either reported or observed (e.g., number of times a husband yelled at his wife). Refer to Figure 1 for further clarification.

By using this 2x2 model, four different types of assessment are available. However, within each of the four categories, several different specific types of assessment can be used, which include, but are not limited to, self-report methods,

<table>
<thead>
<tr>
<th>REPORTER’S FRAME OF REFERENCE</th>
<th>TYPE OF DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insider</td>
<td>Subjective</td>
</tr>
<tr>
<td></td>
<td>Self-report Methods</td>
</tr>
<tr>
<td>Outsider</td>
<td>Objective</td>
</tr>
<tr>
<td></td>
<td>Behavioral Self-Report Methods</td>
</tr>
<tr>
<td></td>
<td>Behavioral Methods</td>
</tr>
<tr>
<td></td>
<td>Observer-Report Methods</td>
</tr>
</tbody>
</table>

Figure 1. Four research approaches.

behavioral self-report methods, observer-report methods, and behavioral methods (Olson, 1976). The assessment method used most frequently to study marital and family interaction is self-report methods, such as questionnaires or interviews (Olson, 1976). Subjective self-reports are based on the insider of the system's feelings or perceptions, as compared to objective self-report based upon an insider's concrete and specific behavioral reports of the system. Observer-report is an outsider's report (e.g., therapist) based on his/her subjective or objective observation of the system.

Outside-observer reports can give information that is unavailable through self-report methods alone. Bakeman and Gottman (1986) stated that self-report measures are unable to specify clearly about interactions, or how these interactions affect the relationship's satisfaction. Ideally, self-report and observer-report could be used jointly in the assessment process. However, Olson (1976) stated that rarely are insider and outsider assessment methods used in conjunction with one another. By using valid observer-reports to assess relationship satisfaction, in addition to standard client self-report methods, additional information is made available which makes the assessment process more complete. Two types of outsider-observer report methods are microanalytic and global assessment. Researchers suggest several advantages to using these outsider-observational assessment techniques and how they provide different information that can be useful (Bakeman & Gottman, 1986; Julien et al., 1989; Krokoiff et al., 1989; Notarius & Markman, 1989; Weiss & Tolman, 1990). These two assessments are discussed below.
Microanalytic Assessment

In the past, microanalytic assessments have been more widely used than global assessment (Weiss & Tolman, 1990). Notarius and Markman (1989) stated that microanalytic assessments are typically used to provide detailed, specific information regarding a particular type of interaction stream that can be used to form specific hypotheses about an interaction process. Bakeman and Gottman (1986) have specified that microanalytic assessment is particularly useful to provide systematic observation of behaviors in order to understand what behaviors preclude or predict others. Some advantages to using a microanalytic assessment are (a) its ability to reveal complex patterns of interactions (Notarius & Markman, 1989) through specific observation of a particular behavior or interaction; (b) its ability to help gain a key understanding of these patterns and dimensions (Julien et al., 1989); and, (c) its ability to provide specificity often needed in doing sequential analysis (e.g., tracking patterns to find specific conditions or probabilities; Bakeman & Gottman, 1986; Weiss & Tolman, 1990).

However, a major disadvantage to using microanalytic assessment is the inability to assess overall functioning of various dimensions in a cost-effective and time-effective manner due to the lengthy specialized training that is needed (Weiss & Tolman, 1990). Bakeman and Gottman (1986) stated that a major concern with microanalytic assessment is training observers so that they produce similar results, given that they observe the same detailed behaviors. Because of the detailed
assessment process, training of observers to ensure accurate and reliable results can take enormous amounts of time and money.

Another disadvantage of microanalytic assessment is the time and detail required in developing an appropriate coding scheme (Bakeman & Gottman, 1986). Because of the time required, this step of microanalytic assessment is often taken casually, with researchers borrowing coding schemes developed for different projects that may not be theoretically appropriate (Bakeman & Gottman, 1986). Finally, another critical disadvantage to these assessments is that, due to their specific focus on a single dimension or behavior, they are often unable to pick up on overall functioning (Weiss & Tolman, 1990). Being unable to pick up on overall functioning, microanalytic assessment, similar to self-report measures, is limited in that interactional effects on the relationship process cannot be fully identified or understood.

Global Assessment

Because of the disadvantages of using microanalysis in assessing overall functioning, interest in the advantages of using global assessment has increased in recent years (Weiss & Tolman, 1990). Notarius and Markman (1989) stated that global assessments focus on a more general interactional dimension, assessing only the relevant parts that influence a particular dimension system, in order to get an idea regarding overall functioning of the particular dimension (e.g., withdrawal, problem solving). Global assessment devices use larger coding units than microanalysis and
have higher levels of inference than do the microanalytic assessments (Julien et al., 1989).

In contrast to microanalytic assessment, several advantages have been reported regarding the use of a global assessment device (DSM-IV, 1994; Julien et al., 1989; Krokoff et al., 1989; Notarius & Markman, 1989; Weiss & Tolman, 1990). One advantage is its usefulness in tracking the clinical progress of clients' overall functioning on a dimension over time with the use of only a single measure (DSM-IV, 1994). Another advantage is the potential cost-effectiveness and time-effectiveness in training and administration of the assessment when only a single measure is needed that is capable of assessing overall functioning (Weiss & Tolman, 1990). Krokoff et al. (1989) have suggested that a more rapid assessment on a global scale would be helpful to researchers who use large samples because it would give them the opportunity to take several variables into account when assessing a dimension such as marital interaction and marital satisfaction.

One other advantage is that global assessments can build upon the factors that are successfully identified through microanalysis in order to make higher levels of inference towards the overall functioning of a dimension (Weiss & Tolman, 1990). One group of researchers has suggested that global assessment could be used to identify the critical features of the patterns identified in microanalytic assessment (Notarius & Markman, 1989). In fact, several researchers have suggested that global assessment should not replace microanalysis, but be used in conjunction with it to
build and extend upon the past and current findings of microanalytic assessment in a way that is cost- and time-effective (Julien et al., 1989; Krokoff et al., 1989; Weiss & Tolman, 1990).

A study was done comparing a microanalytic assessment to a similar global assessment (Julien et al., 1989). This study compared the Interaction Dimensions Coding System (IDCS), which was created to predict relationship quality through assessment of positive and negative dimensions of interaction, with a congruent microanalytic assessment called the Couples Interaction Coding System (CICS). The results showed that the global assessment (IDCS) resulted in reliable and valid measurements of couples' interactions and that both assessments provided similar results, yet both on different levels of functioning (viz., global vs. specific behaviors; Julien et al., 1989). This supports the benefit of having both types of analysis available for assessment purposes. However, in cases where a researcher or clinician only desires to assess the overall, or global functioning of a dimension, the time and expense of doing a detailed and specific behavioral assessment would not be needed, or particularly useful (Krokoff et al., 1989, Weiss & Tolman, 1990). These types of situations are examples of when global assessment would be highly useful and beneficial to the research process.

Another example of a useful global assessment device is the RCISS (Krokoff et al., 1989). The Rapid Couples Interaction Scoring System (RCISS) is a global assessment used to “differentiate the problem-solving interactions of satisfied and
dissatisfied groups” by focusing on “husband and wife speaker and listener behavior” (Krokoff et al., 1989, p. 66). Results from a validation study show that valid global codes were obtained using the RCISS. These codes were able to discriminate between satisfied and dissatisfied married couples (Krokoff et al., 1989). Krokoff et al. (1989) have suggested that one of the uses of the RCISS is to analyze cross-classified events by comparing frequencies from microanalytic codes with episodes identified in the RCISS analysis, which avoids the cost of using microanalysis in areas of non interest.

Julien et al. (1989) stated that if their objectives for assessment were solely to understand the relationship status, they would recommend using a global assessment format due to cost considerations. Weiss and Tolman (1990) reported a study in which a microanalytic assessment (the Marital Interaction Coding System, MICS) was used to create a similar global measure (the Marital Interaction Coding System-Global, MICS-G). The purpose of this measure was to summarize a couple’s interactions through the use of categorical definitions. The global rating helps make an overall inference about the couple’s interaction patterns and to distinguish between distressed and nondistressed couple interactions. The end result is an efficient summary of the more detailed MICS microanalysis. The measure was tested for its cost-effectiveness in comparison to its microanalytic counterpart and for its ability to distinguish between distressed and nondistressed interactions. Results showed that the MICS-G was able to distinguish between distressed and nondistressed interactions
and was cost-effective in this process. The results also showed that the global assessment was more internally consistent on the overall functioning of relationship interactions than its microanalytic counterpart.

From the findings of various researchers regarding the current and potential benefits of global assessment, it is clear that global assessment can be useful in providing important assessment information that might be difficult to obtain in microanalytic ways. Global assessment can help researchers and clinicians better understand the overall patterns and functions of interactions. Global assessments can be used to assess many different global dimensions that impact relationship interaction and satisfaction.

**Global Assessment Devices**

Several different global assessment devices are discussed in the literature, including: the GARF, Marital Interaction Coding System-Global (MICS-G), the Interaction Dimension Coding System (IDCS), the Marital Interaction Rating System (MIRS), the Rapid Couple Interaction Scoring System (RCISS), the Communication Skills Test (CST), the Global Assessment of Functioning (GAF), the Social and Occupational Functioning Assessment Scale (SOFAS), and the Defensive Functioning Scale (DFS, Julien et al., 1989; Weiss & Tolman, 1990; DSM-IV, 1994). Several of these have been used in the past two years, including: the RCISS (cited four times), IDCS (cited two times), and the MICS-G (cited once; Social Science Citation Index,
The GAF, GARF, SOFAS, and the DFS have all been included in the most recent version of the Diagnostic and Statistic Manual of Mental Disorders (DSM-IV, 1994), which all can be used in diagnosing Axis V, which regards global functioning. Global assessment, as defined in the DSM-IV (1994) is the ability to judge the overall functioning level with a single measure. The DSM-IV (1994) Axis V evaluation is based primarily on the global assessment of the clients functioning.

The GARF (DSM-IV, 1994) is believed to give a global assessment of relationship functioning in regards to individuals, couples, and families from an observer’s (e.g., therapist) perspective (DSM-IV, 1994; Cleveland G. Shields, personal communication, November 16, 1995). This measure can be useful as a cost-effective assessment because of its ability to globally assess interaction on various relationship levels. The GARF was constructed for easy administration in various settings, such as emergency rooms, clinics, and a number of service agencies and can be used by both professionals and paraprofessionals (Kaslow, 1993). The GARF can be used as an observational interaction assessment for a specific interview, or an assessment of a specified time period (e.g., the length of the presenting problem; Cleveland G. Shields, personal communication, November 16, 1995). The GARF allows the observing clinician to rate the relationship in regards to meeting or not meeting the affective or instrumental needs in three areas of overall functioning: (1) problem solving, (2) organization, and (3) emotional climate (DSM-IV, 1994).
GARF is included in the *DSM-IV* (1994) to be used as an assessment of clients overall relationship functioning on the Axis V criteria for diagnosis.

Although the GARF had demonstrated reliability and face validity, concurrent validity in its ability to distinguish separately the relationship functioning in individual, couple, and family areas has not been established. In addition, discriminant validity of the GARF in regards to its ability to distinguish between clinical and nonclinical cases has not yet been established. The GAF, another global measure used in the *DSM-IV* (1994), has had prior testing supporting its validity with other standardized measures. Five studies done by Endicott, Spitzer, Fleiss, and Cohen (1976) showed the GAF to have both reliability and validity. They found that it showed greater sensitivity to change over time than did other comparative rating scales and that participants in the studies who had scores lower than 40 on the global assessment measure had higher probability levels of readmission to the hospital than patients with higher scores (Endicott et al., 1976). To this point, however, no such research has been done regarding the concurrent or discriminant validity of the GARF. The validity of the GARF as a global measure, capable of assessing overall relationship functioning, needs to be established in order to increase the overall validity of the measure.

Establishment of concurrent and discriminant validity on the GARF would support usage of global assessments in understanding overall functioning. The GARF could then be used either individually to assess relationship functioning from an observer’s perspective, or in conjunction with other insider or outsider-report
measures. Although microanalytic assessment has been widely used and accepted, establishment of valid global assessment devices such as the GARF would provide alternative ways of assessing functioning on an overall level. Without the use of global assessment, valuable information regarding relationship functioning might never be properly assessed or understood.

Hypotheses

The purpose of this study was to determine the relationship between the selected measures used for the assessment of relationship functioning. More specifically, the study will determine if the GARF can be considered a valid assessment of overall relationship functioning. The hypotheses are as follows:

1. Scores on the self-report assessments (OQ-45.2, DAS) for individual, couple, and family mental health will be positively correlated with scores on their overall relationship functioning as measured by the GARF.

2. There will be statistically significant differences between clinical and volunteer groups on the GARF and self-report measures (OQ-45.2, DAS).
CHAPTER III
METHODOLOGY

Design

In order to establish the concurrent validity of the GARF as a global assessment measure of relationship functioning, a correlational design was used. The correlational design was used to examine the specific relationship between variables by assessing their correlation with each other (Miller, 1986). The reason for using this type of design was an interest in the relationships (similarities and differences) that are observed between the groups and their responses on the assessment devices. Correlations were computed between the various measures given to the groups, particularly, the correlations between the GARF scores and the scores on the OQ-45.2 and the DAS. The purpose was to discover the types of relationships that were revealed through the correlation between the different measures and how those relationships supported or did not support the concurrent and discriminant validity of the GARF. Correlational designs are a standard method of establishing concurrent validity (Peck & Shapiro, 1990).

Respondents

Two groups of respondents were used in this study, including a clinical and a volunteer. Clinical respondents consisted of clients who presented for individual,
couple, or family therapy at the Utah State University Marriage and Family Therapy Clinic. This group also contained clients who presented themselves, and requested therapy for mental health difficulties, at the Family Connection Center of Davis Family Support.

The volunteer participants consisted of individuals, couples, and families who were attending classes at Utah State University at the time of the study. This group was recruited through various methods, including class presentations, newspaper ads, flyers, and word of mouth. Recruitment invitations were made in a total of six Utah State University classes, including three Family and Human Development 120 (Marriage and the American Family) classes and three Family and Human Development 150 (Human Growth and Development) classes. Presentations were made in four of these classes (two FHD 120 classes and two FHD 150 classes) during Winter Quarter 1997 and the final two (one FHD 120 class and one FHD 150 class) were done during the following Spring Quarter 1997. The average class size was 114 students. Both Family and Human Development 120 and 150 are classes used by majors from across the campus to meet general education requirements. Because married couples and entire families were not typically enrolled together in classes, in several cases, spouses or other family members were asked to participate.

During the class presentations a sign-up list was sent around for interested students to sign and put a phone number at which they could be reached and indicate whether they could participate as an individual, couple, or family. Students were
informed of the project and told that if they participated, they would be asked to attend a half-hour interview during which they would fill out self-report questionnaires and would be asked various questions regarding their interpersonal relationship functioning. Class members were also informed that they would receive a free movie ticket for their participation in the project. Students who wrote their names and phone numbers were contacted by phone and appointment times were set up when they could come to the USU Marriage and Family Therapy Clinic for an interview.

Newspaper ads were put in the Statesman (Utah State University's campus newspaper). Flyers were put up in various university buildings and family student housing areas, and were given to the student therapists for distribution. Both the newspaper ads and the flyers explained that individuals, couples, and families were needed to participate in a half-hour interview, and that participants would receive a free movie ticket in return. Student therapists were also asked to talk to friends and neighbors about the project and invite them to participate.

The two groups were of comparable sizes with the clinical group containing 36 cases and the volunteer group containing 35. Each of the two groups was divided into individuals, couples, and families. The clinical group consisted of 14 individuals, 14 couples, and 8 families. The volunteer group consisted of 13 individuals, 16 couples, and 6 families. Two volunteer couples and one clinical family was not used in the analysis due to missing data. Families included one or both parents and only
one adolescent, for purposes of analysis and consistency across the groups (due to the variation of family size). Participants ranged in age from 12 to 57 years old. Table 1 gives an overview of demographic information. As noted, respondents were relatively young, had some college education, had multiple children, and were predominantly Caucasian.

Instrumentation

Global Assessment of Relationship Functioning

Mental health and relationship functioning levels of both groups were assessed using the GARF (DSM-IV, 1994). The GARF is designed as a global measure for relationship functioning from an observer's perspective regarding the mental health of those being observed. The GARF is designed to permit assessment of a relationship's fulfilment of affective or instrumental needs in problem solving, organization, and emotional climate. This measure is designed on a scale from 1-100 in 20-unit increments with each differentially defining relationship functioning. Higher scores on the scale represent better relationship functioning. Levels within each unit are assessed by low, medium, or high numbers within the specific area (e.g., a 63 would be at the low end of functioning in the 61-80 unit of measurement).

Unit scores ranging from 81 to 100 represent a satisfactory functioning relationship. Scores ranging from 61 to 80 represent a somewhat unsatisfactory functioning relationship. Scores ranging from 41 to 60 represent occasional
### Table 1

**Summary of Respondent Characteristics**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Volunteer</th>
<th></th>
<th></th>
<th>Clinical</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
<td>SD</td>
<td>n</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Husband</td>
<td>28</td>
<td>27.32</td>
<td>9.66</td>
<td>20</td>
<td>29.05</td>
<td>8.06</td>
</tr>
<tr>
<td>Wife</td>
<td>29</td>
<td>25.61</td>
<td>9.74</td>
<td>33</td>
<td>30.18</td>
<td>8.01</td>
</tr>
<tr>
<td>Child</td>
<td>6</td>
<td>15.83</td>
<td>2.32</td>
<td>8</td>
<td>14.13</td>
<td>.99</td>
</tr>
<tr>
<td><strong>Years of education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Husband</td>
<td>28</td>
<td>15.57</td>
<td>1.99</td>
<td>20</td>
<td>13.80</td>
<td>1.47</td>
</tr>
<tr>
<td>Wife</td>
<td>29</td>
<td>14.57</td>
<td>1.71</td>
<td>33</td>
<td>14.04</td>
<td>2.07</td>
</tr>
<tr>
<td>Child</td>
<td>6</td>
<td>10.00</td>
<td>2.53</td>
<td>8</td>
<td>8.29</td>
<td>1.11</td>
</tr>
<tr>
<td>Number of children</td>
<td>6</td>
<td>1.31</td>
<td>2.61</td>
<td>8</td>
<td>2.14</td>
<td>1.50</td>
</tr>
<tr>
<td>Years married</td>
<td>6</td>
<td>3.66</td>
<td>6.79</td>
<td>8</td>
<td>5.34</td>
<td>5.86</td>
</tr>
</tbody>
</table>

**Volunteer (n)** | **Clinical (n)**

<table>
<thead>
<tr>
<th>Husband’s ethnicity</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian</td>
<td>27</td>
<td>17</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Asian American</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

*Table continues*
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Volunteer (n)</th>
<th>Clinical (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wife’s ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>Hispanic</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Asian American</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Native American</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Child’s ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Previous marriages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Husband</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>26</td>
<td>18</td>
</tr>
<tr>
<td>One</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Two</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Wife</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>26</td>
<td>20</td>
</tr>
<tr>
<td>One</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Child gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Female</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
relationship functioning with predominant dysfunction in the relationship. Scores ranging from 21 to 40 represent a seriously dysfunctional relationship with only minor areas of relationship functioning. Finally, scores from 1 to 20 represent a relational unit that is too dysfunctional for continued contact and attachment.

GARF scores can be calculated solely on the subjective/objective level of the observer by giving the individual, couple, or family a score based on observation and interaction. Another way of calculating GARF scores is to use a computer created program (Dennis, 1994) which asks 30 multiple choice questions regarding what was observed regarding the relationship functioning. The computer then calculates a GARF score based on the observer's responses to the various questions. The computer-based program was used for the purposes of this study in order to help establish consistency between trained observers.

**Outcome Questionnaire-45 2**

Concurrent validity, regarding the GARF's ability to distinguish between relationship functioning on the individual, couple, and family levels, was assessed by correlating the GARF's results to well accepted measures of individual, couple, and family relationship satisfaction and mental health functioning. The same measures were also used to assess the second hypothesis regarding the GARF's discriminant validity and its ability to distinguish between clinical and nonclinical cases. The measure used to assess the validity of the GARF in relation to the individual and
family functioning was the OQ-45.2 (Lambert et al., 1996). This measure has 45 total questions measured on a 5-point Likert scale from “never” to “almost always.”

The OQ-45.2 is designed to measure three specific aspects of mental functioning for individual and family functioning including: (1) symptom distress, (2) interpersonal relationships, and (3) social role performance (Lambert et al., 1996). The symptom distress subscale is designed to measure the amount of distress subjects are feeling based on symptoms they are experiencing that are all common to a variety of disorders. This subscale consists of 22 items with scores ranging from 0-88. The clinical cutoff for this subscale is 30, with scores 30 and above representing clinical disturbance in this area (Lambert et al., 1996).

The interpersonal relations subscale is designed to measure interpersonal complaints such as loneliness, interpersonal conflicts with family or others, and other related interpersonal problems. This subscale consists of 11 items with scores ranging from 0-44, with scores 15 and above representing the clinical cutoff. The final subscale, social role performance, relates to how a person is doing with important life tasks. This subscale has nine items, with scores ranging from 0-36 and the clinical cutoff being 12 and above.

Reliability on the OQ-45.2 was assessed by using 157 students from a large western university (Lambert et al., 1996). The test-retest values ranged from $r = .78$ to $r = .84$ and were found to be significant ($p < .01$) with the sample population. Internal consistency was assessed using 298 patients from an employee assistance...
program and the previous 157 student sample. The overall and subscale scores on internal consistency are considered to be relatively high for this measure. Internal consistency scores for the symptom distress subscale were alpha = .92 for the student group and alpha = .91 for the patient group. Scores for the interpersonal subscale were alpha = .74 for both student and patient groups. On the social role subscale, internal consistency scores ranged from alpha = .70 to alpha = .71 for student and patient groups, respectively. The total internal consistency score for the combined subscales was .93 for both groups.

The student sample was used to assess concurrent validity by using Pearson correlation coefficients on the OQ-45.2 total score and individual domain scores (Lambert et al., 1996). These scores were compared to scores obtained on similar measures (e.g., Symptom Checklist-90-R, Inventory of Interpersonal problems, Social Adjustment Scale, and five other measures related to symptom distress). When results on the OQ-45.2 and their respective counterparts were obtained, all correlations for the OQ-45.2 and the individual domains were statistically significant (p > .01) with the compared measures (Lambert et al., 1996).

Construct validity in relation to the OQ-45.2’s ability to show change throughout therapy was assessed (Lambert et al., 1996). A paired t test between pre-test and posttest scores of the OQ-45.2 given to 40 patients who had at least seven therapy sessions showed a statistically significant improvement in scores between the first and seventh session. Construct validity was further demonstrated by showing
statistically significant differences between clinical sample scores and community undergraduate sample scores. These results suggest that the OQ-45.2 can distinguish between nonclinical and clinical samples.

**Dyadic Adjustment Scale**

The measure used to assess functioning on the couple level was the DAS (Spanier, 1976). This is a 32-item self-report measure that can be given to married or nonmarried couples to assess the quality of the relationship dyad. Scores can range from 0-151, with higher scores representing higher relationship adjustment. The DAS has been used in numerous studies since its development in 1976 and is currently one of the most widely used measures to assess marital adjustment and satisfaction (Busby, Christensen, Crane, & Larson, 1995; L’Abate & Bagarozzi, 1993). This measure has a reliability coefficient of alpha = .96 for internal consistency with subscales ranging from alpha = .73 to alpha = .94. The concurrent validity was assessed by comparing the DAS with the Locke-Wallace Marital Adjustment Scale. Correlations between the two on divorced couples were $r = .88$, whereas correlations on married couples between the two measurements were $r = .86$. The DAS was found to discriminate between samples of divorced and married couples with overall scores on the measure averaging 70.7 for divorcees and 114.8 for married couples.
Procedures

The data for the clinical respondents were collected by giving clients at The Utah State University Marriage and Family Therapy Clinic each assessment measure at the time of intake as standard procedure for the entire clinic group. Clients who entered the clinic between November 1996 and May 1997 were given the measures to fill out. Data were collected at The Family Connection Center by giving clients seen by the Marriage and Family Therapy Intern the stated assessment measures during their first therapy session during the month of January 1997 until May 1997. Each measure from both therapy clinics was collected and scored. Each client was given a computer-rated GARF score by their first- or second-year therapist after their first 50-minute therapy session. To further ensure consistency and accuracy of the GARF score, some teammates (who were marriage and family therapy graduate students assisting on the therapy case) who watched the therapy cases from behind a two-way mirror also gave clients a GARF score. GARF scores were to be completed independently by the therapist and the teammates observing the case.

Interrater reliability scores between the therapist and teammate GARF scores were calculated using Cohen’s Kappa. This was chosen over other agreement percentage scores because of Kappa’s ability to account for the observed agreement which is due just to chance (Bakeman & Gottman, 1986). Cohen’s Kappa for this study was $k = .68$. Some researchers suggest that Kappas below .7 should be viewed...
with some concern; however, others have classified kappa levels .60 to .75 as good (Bakeman & Gottman, 1986).

The volunteer participants came to the USU Marriage and Family Therapy Clinic for a 30-minute face-to-face interview during which they completed the OQ-45.2 and the DAS (couples only). After completing the self-report measures, the participants remained for approximately 15 minutes during which time they were asked various questions regarding their relationship functioning in order to obtain information for assigning participants an appropriate GARF score based on their reported and observed relationship functioning. All of the volunteers were asked if they were currently attending therapy for a mental health problem in order to ensure their nonclinical status. None of the volunteer participants in this study reported currently receiving mental health services.

In order to obtain consistent data on the family assessments, only the parents and one child above the age of 12 were included in the assessment process. A cutoff age of 12 was chosen because symptom indices are commonly used down to this age (Derogatis, 1992). Recently (April 1997), the authors of the OQ-45.2 have developed a checklist to be used specifically for adolescents. This measure, however, was unavailable prior to data collection, and therefore was not used in this study. Scores for family assessment obtained on the OQ-45.2 were scored separately for the child and each parent, then averaged into one family score. The couple scores on the DAS were averaged together for husbands and wives in order to obtain a combined
satisfaction level for the couple (L’Abate & Bagarozzi, 1993). Scores for the OQ-45.2 symptom distress subscale and interpersonal response subscale were recoded in the opposite direction for purposes of the correlation analysis.

The graduate marriage and family therapy students who were either in their first or second year of training were trained in performing the volunteer interviews and in completing the computer-rated GARF scale. The graduate students also were trained in obtaining pertinent information relating to relationship functioning through receiving various handouts outlining the procedures as well as participating in a group meeting where the procedures were discussed. The procedures for completing the GARF on the computer were clarified as well. Any questions were then directed toward the student researcher.
CHAPTER IV
RESULTS

The scores from the GARF and the self-report measures were treated as interval-level data for purposes of analysis. The first hypothesis was that the scores on self-report assessments for individuals, couples and family’s mental health would be positively correlated with the overall relationship functioning scores obtained on the GARF. To address the first hypothesis, a correlational analysis was done comparing the overall results obtained on the self-report measures with the overall results on the GARF. The GARF was compared on the individual and family level with scores obtained on the OQ-45.2 to assess the extent to which concurrent validity for the GARF was supported through these correlations.

For individuals, the OQ-45.2 symptom response subscale scores were correlated with GARF scores. For couples, the overall scores from the DAS were averaged and then correlated with GARF scores. Couple’s averages were computed by summing the husband and wife’s total DAS scores and then dividing them by two. For families, the OQ-45.2 interpersonal response subscale scores were averaged and then correlated with GARF scores. The family averages were computed by taking the sum of the interpersonal subscale scores for each member of the family who filled out the measure, and dividing this by the number of family members who filled out the measure. Table 2 gives a summary of the correlations for both the clinical and volunteer groups between the GARF and the self-report standardized measures.
### Table 2

**Summary of Correlations Between GARF and Self-Report Assessments**

<table>
<thead>
<tr>
<th>Therapist GARF scores</th>
<th>Individual Symptom Distress Subscale OQ-45.2 scores</th>
<th>Couple DAS scores</th>
<th>Clinical group</th>
<th>Family Interpersonal Response Subscale OQ-45.2 scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r</td>
<td>.01</td>
<td>.47</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>14</td>
<td>14</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

**Volunteer group**

<table>
<thead>
<tr>
<th>r</th>
<th>.69**</th>
<th>.81***</th>
<th>.43</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>13</td>
<td>14</td>
<td>6</td>
</tr>
</tbody>
</table>

** p < .01  *** p < .001

The first hypothesis was only supported by correlations for volunteer individual’s and couple’s self-report scores and the GARF scores. A statistically significant correlation ($r = .69; p < .01$) was found in the expected direction between the volunteer individuals OQ-45.2 symptom distress subscale scores and their GARF scores. The shared variance between the two measures for this population was 48% ($r^2 = .48$). The $r$ square of .48 shows that nearly 50% of the variance is shared and can be accounted for through the relationship between the two measures. A statistically significant correlation ($r = .81; p < .01$) was found in the expected
direction between volunteer couples’ DAS scores and their GARF scores. The shared variance between the two measures for this population was 66% ($r^2 = .66$).

Correlations between the GARF and the averaged OQ-45.2 scores for the volunteer families were in the expected direction ($r = .43$) but did not reach statistical significance ($p > .05$).

None of the correlations for the clinical population reached statistical significance. The correlation for clinical individuals between their OQ-45.2 symptom distress subscale scores and their GARF scores was close to zero ($r = .01; p > .05$).

Correlations between clinical couples’ couple scores for the DAS and the GARF were in the expected direction ($r = .47$), but did not reach statistical significance ($p > .05$).

The correlation for clinical families between scores on the OQ-45.2 interpersonal response subscale and their GARF scores were also very low ($r = .01; p > .05$).

The second hypothesis was that a statistically significant difference would be found between clinical and volunteer groups on the GARF and self-report measures. To assess the second hypothesis, the data were compared using an independent-groups $t$ test to compare the two groups in regards to relationship functioning and mental health functioning based on therapist observation and participant self-report.

The independent-groups $t$ test was chosen because there were two groups and the dependent variable was interval level. Assumptions for using an independent $t$ test are (a) random sampling, (b) homogeneity of variance, and (c) normal sample distribution (Couch, 1987). Although this study meets the second assumption, the
first assumption was violated. Subjects were selected on the basis of convenience; no sampling techniques were employed. Because the distribution shape of the population is unknown, Couch (1987) recommended that sample sizes of more than 20 be used in order to better approximate a normal distribution.

Nevertheless, the results of the $t$ tests are summarized in Table 3. For the individuals group, mean scores between the clinical ($M = 35.50, SD = 13.78$) and volunteer ($M = 32.00, SD = 9.64$) participants on the OQ-45.2 symptom distress scale were not statistically significantly different ($t = .76, p > .05$). The mean scores for both the client and volunteer groups were within the clinical range on the measure (clinical cutoff > 30; Lambert et al., 1996). There was, however, a statistically significant difference ($t = -4.08, p < .001$) in individual group scores on the GARF between the clinical ($M = 46.35, SD = 16.53$) and volunteer ($M = 71.92, SD = 15.97$).

These results show that the therapists viewed the groups differently, assigning the clinical group much lower GARF scores than the volunteer group. Self-report scores of the individuals, however, showed that there was not a statistically significant difference between the groups, with the mean scores being virtually the same. Effect sizes were calculated by taking the difference between the two means and dividing this by the pooled standard deviation. Effect sizes were calculated to measure the practical significance of the findings in relation to their statistical significance (Cohen, 1994; Shaver, 1993). Effect sizes for the individual respondents were consistent
Table 3

Summary of Independent t Tests for Individuals, Couples, and Families on Self-Report Assessments and GARF

<table>
<thead>
<tr>
<th>Category</th>
<th>Clinical Group</th>
<th>Volunteer Group</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Individual respondents</td>
<td>(n = 14)</td>
<td></td>
<td>(n = 13)</td>
</tr>
<tr>
<td>OQ-45.2 (SD)</td>
<td>35.50</td>
<td>13.78</td>
<td>32.00</td>
</tr>
<tr>
<td>GARF</td>
<td>46.35</td>
<td>16.53</td>
<td>71.92</td>
</tr>
<tr>
<td>Couple respondents</td>
<td>(n = 14)</td>
<td></td>
<td>(n = 16)</td>
</tr>
<tr>
<td>DAS</td>
<td>102.39</td>
<td>13.71</td>
<td>118.54</td>
</tr>
<tr>
<td>GARF</td>
<td>58.00</td>
<td>10.66</td>
<td>79.06</td>
</tr>
<tr>
<td>Family respondents</td>
<td>(n = 8)</td>
<td></td>
<td>(n = 6)</td>
</tr>
<tr>
<td>OQ-45.2 (IR)</td>
<td>15.57</td>
<td>4.21</td>
<td>8.44</td>
</tr>
<tr>
<td>GARF</td>
<td>55.25</td>
<td>14.29</td>
<td>84.33</td>
</tr>
</tbody>
</table>

a Symptom distress subscale
b Interpersonal response subscale

** p < .01 *** p < .001
with the results of the t tests, with the individuals’ OQ-45.2 symptom distress subscale scores suggesting little practical significance (d = .30) and the therapists’ GARF scores suggesting practical significance (d = 1.24).

With the couples, the t-test results were in the expected direction with the participants self-reported scores on the DAS showing a statistically significant difference in the clinical (M = 102.39) and volunteer (M = 118.54) groups and the therapists reporting a similar difference in their assigned GARF scores (clinical M = 58.00; volunteer M = 79.06). The clinical cutoff level for the DAS is less than or equal to 107 (Crane, Allgood, Larson, & Griffin, 1990). The 107 cutoff level supports the couple scores in this study, with clinical couples’ reported mean levels of satisfaction being below the clinical cutoff and the volunteer couples’ reported mean satisfaction levels being above the clinical cutoff range. The effect sizes for the couple respondents were consistent with the results of the t tests, with both the therapists’ GARF (d = 1.33) and the couples’ DAS (d = 1.13) scores suggesting practical significance.

The scores on the family participants also showed statistical differences between the groups (see Table 3). The clinical families mean score on the OQ-45.2 interpersonal response subscale was 15.57, reaching the clinical cutoff point of 15 or above. The volunteer families had a much lower mean score of 8.44 on the interpersonal response subscale. The therapist’s GARF scores supported the results of the self-report measures with mean scores of 55.25 for clinical families and 84.33
for volunteer families. Differences between the clinical and volunteer groups on both the GARF and the interpersonal subscale of the OQ-45.2 were statistically significant (see Table 3). The effect size calculations corroborated the results of the t test for the family group. The effect size for the OQ-45.2 interpersonal response subscale (d = 1.40) suggested practical significance between the clinical and volunteer groups. The effect size for the therapist’s GARF score (d = 1.57) also suggested practical significance between the two groups.

These results, however, are very limited in their generalization because the group size was very small, with only eight clinical families (seven for the OQ-45.2 due to missing data) and six volunteer families. The scores obtained from both the couples and families used in this study did, with limitations, support that there was a statistically significant difference between clinical and volunteer groups on both the GARF and the self-report measures, both in the expected direction. The results of the effect size calculations showed practical significance between the clinical and volunteer couples and families on both the GARF and the self-report measure.
CHAPTER V

DISCUSSION

Correlations

The first hypothesis, that the results would show a positive correlation between the self-report assessment scores and the therapist-assigned GARF scores, was only partially supported. The correlations were low to zero between the two assessment styles (observer vs. self-report) for the clinical population. The only data for this group that showed a moderate correlation were the data obtained from the clinical couples. The correlation between the DAS scores and the GARF scores for this group was \( r = 0.47 \), which was in the expected direction, but was statistically nonsignificant. The volunteer group, however, yielded stronger support for the first hypothesis than the clinical group. Correlations between the therapist GARF scores and both the individuals' symptom distress scores and the couples' DAS scores were statistically significant.

This finding creates an interesting contrast. One interpretation is that more accurate observations are being made by therapists on volunteers than those who are seeking treatment. If this is the case, the GARF may not be an adequate global assessment to use in assessing relationship dysfunction, but only in assessing if a relationship is functioning well.
In addition, it may be that the GARF is not tapping into the most important aspects of relationship functioning for people in distress. The GARF is a fairly new global measure, and was not included in the DSM until its most recent version in 1994 (Kaslow, 1993). With the GARF being a fairly new assessment in the field, with limited use and development to this point in time, it could be that revisions need to be made in order for it to better assess the central aspects of relationship functioning.

There are several assessment devices that have been developed to assess aspects of relationship functioning such as the Circumplex Model (Olson, 1993), Beaver's Model of family functioning (Beaver & Hampson, 1993), and The McMaster Model (Epstein, Bishop, Ryan, Miller, & Keitner, 1993). Each of these assessments focuses on key elements that their designers felt were the most significant predictors of relationship functioning. For instance, the Circumplex Model focuses on family cohesion, flexibility, and communication (Olson, 1993). Beaver’s Model focuses on a stylistic dimension, adaptability, and affect (Beaver & Hampson, 1993). The McMaster Model focuses on affective involvement, behavior control, problem solving, roles, communication, and affective responsiveness (Epstein et al., 1993). Each model emphasizes different components, yet each is assessing relationship functioning, similar to the GARF.

The GARF focuses on problem solving, organization, and emotional climate (DSM-IV, 1994). These elements are found in some of the other relationship assessment models mentioned above. The GARF is similar to the Circumplex Model
with its focus on problem solving, and has a similar focus as all three regarding their focus on the affective (emotional) component of the relationship. The GARF’s organizational component is somewhat related to components such as roles and behavior control contained in the McMaster Model. For the most part, each of the mentioned assessments has more similarities than differences.

The GARF, however, differs from the other models in some key aspects. The GARF puts limited emphasis on areas such as flexibility and cohesion, although these are implied. With all four models discussed each having slightly different focuses regarding relationship functioning, the question then becomes, “Which of all the components presented by the various models are most important in assessing relationship functioning?”

A project funded by the U.S. Department of Health and Human Services and the Office of the Assistant Secretary for Planning and Evaluation looked at various traits considered to be important in identifying successful families (Krysan, Moore, & Zill, 1990). Of the nine traits identified, several were not mentioned by any of the four models presented above, including (a) encouragement of individuals, (b) expression of appreciation, (c) spiritual orientation, and (d) social connectedness. It may be that the small correlations between the GARF and the self-report measures could be accounted for because the GARF is focusing on issues that are not significant indicators of relationship functioning, or that it is missing information that plays a key role to overall relationship functioning.
These assumptions will require more research in order to know whether or not the GARF needs to be revised, or what types of changes should be made. Continued research regarding its ability to assess key components of relationship functioning, and then a possible revision according to the research findings, may be beneficial in showing stronger support for the validity of the GARF in the future.

On the other hand, the measure itself may not be at fault, but the therapist’s use of the measure. Given that the instructions for using the GARF are very limited, the therapists may have not appropriately used the GARF as it was designed to be used. The therapists in this study were student therapists. The procedures in accurately completing an assessment device may vary given the specific purpose of the device. In the DSM-IV, these aspects of the GARF are only minimally clarified. Maybe more extensive clinical training or specific training in the use of the GARF is necessary for effective use of the measure.

A GARF score, however, can be calculated using one of two methods. It is possible that the GARF scores assessed through the computer program are less accurate than the regular method of assigning a GARF score based on the criteria found in the DSM-IV. On several questions with the computer-rated GARF, the difference between assigning the family a two or a three on a question is minimal. For example, question number eight about communication differentiates between a two and a three by giving a number two score if members are generally aware of each other’s location and a number three score if families are often aware of other
members' locations. The difference between these two seems very minimal and may be confusing to an untrained rater.

In addition, some questions are confusing in the sense that two or more answers may be correct for a given individual, couple, or family. For instance, questions two and three relate to family rules. The questions answered on the computer by the therapist can range from 1 to 5 on a Likert scale. A number 2 is assigned to a system where rules are produced by the parental subsystem, whereas a number 3 is assigned to a system in which there is some confusion about acceptable behavior. In any given individual, couple, or family, both of these could be true at the same time. For example, a rule that sexual topics are not to be discussed in the home may be a rule established by the parental subsystem. However, children within the home could still be confused regarding what is or is not acceptable sexual behavior because they are never able to clarify it due to the parents' established rule about sexual discussions.

Therefore, since both cases are true with this family, it may be difficult for the observer to decide which score is most appropriate to assign the family. With both this, and the previously stated confusing factors regarding questions on the computer-rated GARF test, it may be difficult for an observer to assign an accurate GARF score to the members being observed. These possible scoring problems, however, would only account for a minimal number of points on the GARF at best. These potential
problems, however, emphasize the possibility that the computer version may need to be reviewed.

In the past, students have had a tendency to overrate clients using the subjective/objective method of calculating the GARF, and then, after reviewing the cases in supervision, dropped their scores (Scot Allgood, personal communication, July 17, 1997). Therefore, in this study, the GARF scores were calculated using the computer method in order to obtain more consistent and accurate results. At this point in time, information pertaining to the relationship between the computer-rated GARF and the subjective/objective method used in the DSM-IV is still very limited. Therefore, it is uncertain if this factor significantly influenced the results of this study.

The table used to calculate the Cohen's Kappa level showed that when discrepancies existed between the therapist's and observer's GARF scores, lower scores were typically assigned by the therapists than the observers. This difference makes it difficult to determine which perspective produced a more accurate GARF score. Maybe the therapists have a more accurate perspective, due to their direct communication with the interviewees, therefore making the observers less sensitive to the relationship functioning because they were not personally involved in the conversation. The opposite of this could also be true if the observers, because of their outside position, were able to pick up on physical or verbal clues that the therapist did not because they were too involved in the conversation or in thinking what question to ask next, and so forth.
Given the discrepancy in the assigned GARF scores by the therapists and the observers, it is unclear which scores were more accurate. This problem, however, might be resolved through using a team approach to the GARF. Kuhlman, Sincaban, and Bernstein (1990) did a study in which the Global Assessment Scale (GAS) was used through a team approach. A multidisciplinary team was set up for each client, with each member independently assessing a client’s GAS score. Scores from the group were then averaged for each client to develop a mean GAS score. This method was found to be reliable, valid, and cost-effective. A mean GARF score could be calculated in a similar manner, allowing both the therapist and teammate observations to be accounted for in a single GARF score.

\[t\] Tests

The second hypothesis, that statistically significant differences would be found between clinical and volunteer groups on both the GARF and self-report measures in a similar direction, was supported in five of the six comparisons. The \(t\) tests showed statistically significant differences between the clinical and volunteer groups on the GARF and both the couple DAS scores and the families OQ-45.2 interpersonal response subscale scores. The \(t\)-test results for the individual groups did not support the stated hypothesis on the symptom distress subscale of the OQ-45.2, although, a statistically significant difference was found in GARF scores between the clinical and volunteer groups.
The support from the couple and family data regarding the second hypothesis is important for several reasons. Since the findings were in the expected direction, with significant differences being found between the clinical and volunteer groups on both the GARF and the self-report measures, the assumption could be made that the GARF is a valid measure of relationship functioning when relational interaction can be observed directly by the therapist. These findings support the purpose of this study, which was to assess whether the GARF has discriminant validity in relation to other standardized assessment measures. The discriminant validity, however, was only supported in two of the three groups involved in the study, with the results showing no support for discriminant validity with the individual groups.

Interestingly, in both of the other groups, the therapist was able to listen to interaction between the couple or the family members, rather than just listen to them talk about their interactions. Individuals in the study were only able to discuss their interactions, which limited the information the therapist was able to obtain regarding relationship functioning. With couples and families, where interactions could be directly observed, similar results were found between the observation assessment method (GARF) and the participant’s self-report assessment. The fact that the t-test results were statistically significant in the expected direction for both couples and families provides support for the GARF being a valid assessment of relationship functioning when interviewing two or more family members at a time. Caution is
warranted, however, because of the various limitations discussed later regarding group size and standard deviations.

Given that the second hypothesis was only supported statistically in the couple and family data, but not with the individuals involved in the study, it may be that the GARF is not an effective measure when only one individual family member can be observed in the interview. This study showed no evidence that the GARF should be used in assessing relationship functioning for individuals. The statistically significant findings on the t tests for the couple and family data provide support that the GARF is a useful measure for assessing relationship functioning when more than one person is present for the interview. These findings also support the discriminant validity of the GARF to distinguish between clinical and nonclinical cases involving couples and families.

As previously stated, the individual groups did not support the second hypothesis. Although there are several possible explanations for lack of support on the t test, the most likely explanation, perhaps, is that the therapist’s prior knowledge of which individuals, couples, or families were clinical and which were volunteer could have biased their observations, and thus, their assigned GARF score. When experimenters in a study influence the results based on their prior knowledge regarding the purpose of the study, this is often called the Rosenthal effect (Martin, 1977). In this study, the therapists were aware at the time of the interview whether the person or persons were clinical or volunteer.
The interpersonal response scores on the OQ-45.2 for both clinical and volunteer individuals involved in the study were within the clinical range, indicating that both groups may have had mental health concerns. Regardless, the therapists assigned the volunteer individual group GARF scores, which were statistically significantly higher than their clinical counterpart. This could have resulted directly from the fact that the therapists knew beforehand which families were seeking therapy and which families were not. Therefore, the therapists in this study might have made the assumption that the volunteer families were functioning better based on the fact that they were not openly seeking therapy.

Barber (1976) discussed experimenter unintentional expectancy effects and the results this type of problem can have on results. He concluded that experimenter expectancy can modify the data and outcome of a study. Although the experimenter (or researcher) in this study did not do all the interviews, the other interviewers were aware of the purpose of the study, which may have biased their observations. Leedy (1974) described bias as any influence or condition that causes a distortion or aberration of the data. In addition, Leedy stated that bias can easily influence the results of research projects. The observers' bias, due to their knowledge regarding the study and the status of the interviewee, could have altered the results of the GARF. A possibility for future studies would be for the therapists to be blind to the status (clinical or nonclinical) of the interviewee. Observational bias could not be avoided in this study because of factors of confidentiality and ethics.
In support of the assumption that the therapists may have been biased based on prior knowledge regarding the status (clinical or volunteer) of the interviewee(s), Charny (1980) discussed ways in which therapists view people. He stated that the prevailing model in the past has been to view things in a right or wrong perspective, assuming that there was a right and a wrong way to do things. With this perspective, Charny stated that emotionally disturbed people were seen as people who were doing more wrong things than right, and that the therapist’s job was to teach clients how to do more things right. It may be that the therapists in this study may have viewed the clinical population as experiencing problematic relationship functioning and the volunteer population as experiencing satisfactory interpersonal relationships, based mostly upon their status as clinical or volunteer.

Another possible explanation for these findings may be that the self-report measures did not assess the global issues that the GARF was able to account for. For instance, individuals may have been experiencing symptom distress within the clinical range of functioning, but they may possess other global factors that counterbalance these factors and allow these individuals to function adequately in their relationships. Much of the research that has been done on other global assessments has shown them to be as good or better detectors of general functioning than their microanalytic countermeasures (Julien et al., 1989; Krokoff et al., 1989; Weiss & Tolman, 1990). This being the case, the GARF may be measuring factors that the self-report questionnaires are not accounting for. The possibility of this factor is unlikely
because the self-report measures used in this study were all well established and accepted global assessment measures.

A final explanation for the lack of support with the individual population may be that the groups had demographic differences which influenced the findings. The ages of the individuals ($M = 24$) and couples ($M = 26$) were similar, but the families were much older ($M = 40$). Couples, however, were in established relationships with child responsibilities, where the individuals were typically single. It may be that the individuals were not as accurate in their memory and reporting of their relationship functioning due to their youthful age and their single lifestyle. The couples and families, who were in daily contact with their significant others, may have had a clearer perception of relationship functioning during the interview.

The results of this study indicate that the concurrent and discriminant validity of the GARF is still not clear. Concurrent validity was supported through statistically significant score correlations between the therapist GARF scores and the OQ-45.2 and the DAS scores for volunteer individuals and families. All other correlations, however, for both the clinical and volunteer groups were not statistically significant, thus questioning the concurrent validity of the GARF. Since the correlations in the study produced inconsistent results between the GARF and the self-report assessments, further establishment of concurrent validity is necessary before making conclusions.
This study showed no support that the GARF was able to validly assess the relationship functioning of individuals when compared to the results of other well-established self-report measures. The results did, however, show some support for the discriminant validity between the GARF and the self-report measures for couples and families. The results of this study indicated that the GARF was more valid when assessing couple and family functioning, with more than one member present, than it was for assessing individual relationship functioning with only the individual present at the interview. Overall, the GARF produced results in the expected direction, with higher GARF scores being reported for the volunteer groups than the clinical groups.

Two types of validity were being assessed in this study regarding the GARF: concurrent and discriminant. Concurrent validity for the GARF was supported by the volunteer individuals and couples. In all four other groups, statistical significance was not obtained. Discriminant validity was supported by the couple and family groups. Therefore, concurrent validity was supported for two of the six groups and discriminant validity was supported for four of the six, with only one group being supported consistently by both concurrent and discriminant validity. The support for the two types of validity is very inconsistent in this study. Based on these results, the validity of the GARF as a global assessment measure in its current computer form is questionable.
Limitations

A major limitation in this study was the lack of a random sampling strategy. This study was done on a respondent basis, which did not allow for random sampling. This limitation increases the possibility of bias, decreasing its generalization to others. In addition, although substantial efforts were made to involve more families in the study, participation was low, which was likely due to the limited number of families attending Utah State University with children over the age of 12. This limits the generalization of the findings to other families external to this study. The other groups also were small, which therefore limits the generalization of the findings with the entire study.

Another limitation regarding the couple data is the large standard deviations in both the DAS scores and the GARF scores. Because the mean scores of the DAS with both the clinical (102.39) and volunteer (118.54) groups were close to the clinical cutoff range (107), the large standard deviations show that the difference may not be as significant as it appears when looking solely at the mean scores. The effect sizes, however, for the couple population did agree with the t-test results on the level of significant difference between the two groups on the DAS, with the effect size suggesting practical differences between the two groups of couples. Therefore, although caution must be used because of the large amount of variance found in the group scores, the effect size suggests that the large amount of variance did not affect
the significance of the couples' results. The large variance in score results is true not only for the couple population, but also for the individuals and families as well.

The volunteers used in this study were self-selected. The student researcher recruited participants from six different university classes consisting of an average of 114 students per class. Only 35 of these were included in this study, indicating a small response rate. Because the volunteers were self-selected, there is a possibility that this group was different from the average university population. One possibility is that these participants were in need of therapy and wanted to find out more about the Marriage and Family Therapy Clinic. Another possibility is that these participants were more highly motivated and educated, increasing their desire to participate. These factors, if true, would have biased the results of the study, lessening the generalization of the results.

One of the purposes of this study was to show that the GARF could account for overall relationship functioning similarly with individuals, couples, or families. Because of the many limitations of this study, in addition to the lack of significance between individuals scores on the GARF and their self-report scores on the OQ-45.2 symptom distress subscale, it is premature to conclude that the GARF can assess individual, couple, and family populations equally well.
Implications for Therapy

Although there are many limitations to this study, some implications for therapy can still be made based on the existing results. One implication of the findings is that therapists need to be cautious of their perceptions in order to avoid bias. The suggestion given by Charny (1980), that therapists should not view clients in a right and wrong way, but in more constructive ways that allow families to balance between the good and the bad and that allow them to find an environment which works best for them, is advice that could be useful in therapy. Although the results of this study do not give any conclusive evidence of therapist bias on the therapists' assigned GARF scores, it could be a potential problem that may affect the validity of the GARF scores.

These findings imply that it may be beneficial, as suggested by several researchers (Olson, 1976; Peck & Shapiro, 1990), that self-report and observational assessments be used in conjunction with one another to best assess client functioning. An approach used often by researchers to assess convergent validity between assessment measures is multitrait-multimethod (Morey & LeVine, 1988; Reichardt & Coleman, 1995; Sidani & Jones, 1995). In a multitrait-multimethod approach, various traits are assessed using multiple methods of assessment. Sidani and Jones (1995) used the multitrait-multimethod to analyze family relational data by comparing family functioning among deaf- and hearing-parented families. Using this method, the
researchers were able to determine the pattern of agreement among family members regarding the parents’ childrearing skills.

The multitrait-multimethod research approach is similar to Figure 1 presented earlier regarding four research approaches (Olson, 1976). The figure emphasized that there is a variety of methods of doing research, each tapping into different aspects. Through using solely the GARF, important symptoms may be missed which are critical to the assessment and intervention process. Through using solely client self-report, the overall picture may be missed, which may potentially slow the intervention process of therapy. By using both styles of assessment, both sides of the coin can be assessed, giving a more complete picture for developing and implementing a treatment plan.

A final implication for therapy is that caution may need to be used when using the GARF as a global assessment for relationship functioning with individuals. In this study, the GARF scores were unrelated to the scores obtained on the symptom distress scale of the OQ-45.2 for individuals. It may be difficult for a therapist to pick up on interactional patterns of individuals because of the inability to directly observe these interactions in process. Therefore, with individuals, therapists may need to rely more on self-report assessments and use the GARF as a secondary measurement of functioning.
Recommendations

This study has contributed to the current body of research that has been done on global assessment measures, particularly on the effectiveness of the GARF as a global assessment of relationship functioning for individuals, couples, and families. However, due to the many limitations of this study, conclusions about the concurrent and discriminant validity of the GARF as a global assessment measure are still tentative. Further research is necessary in order to define more clearly if the GARF can adequately assess overall relationship functioning, as well as its concurrent validity with standardized assessment measures.

Since both concurrent and discriminant validity of the GARF as a global assessment still need more establishment, therapists should be cautious of using the GARF as their primary assessment of relationship functioning. One recommendation is that the GARF be used in conjunction with other standardized assessments in order to check for accurateness of the GARF scores obtained through the therapist’s observations. This will help ensure that therapist bias does not negatively influence the therapeutic intervention process.

In addition, a recommendation is that therapists who use the GARF familiarize themselves well with the procedures of obtaining an accurate GARF score. The procedures to obtain a GARF score based on observations are outlined in the DSM-IV (1994). Although these procedures are somewhat vague and limited in their criteria for obtaining a GARF score, they give the therapist a base from which to
properly use the measure. Specific information is necessary in order for the therapist to calculate an accurate GARF score. Without knowing what information is important to obtain, a therapist could easily make inaccurate assumptions, leading to an invalid GARF score. Adequate training and understanding of the GARF is essential for effectively using the measure in the way it was designed.

In order to best meet this recommendation, additional information will need to become available outlining more clearly the instructions for properly using the GARF as a global assessment. With additional clarification on the usage of the GARF assessment device, the GARF could become a valuable and reliable assessment that could serve as a complementary measure to other microanalytic and global assessments.
REFERENCES


Charny, I. W. (1980). Why are so many (if not really all) people and families disturbed? *Journal of Marital and Family Therapy, 6*(1), 37-47.


Social Science Citation Index (Vols. 1-3) (1994). Philadelphia: Institute for Scientific Information.


APPENDIX
Letter of Permission to Use
Copyrighted Material
April 7, 1997

Dr. David H. Olson
Family & Social Sciences
290 McNeal Hall
University of Minnesota
1985 Buford Ave.
St Paul, MN 55108

Dear Dr. Olson:

I am in the process of preparing my thesis in the Family and Human Development Department at Utah State University. I hope to complete in the Spring of 1997.

I am requesting your permission to include your "Figure 1: Four Research Approaches" found on page 518 in your book entitled "Treating Relationships (Olson, 1979)." I will include acknowledgements and/or appropriate citations to your work as shown and copyright and reprint rights information in a special appendix. The bibliographical citation will appear at the end of the manuscript as shown. Please advise me of any changes you require.

Please indicate your approval of this request by signing in the space provided, attaching any other form or instruction necessary to confirm permission. If you have any questions, please call me (Ben Ashcraft (801)-797-6683) or my advisor (Scot Allgood (801)-753-5895) at the numbers listed.

I hope you will be able to reply immediately. If you are not the copyright holder, please forward my request to the appropriate person or institution.

Thank you for your cooperation,

Ben Ashcraft, MFT Graduate Student

Scot Allgood, Faculty Advisor
I hereby give permission to Ben L. Ashcraft to reprint the following material in his master's thesis.


(Signed)  

[Signature]

Family Social Science

U.S. Minnesota

St. Paul, MN 55108