2016-06-01

Initial Development and Validation of the Clinically Adaptive Multidimensional Outcome Survey

Jason Andrew McBride
Brigham Young University

Initial Development and Validation of the Clinically Adaptive Multidimensional Outcome Survey

Jason Andrew McBride

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

P. Scott Richards, Chair
Brent D. Slife
Richard R. Sudweeks
G.E. Kawika Allen
Tim Smith

Department of Counseling Psychology and Special Education

Brigham Young University

June 2016

ABSTRACT

Initial Development and Validation of the Clinically
Adaptive Multidimensional
Outcome Survey

Jason Andrew McBride
Department of Counseling Psychology and Special Education, BYU
Doctor of Philosophy

There has been a long-standing need in the field of psychotherapy to document progress and show effectiveness. The evidence-based practice (EBP) movement has had considerable influence in the field of psychology as evidenced by the APA task force that adopted the stance of evidence-based practice in psychology (EBPP) to ensure quality and accountability for psychological services as well as the integration of science and practice. One of the primary components of EBPP is the use of routine outcome measures (ROMs), which seek to integrate research with practice while simultaneously documenting progress and enhancing treatment. Despite the wave of ROM in the field, implementation rates have remained low. Research has brought forth many practical and philosophical concerns of therapists using these measures in routine practice including time burden and local validity.

The Clinically Adaptive Multidimensional Outcome Survey (CAMOS) was created to directly address clinicians' concerns with a specific focus on concerns of local validity. The CAMOS was designed to monitor several dimensions of functions, thus covering a wide range of issues. In this study the item pool proposed for the CAMOS was factor analyzed, and acceptable fit was found for a 6-factor model that contained 42 items. The 6 factors include (a) psychological distress, (b) relationship distress, (c) therapy expectations, (d) spiritual distress, (e) physical health distress, and (f) work/school distress. It is of note that spirituality emerged as a distinct factor with this data set and the implications and applications are discussed. With this multidimensional foundation, clinicians could more flexibly use the CAMOS to increase local validity. Clinical applications and future directions are discussed.

Keywords: outcome measure, clinically adaptive multidimensional outcome survey, camos, evidence based practice, spirituality, research-practice gap

ACKNOWLEDGEMENTS

I would like to thank my father, who expressed a vote of confidence in me early on in my educational path that opened my mind to options I had preemptively foreclosed on. I would like to thank my committee members for their feedback and suggestions, which significantly strengthened the statistical analyses and deepened the theoretical rationale. I would like to express my gratitude for my good friend Peter Sanders with whom I worked very closely to complete this project and whose brilliance and creativity substantially altered its course. I would also like to thank my chair Scott Richards whose vision and sagacity set in motion a project of great importance and high impact. I have appreciated the confidence, trust, and autonomy he has given me in taking a small part of it. Finally, I would like to thank my wife who supported me every step of the way with patience and perseverance.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS iii

TABLE OF CONTENTS iv

LIST OF TABLES vii

LIST OF FIGURES viii

DESCRIPTION OF DISSERTATION STRUCTURE AND CONTENT ix

Introduction 1

Practice-Based Evidence 3

Routine Outcome Measurement 4

Limitations of Routine Outcome Measurement 5

The Clinically Adaptive Multidimensional Outcome Survey 7

Statement of the Problem 8

Statement of Purpose 9

Research Questions 9

Method 10

Development of the Clinically Adaptive Multidimensional Outcome Survey 10

Brigham Young University–Idaho Sample 12

Training/integration 12

Data collection procedures 12

Participants 14

Center for Change Sample 14

Data collection procedures 14

Participants 15

Data Analysis 15

Data sets 15

Factor analysis 16

Results 18

Exploratory Factor Analysis 1 18

Item Total Correlations and EFA 2 19

Evaluation of a Potential Method Effect 20

EFA 3 21

Confirmatory Factor Analysis 22

Discussion 28

Factor Structure and Multidimensionality 28

Method bias 30

Meaningful subscales 31

Spiritual Distress Dimension 32

Brevity and Flexibility 32

Recommendations for Clinical Practice 33

Limitations 39

Strengths 40

Future Research Directions 40

Conclusion 44

References 45

APPENDIX A: REVIEW OF THE LITERATURE 57

References 83

APPENDIX B: SUPPLEMENTARY TABLES AND FIGURES 94

LIST OF TABLES

Table 1 *Parallel Analysis for EFA One* 18

Table 2 *Confirmatory Factor Analysis Fit Statistics* 23

Table 3 *Intercorrelations Among Factor* 24

Table 4 *Cronbach's Alpha for Final Factors* 25

Table 5 *Confirmatory Factor Analysis: Factor Loadings* 26

LIST OF FIGURES

Figure 1 Comparison of the scree plot of random eigenvalues versus the scree plot of eigenvalues from the observed data. 19

DESCRIPTION OF DISSERTATION STRUCTURE AND CONTENT

This dissertation, *The Initial Development and Validation of the Clinically Adaptive Multidimensional Outcome Survey*, is written in a hybrid format. This hybrid format integrates traditional dissertation requirements and journal publication formats.

The initial pages of this dissertation are for the purpose of fulfilling requirements for submission to the university. The remainder of the dissertation is written in a format that will allow it to be converted for a journal submission.

The review of the literature is included in Appendix A. Appendix B contains supplementary tables and figures that were deemed helpful but not necessary in the body of the dissertation.

There are two reference lists contained in this dissertation. The fist reference list contains references included in the journal-ready article. The second reference list includes all the references for the review of the literature.

Introduction

During the past 50 years, a great deal of research has been done in the field of psychotherapy (Lambert, Bergin, & Garfield, 2004). This research has provided much insight into the processes and outcomes of psychotherapy. It has also provided support for the effectiveness of psychotherapy in general and for many specific types of psychotherapy, such as behavioral, cognitive, humanistic, family, and group psychotherapy approaches (Lambert, Bergin, & Garfield, 2013).

Despite the large amount of research supporting the effectiveness of various types of psychotherapy, relatively little is known about what types of psychotherapy approaches are most effective for what types of clients and what types of clinical issues (Roth & Fonagy, 2005). The lack of specificity in the knowledge base about what approaches are most effective for specific types of clients and clinical issues has contributed to the infamous research–practice gap in the field of psychotherapy (Drabick & Goldfried, 2000; Kazdin, 2008; Ogilvie, 2012). The research–practice gap refers in part to the fact that empirical research has little influence on most practitioners. The treatment interventions which practitioners select and use with clients are based more on theoretical rationales and clinical experience than on the findings of psychotherapy research (Beutler, Williams, Wakefield, & Entwistle, 1995; Cohen, Sargent, & Sechrest, 1986; Cook, Schnurr, Biyanova, & Coyne, 2009; Lucock, Hall, & Noble, 2006; Morrow-Bradley & Elliot, 1986; Stewart & Chambless, 2007). The research–practice gap also refers to the lack of communication between researchers and practitioners. Researchers often fail to include practitioners in the process of designing their studies, which leads to research questions, methods, and findings that practitioners perceive are irrelevant to practice-related needs.

In response to concerns about the research–practice gap, and because of frustration that research findings were having so little influence on psychotherapy practice, an empirically supported treatment movement began within Division 12 of the American Psychological Association (APA) during the 1990s. Division 12 created a task force that developed a list of psychotherapy approaches they considered empirically supported treatments (APA, 1995). The list of empirically supported treatments, and the criteria used by the task force to select these treatments, bred substantial controversy and criticism (Addis, 2002; Addis & Krasnow, 2000; Goldfried & Wolfe, 1998; Kettlewell, 2004; Persons & Silberschatz, 1998).

Seeing the need for revision of how evidence-based practice (EBP) was defined in psychology, the APA assembled a presidential task force that established a definition for EBP that focused on principles rather than specific treatments. This report emphasized the importance of using multiple methods for building an evidence base. This is a substantial shift from empirically supported treatments (EST), where treatments could only be validated by the use of rigorous randomized controlled trials (RCT). There is no doubt that RCTs play an important role in establishing best practice, but there are certainly other methods of establishing evidence for effective treatment and outcome. Specifically they mention clinical observation, qualitative research, systematic case studies, single-case experimental designs, public health and ethnographic research, process/outcome studies, effectiveness research, and meta-analyses (APA, 2006). Clearly the APA's current stance is that a wide variety of research methods should be used in establishing a robust and diverse evidence base in psychology. It is interesting to note that among the list of research designs that the task force article highlighted, most of the designs fall under the category of practice-based evidence (PBE; e.g., effectiveness research, public

health and ethnographic research), and most of these designs require repeated measurement during treatment (e.g., process/outcome research, single case experimental designs).

Practice-Based Evidence

Practice-based evidence is one form of research that is significantly different from the traditional RCT methodology, but in many ways is complementary (Barkham, Hardy, & Mellor-Clark, 2010). One of the main differences from RCTs is that practice-based evidence focuses on collecting data in a naturalistic setting, which significantly increases its generalizability—a significant limitation of RCTs. In PBE studies it is important that therapists continue their treatment as usual and that they do not use a treatment manual approach. Additionally, it is not necessary to monitor the implementation of that treatment to ensure that it meets a certain criteria or protocol. Not only do these guidelines make results more generalizable and applicable, but they also make it feasible for studies to be implemented more often and for longer durations. Another feature of PBE is that all subjects presenting for treatment are included in the study, which creates heterogeneity with regard to personal characteristics as well as presenting problems. This not only allows for quicker data collection, but also directly results in greater applicability of findings. Ownership by practitioners is an important feature of PBE. Practitioners and site managers are the driving force behind the research questions being addressed, and they are the ones that have ownership of the data. This is in stark contrast to RCTs where every element of the study is designated by the researchers. Other important aspects of PBE include the formulation of research questions, the focus on practice improvement, benchmarking, and large data sets.

While many of the rigorous criteria of RCTs have been relaxed, it is important to note that there is still a strong emphasis on using reliable and valid measures. However, their application takes a more practical focus, accommodating for the site and existing procedures,

rather than focusing on experimental control. Therefore reliable and valid routine outcome measures are essential to most types of PBE designs. It is through these measures that data can be collected in a naturalistic setting from every client.

Routine Outcome Measurement

This presidential task force defined evidence-based practice in psychology (EBPP) as "the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (APA, 2006, p. 273). With the current definition of EBPP the burden falls on practitioners to show they are incorporating research with clinical expertise while taking into account specific client characteristics. The report from the task force also stated that "ongoing monitoring of patient progress and adjustment of treatment as needed are essential to EBPP" (APA, 2006, p. 280). In order to monitor the progress of patients, it is essential to engage in routine outcome measurement (ROM). ROM simply means that practitioners and researchers collect quantitative and qualitative measurements of patient processes and outcomes during the course of treatment. ROM can help establish the effectiveness of treatment, document progress, and serve as an aid in decision making for therapists (Asay, Lambert, Gregersen, & Goates, 2002; Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004; Lambert et al., 2003; Leon, Kopta, Howard, & Lutz, 1999; Lueger et al., 2001; Miller, Duncan, Sorrell, & Brown, 2005; Slade, Lambert, Harmon, Smart, & Bailey, 2008; Whipple et al., 2003). Additionally, therapists have reported using specific ROM questions to stimulate relevant dialogue in session (Callaly & Hallebone, 2001; Garland, Kruse, & Aarons, 2003; Trauer, 2010).

The APA taskforce statement specifically includes recommendations that there is a need for ". . . developing well-normed measures that clinicians can use to quantify their diagnostic judgments, measure therapeutic progress over time, and assess the therapeutic process . . ." (APA, 2006, p. 278). ROM addresses the three major decision-making factors outlined by the

EBPP task force to varying degrees of success. ROMs incorporate research as far as the ROM itself is research based, reliable, and valid. Analyzing the aggregate results of the ROM used perpetuates the research process as well. ROM can also enhance clinical judgment by providing the clinician with information about patient progress, allowing them increased responsiveness to potential problems. In this way the ROM allows the therapists access to up-to-date client characteristics that can inform the therapeutic process.

Limitations of Routine Outcome Measurement

Many research studies have been conducted to create and validate ROMs. However, the implementation rate of these ROMs by practitioners remains relatively low (Bewick, Trusler, Mullin, Grant, & Mothersole, 2006; Bickman et al., 2000; Brand, 2008; Garland et al., 2003; Gerdes, Edmonds, Haslam, & McCartney, 1996; Gilbody, House, & Sheldon, 2002; Hatfield & Ogles, 2004; Patterson, Matthey, & Baker, 2006; Ventimiglia, Marschke, Carmichael, & Loew, 2000; Zimmerman & McGilchey, 2008). Implementation rates are so low, in fact, that there are substantial amounts of research studies dedicated to investigating the reasons behind the lack of implementation of these empirically validated ROMs (Abrahamson, 1999; Garland et al., 2003; Gilbody et al., 2002; Lakeman, 2004; Roth & Fonagy, 2005; Trauer, 2010; Zimmerman & McGilchey, 2008). This research has yielded several reasons for lack of implementation including concerns with practicality, local validity, and feasibility.

In a study done by Garland and colleagues (2003), therapists reported that the outcomes prescribed by the ROM simply didn't capture the nuances and complexities of their clients or their therapy. This is a profound concern that challenges the adequacy of the principles of standardization and generalizability in psychological assessment and outcome measurement. These principles are grounded in the philosophy of abstractionism, which is one of the naturalistic assumptions that often goes unexamined in psychology and science (Slife, 2004).

When grounded in the philosophy of abstractionism, the very nature of psychological measurement is to overlook the nuances and complexities of each individual in favor of standardization and generalizability (Sonnanburg, 1996).

ROMs follow these assumptions and principles, which is demonstrated by the fact that they are developed and designed explicitly to apply to all persons, contexts, races, and religions; every socioeconomic status; and every presenting concern. This methodology has its place and yields valuable information, but it should not be seen as the *only* way to measure outcome. This is a glaring example of how defining outcome through a solely universalistic paradigm provides a one-sided view of outcome and consequently an impoverished understanding of assessment and treatment.

The assumption of abstractionism is of obvious concern to the practitioners in the studies cited earlier as evidenced by their complaints of local validity and the overall lack of implementation of these ROMs, but it should also pose significant concern to the APA considering the theoretical conflicts inherent in statistical standardization. This solely abstractionist approach to ROM is in direct conflict with the definition of EBPP put forth by the APA which states that "Psychological services are most effective when responsive to patient's specific problems, strengths, personality, sociocultural context, and preferences" (APA, 2006, p. 284). If researchers were to include these characteristics into their ROM, then practitioners would be more likely to see their value and utility, which would result in higher implementation rates, and the aims of EBPP would be better realized.

Another issue that arises when discussing the assessment of population-specific or individualized assessment is feasibility. It is virtually impossible to create a measure that takes into account all possible population-specific characteristics, cultures, or preferences. A possible

solution to this problem has been to use several different measures that have been designed for specific populations or problems. This does address concerns of local validity and begins to step away from a solely abstractionist approach, yet this becomes cumbersome to administer and can be quite lengthy if a client is taking multiple measures every week. Thus, the problem remains that researchers seek to develop standardized measures that meet criteria for reliability and validity (abstractionism), but practitioners feel the outcomes being measured are not relevant or specific enough (contextualism) to be useful. A potential solution for this dilemma is to create a hybrid assessment system that maintains universality while also incorporating a greater level of contextual applicability.

The Clinically Adaptive Multidimensional Outcome Survey

The Clinically Adaptive Multidimensional Outcome Survey (CAMOS) was designed for this purpose. The CAMOS builds upon two theoretical ideas from two different paradigms. The first feature of the CAMOS is designed to incorporate traditional standardized statistical assessment. The eight dimensions of the CAMOS globally assess eight aspects of human experience: (a) therapy progress, (b) relationships, (c) distressing behaviors, (d) distressing thoughts, (e) distressing emotions, (f) spirituality, (g) work/school, and (h) physical health. This multisystem structure was created using a theoretical rationale (Lazarus, 1973, 2007; Richards & Bergin, 2005) and is also generally supported empirically (Miovic et al., 2006; Slade, 2002). It is important to note the inclusion of a dimension that assesses spirituality, which is another way in which the CAMOS increases relevance and local validity. By assessing spirituality the CAMOS taps into an aspect of life relevant to treatment that most outcome measures ignore. In this way even the standardized portion of the CAMOS will have greater applicability and breadth than most outcome measures by taking into account this aspect of a clients' "characteristics, culture, and preferences (APA, 2006, p. 273)." Further explanation of the development and definitions of

the dimensions will be provided in the methods section, but suffice it to say here that the eight core dimensions are designed to be applicable across culture, race, ethnicity, and religion.

The second feature of the CAMOS provides opportunities for the therapist to assess clients through a contextual and relational approach. The CAMOS's global, eight-dimensional structure can be expanded to include population specific dimensions (e.g., eating disorder dimension, cultural sensitivity dimension). It also allows therapists and clients to tailor the assessment to their unique context and relationship by creating unique items that apply only to that client and their work in therapy, an idea that is based in the area of individualized assessment (Haynes, Mumma, & Pinson, 2009).

The CAMOS seeks to combine these two very different paradigms into an integrated system that combines standardization (facilitating comparability) with customization (facilitating local relevance). The first step in this process is to establish a multidimensional standardized base that will allow for flexibility and adaptability.

Statement of the Problem

Evidence-based practice in psychology has established a foundation for quality in the field of psychology. Routine outcome measures have made significant progress in defining EBPP, including using research, clinical expertise, and patient characteristics. However, several limitations of ROMs, including limitations of local validity and feasibility, are inhibiting their overall effectiveness, acceptance, and use in mainstream practice. Currently, there are no routine outcome measures that directly address these practical and theoretical concerns posed by practitioners, although the CAMOS has been developed for this purpose. However, in order for the CAMOS to be of any value to the field, it is necessary to better understand its psychometric properties.

Statement of Purpose

The purpose of the present study is to investigate the factor structure of the CAMOS and find the best fitting factor model for the data sets. This will help shorten its length and increase inter-item reliability by eliminating items that do not load cleanly into the best fitting factor model. Examining the factor structure will help in determining the quality of the individual items, test the theoretical eight-factor model, and determine whether there are empirically derived factor models that better fit the item pool. This will help us determine whether the CAMOS will be useful as a ROM suitable for research and practice in the psychotherapy field.

Research Questions

1. What is the factor structure of the CAMOS data set?
2. Is spirituality distinct from other areas of life, and will it emerge as a distinct factor?

Method

Development of the Clinically Adaptive Multidimensional Outcome Survey

Initial structural foundations and conceptualization of the Clinically Adaptive Multidimensional Outcome Survey (CAMOS) builds on the theory of multimodal therapy (Lazarus, 1973, 2007), which argues a rationale for monitoring several modalities of human functioning in client assessment and treatment. In Lazarus's model there are seven distinct modalities including (a) behavior, (b) affect, (c) sensation, (d) imagery, (e) cognition, (f) interpersonal relationships, and (g) drugs. Richards and Bergin build on this idea with their theory of multilevel, multidimensional assessment strategy (2005). This theory postulates that there are seven different systems that therapists should assess when working with clients. These systems include (a) emotional, (b) social, (c) physical, (d) educational or occupational, (e) behavioral, (f) spiritual, and (g) intellectual or cognitive concerns.

The inclusion of a spirituality dimension is a unique addition that is not part of Lazarus's research nor is it a common area of assessment. However, there are significant theoretical rationales and empirical findings that suggest assessing spirituality in routine practice is needed (Richards & Bergin, 2005; Smith, Bartz, & Richards, 2007). For the CAMOS we adopted and adapted the seven theoretical systems proposed by Richards and Bergin and created a dimension that assessed each system. Additionally, a therapy progress dimension was created to monitor the therapeutic relationship. In addition to the theoretical rationale, there is general empirical support for the inclusion of these dimensions in assessing and treating clients. Karstin Slade reviewed 16 studies that suggested outcome domains for use in mental health services (2002). Slade analyzed and grouped these proposals into seven emergent categories (i.e., wellbeing, cognition/emotion, behavior, physical health, interpersonal, society, and services). Out of the seven domains identified, five corresponded directly with the content of the proposed CAMOS domains, and

seven out of the eight domains of the CAMOS were represented among Slade's seven categories. In a study done by Miovic and colleagues, 17 domains of discussion in psychotherapy were ranked by importance for treatment by both clinicians and clients (2006). Of the top 10 domains of relevant discussion, 7 directly corresponded with the proposed 8 dimensions of the CAMOS. Please see Table A1 in Appendix A for a more detailed description of Slade's and Miovic's domains.

Initial item creation for the CAMOS followed the theoretical framework described above, and individual items were created in an effort to assess each of the eight systems. A team of three psychologists and three graduate students created the initial pool of items to assess each dimension. Specific efforts were made to create questions that assessed a wide range of concerns within each dimension that focused only on identifying the level of distress in each system. These items were then presented to three other PhD level therapists who own a private practice in Utah. These therapists gave feedback and offered suggestions to change questions or add additional questions. The revised items were then presented to 14 therapists at the BYU–Idaho (BYUI) counseling center for further review and revision. Specifically, these therapists were also asked to give input on what questions they would like to ask clients from session to session within each dimension. One of the most significant suggestions made by these therapists was to add positively worded items. They expressed that the measure seemed overly negativistic and that measuring positive responses would add unique and rich information that could not be obtained through questions that are solely negatively worded. Subsequently we added items to each dimension or changed the valence of existing questions to accommodate this suggestion. With regard to the dimensions, the therapists expressed agreement that the eight dimensions measured were relevant and useful for assessment and treatment.

Brigham Young University–Idaho Sample

Training/integration. After the initial creation of the questions included in the CAMOS we met with the Brigham Young University – Idaho (BYUI) counselors and secretaries to delineate the most efficient and practical way to administer the CAMOS to students. We wanted to ensure that we tailored the administration process to the needs and existing procedures of the center.

Data collection procedures. It was decided that clients would be entered into to the Qualtrics online survey software panel via an online survey (Secretary Panel Entry) that the secretaries would access. When a student set up a counseling appointment the secretaries filled out the Secretary Panel Entry survey which included their BYUI student ID and the therapist ID, which we created specifically for this study. The BYUI ID is a nine-digit number and the therapist ID is a simple two-digit number. Names were not used in the CAMOS system to protect the confidentiality of the students. Upon completion of the Secretary Panel Entry survey the clients' data were entered into a panel that served two purposes: (a) This panel is referenced when a client enters their BYUI ID when taking the CAMOS survey. If their student ID is not in the panel then they cannot take the survey. This helps to ensure that only students who have signed up for counseling and agreed to participate in the study will take the survey. (b) This panel list also allowed the researchers to be alerted to new clients so they could create an individualized report for therapists.

When students of participating therapists presented for their first psychotherapy session at the BYUI counseling center, they were informed by the secretary that certain therapists at the treatment center were participating in a research study in collaboration with Dr. Scott Richards, at Brigham Young University, to assess outcomes of psychotherapy treatment, and that they

were inviting all of their clients to participate in the study with them. Potential participants were given the informed consent form, which explained the procedure and purpose of the study and the fact that it was voluntary. If they agreed to participate they were asked to sign the document and then given a Kindle Fire HD to fill out the online survey. Secretaries typically entered the student ID at the first page of the survey for convenience. The student then proceeded to take the CAMOS, which included demographic questions such as age, gender, ethnicity, and religious affiliation. The CAMOS itself included about 70 questions that span the eight dimensions (e.g., relationships, behaviors, emotions). The client filled out the survey using Qualtrics, an online survey software. The CAMOS took about 15 minutes to complete at intake and therefore the students were instructed by the secretaries to arrive approximately 15 minutes early to their first session.

It is important to note that the therapists had access to the results of their clients' answers via an online report created and distributed by the researchers through the Qualtrics online survey software. These reports were only provided as a potential benefit for the therapist and client, and therefore had no effect on the intake administration and were irrelevant to the current study. It is also important to note that clients took the adaptive form of the CAMOS before each subsequent counseling session. The therapy expectations questions were adapted from future tense wording to present tense wording to assess "therapy progress" rather than "therapy expectations." The adaptive form was created to shorten the amount of time it takes for the client to fill out the survey, thus making it more practical. However, validating this adaptive feature of the CAMOS was not a focus of the current study, and only the results acquired at intake were aggregated and analyzed.

Participants. Therapists were selected from the BYUI counseling center to participate in the study. Seven therapists agreed to participate, each with a doctoral or master's degree. All clients of participating therapists that presented for treatment from March 2013 to March 2014 were invited to participate in the study. The total number of students included in the study was 304. The mean age of clients was 21.8 with a standard deviation of 3.9. There were 116 males (38.2%) and 186 females (61.2%) included in the study. The great majority were members of the Church of Jesus Christ of Latter-day Saints (98.7%). The following ethnicities were represented in their respective proportions: American Indian (1.0%), African American (0.7%), Asian (1.7%), Latino/a (7.9%), Hawaiian Native/Pacific Islander (0.3%), White/Caucasian (76.3%), and missing (12.1%). Clients presented with a wide range of concerns including depression, obsessive-compulsive disorder, generalized anxiety disorder, and impulse control disorder.

Center for Change Sample

Data collection procedures. Data were also collected in Utah at the Center for Change (CFC) which is a residential treatment facility for women with eating disorders. Data collection procedures at the (CFC) were somewhat different given that the center was already collecting data using other assessment batteries. The existing methods for data collection at the center were varied and included multiple forms of computer administration and paper and pencil administration of various instruments. Due to the varied nature of instrument administration and data collection it was determined that it would be beneficial to convert most of their existing assessments into Qualtrics surveys. Surveys were administered as part of the treatment protocol within the first few days of admittance into the center by Care Techs.

Reports were also created in Qualtrics that relayed real time results to the therapists. With this new system the CAMOS was easily integrated into the assessment procedures for new patients. Every patient admitted to the CFC from May 9, 2013, to June 17, 2014, was asked to

complete the CAMOS as part of their treatment at the center. The Care Tech filled out a Qualtrics survey, which entered the patient medical record number into the Qualtrics system. This then allowed the patient to use their medical record number to access the CAMOS survey online. As per the existing CFC procedures, the patients filled out the assessments within the first few days they were admitted.

Participants. The total number of patients that completed the survey and thus were included in the study was 209. The mean age of clients was 23.1 with a standard deviation of 9.4. As the CFC is a treatment facility for women, 100% of respondents were female. Religious affiliation was represented as follows: Latter-day Saint (25.6%), Protestant Christian (9.5%), Roman Catholic (7.1%), Jewish (1.4%), Buddhist (0.9%), Hindu (.5%). Some respondents marked the "other" category (5.7%), 13.7% of respondents indicated that they were spiritual, but not religious, 12.8% reported being non-religious, and 22.7% of respondents left this item blank. The following ethnicities were represented as follows: American Indian (0.5%), African American (0.9%), Asian (0.5%), Latino/a (0.9%), Hawaiian Native/Pacific Islander (0.5%), White/Caucasian (67.3%), Alaskan Native (0.5%), Multiracial (Latino/a and White/Caucasian; 0.9%), Multiracial (African American and Asian; 0.5%), Multiracial (Asian and White/Caucasian; 0.5%), other (2.8%), missing (22.7%).

Data Analysis

Data sets. In order to factor analyze the data, the two samples were combined and then separated to form two data sets with subjects from each sample represented in each data set. The BYUI sample was stratified based on gender and then split randomly into two groups to ensure a comparable number of men and women in each data set. The CFC sample was split randomly into two data sets, and the two CFC data sets were combined with the two BYUI data sets to form two data sets that contained subjects from the BYUI sample (with a comparable number of

men and women) and subjects from the CFC sample. IBM SPSS Statistics Version 21 was used in separating and organizing the data. The first data set will be referred to as Data Set 1 and was used primarily to conduct exploratory factor analyses. The second data set will be referred to as Data Set 2 and was used to conduct confirmatory factor analyses (CFA).

Factor analysis. The first step was to categorize all the items into meaningful and interpretable factors. An exploratory factor analysis (EFA) was conducted on the 70 original items. The eight hypothesized dimensions were based primarily on a conceptual practical rationale, but lacked an empirically based rationale. It was not expected that the eight dimensions would emerge as eight factors; therefore an EFA would provide valuable empirical information about the factor structure of the CAMOS item pool. Items were treated as ordered categories and the weighted least squares means and variance adjusted (WLSMV) estimator was used due to the Likert scale format of the questions. Mplus statistical software was used for all EFAs and CFAs (Version 7.3, Muthen & Muthen, 1998–2014). Geomin (oblique) rotation was used as well as the weighted least squares extraction setting. Factor retention was based on three criteria: (a) Eigenvalues greater than 1.0 (Kaiser, 1958), (b) Scree plot analysis (where factors above the bend or elbow of the chart are retained; Cattell, 1966), (c) Parallel analysis (compares obtained eigenvalues to randomly generated eigenvalues; Horn, 1965). An analysis was run on Data Set 1 to determine the eigenvalues of the data set.

Once the number of factors was determined, item total correlations were obtained using IBM SPSS Statistics Software Version 21. The reliability analysis was used to calculate the corrected item to total correlation for each obtained factor in order to eliminate items that did not correlate well within each factor. Items were excluded if the corrected item to total correlation was less than 0.4.

It was anticipated that due to the presence of positively worded items a potential method bias might affect the data in unintended ways. This was addressed specifically in two ways: (a) conducting EFAs on an individual dimension if positively worded items grouped into a single factor, and (b) creating a method effect factor in a CFA and analyzing factor loadings.

The remaining items were subjected to EFAs with the same three-part factor retention criteria with an emphasis on the parallel analysis. Once a potential factor structure was found exclusion criteria were used to identify items that (a) did not load high on any factor (< 0.32) and (b) had their primary and secondary factor loading difference less than 0.2, suggesting that it did not cleanly load into one factor. Conceptual analyses were also conducted to make decisions about items that were close to the proposed cutoffs. Explicit efforts were made to retain as many useful items as possible, as the goal was to maintain a relatively large number of items that would be useful for clinicians to obtain a broad understanding of their clients. The main focus was to cleanly delineate factors to provide a meaningful categorization of items to facilitate therapist understanding of client functioning in multiple areas.

The obtained factor structure from the subsequent EFAs was then tested using confirmatory factor analysis on Data Set 2. At this point the remaining positively worded items were subjected to further analyses based on the possibility of a method effect. Modification indices were also examined and adjustments were made after a conceptual analysis of the suggestions.

Results

Exploratory Factor Analysis 1

Eigenvalues were obtained for Data Set 1, and factor retention based on eigenvalues greater than 1.0 suggested retaining 15 factors, while a visual analysis of the scree plot suggested retaining two factors. A Monte Carlo PCA was used to obtain random eigenvalues to be used in a parallel analysis, which yielded a result that fell between these two extremes and suggested retaining eight factors (see Table 1 and Figure 1).

Table 1

Parallel Analysis for EFA One

		Monte Carlo PCA	
Step	EFA One Eigenvalue	Random Eigenvalue	SD
1	21.775	2.242	0.067
2	4.423	2.129	0.050
3	3.437	2.048	0.040
4	3.221	1.978	0.036
5	2.863	1.919	0.032
6	2.412	1.863	0.030
7	2.115	1.811	0.028
8	**1.894**	**1.763**	**0.027**
9	1.591	1.717	0.026
10	1.506	1.674	0.025
11	1.344	1.633	0.024
12	1.257	1.593	0.023
13	1.228	1.554	0.021

Note. The number of factors to be retained is determined by the highest number of the step where the EFA One Eigenvalue is still greater than the Random Eigenvalue.

Among the three retention methods, parallel analysis is the least subjective method because it is statistically based and therefore is considered the best estimator for factor retention (Hayton, Allen, & Scarpello, 2004; O'Connor, 2000; Schmitt, 2011). Eight factors were retained, and items were categorized into the factor in which they loaded most highly.

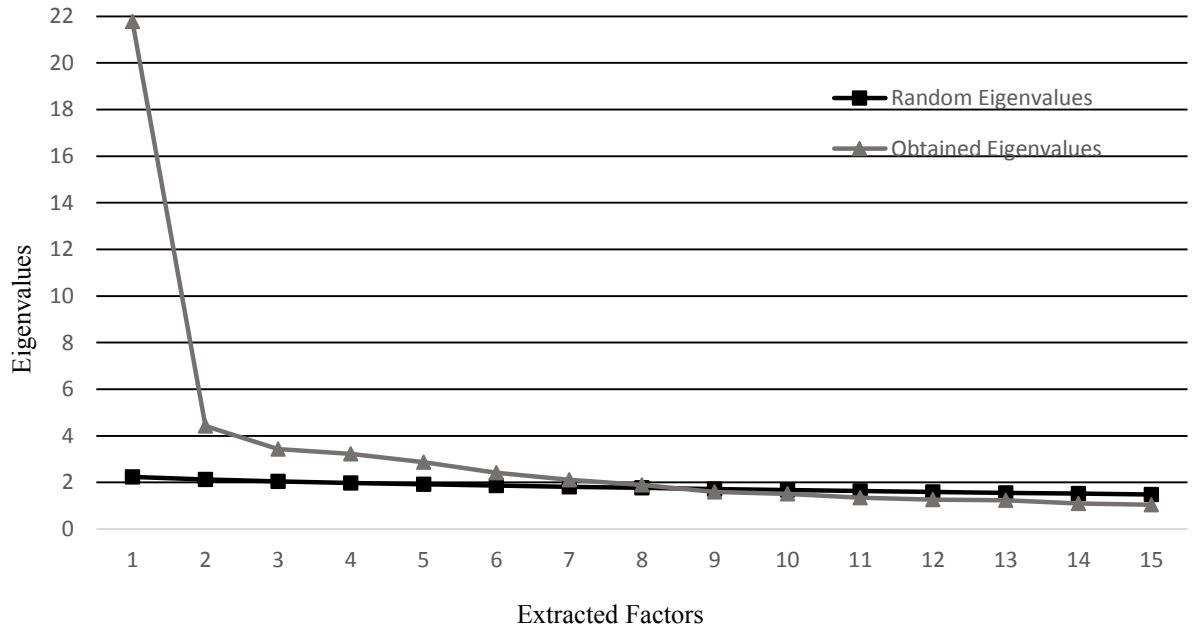

Figure 1. Comparison of the scree plot of random eigenvalues versus the scree plot of eigenvalues from the observed data.

Item Total Correlations and EFA 2

Item total correlations were calculated, and items were excluded if their corrected item-to-total correlation was less than 0.4. This resulted in the exclusion of four items from four different factors (i.e., psychological distress, behavioral concerns, relationships, physical concerns; see Table A2 in Appendix A). A second EFA was conducted including the 66 remaining items and resulted in a similar eight-factor structure based on parallel analysis.

Conceptual analysis showed distinct factors including: (a) therapy expectations, (b) relationships, (c) psychological distress, (d) behavioral concerns, (e) physical concerns, (f) spiritual concerns, and (g) work/school concerns. These factors were closely related to the theorized factor structure that was used when the items were created. However, one factor included only positively worded items. Due to the fact that there were only 15 positively worded items out of a total of 70 items the possibility of a method effect had to be explored.

Evaluation of a Potential Method Effect

Analyses were conducted to determine whether the identified positive dimension had any substantive value or if it represented a method effect. This analysis is based on the idea that these positively worded items may cluster together for two different reasons. One explanation is that there is a method bias and the reason for their correlations is not because of the content of the questions, but simply because they are positively worded. An alternative explanation is that the questions correlate because they are measuring a distinct and meaningful factor. To explore this possibility we analyzed all of the positively worded items that loaded more highly on the positive item dimension. It is important to note that some of the positively worded items loaded more highly onto other factors and these items were left in those factors, so only a subset of the positively worded items (9) was included in this dimension. This dimension was analyzed to determine if an EFA would suggest a multiple factor structure for this data set. If a multiple factor structure were indeed found then that may be evidence to support the idea that these items only correlated well together in the presence of the negatively worded items and were not measuring a distinct factor in and of itself. The second possibility is that these items would still correlate well together in the absence of the negatively worded items. Conceptually it is possible that the dimension could be measuring optimism or self-efficacy. When more than one item was included from a similar dimension correlations were high enough among those items to suggest a

multiple factor structure. However, when there were no two items from the same dimension a single factor structure emerged as the best fit for the data, suggesting the possibility that these items load onto a single factor due to their measuring the same latent variable and not just because they were positively worded. This resulted in a six-item dimension that was hypothesized to measure self-efficacy.

This dimension was then included in another EFA with 63 items to see if this would yield a viable factor structure. Parallel analysis again suggested retaining eight factors; however, eight-, seven-, and six-factor models fell apart when the exclusion criteria (i.e., primary factor loading less than 0.32, primary and secondary factor loading difference less than 0.2) were imposed. In light of these results, it was determined that method bias may be significantly affecting the analyses, and it may be best to exclude the positively worded items that loaded higher on the identified positive dimension. This left three positively worded items that seemed to load well into factors that made conceptual sense.

EFA 3

An EFA was conducted on the remaining 57 items, which indicated retention of seven factors. Items were then organized into the factors on which they loaded highest. The exclusion criteria were then imposed, and the dimension that was identified preliminarily as perhaps measuring behavioral concerns collapsed and had too few items to interpret meaningfully. The six-factor structure was then examined and after items were sorted and exclusion criteria imposed there were enough items in each dimension to support an acceptable factor structure. Thirteen items were identified as questionable based on the exclusion criteria. One item (WOR11) was retained despite meeting one of the exclusion criteria. It was retained because it had a decent factor loading (0.503) in its dimension, strengthened the factor conceptually, and its primary factor loading was 0.194 higher than its secondary factor loading which meant that it

missed the criterion by only 0.006. Another item (THO2101) was retained not only because it missed the primary and secondary difference of 0.2 exclusion criterion by only .045, but also because it was conceptually helpful for the factor to which it is connected. This resulted in 46 items spanning six dimensions (i.e., therapy expectations, relationship, psychological, spiritual, physical, work/school).

Confirmatory Factor Analysis

A CFA was then conducted testing the fit of the proposed factor structure established in the previous EFA. The initial results yielded fit statistics that were slightly below accepted standards and are reported in Table 2. Having retained three positively worded items, method bias was still of concern. This possibility was tested by creating a method factor that was composed of just the positively worded items. These items were also allowed to correlate with their respective factors to which they loaded most highly. Factor loadings were then analyzed to determine if the factor loadings were higher on the original factors or on the method factor. It was determined that two of the positively worded items correlated substantially higher with each other when compared to their factor loading in their respective factors. This may indicate that their shared variance is greater than the variance shared within their factors as to indicate the potential for significant method bias. This shared variance seems to be due to the positive wording of these items, and it was therefore determined that all three positive items should be removed due to probability of method bias. Another CFA was conducted without the suspected method bias factor, and model fit increased slightly after the removal of these items. Modification indices suggested a correlated between two items from the psychological distress factor, both of which were initially created as screening items and therefore have almost identical wording. These items also loaded into the same factor, and we determined that we could retain adequate construct coverage by deleting one of the items. This resulted in the final model for the

data set for which the fit statistics are shown in Table 2. Factor inter-correlations range from 0.30 to 0.69 (see Table 3). The final model results included 42 items and six dimensions: (a) therapy expectations, (b) relationship distress, (c) psychological distress, (d) physical health distress, (e) spiritual distress, (f) work/school distress (see Table 4). Reliability estimates by dimension using Cronbach's alpha yielded coefficients between 0.76 and 0.92 (see Table 4). Four out of the six dimensions fall above the optimal minimum standard of > 0.8 (Streiner, 2003) and the other two dimensions meet an acceptable standard of > 0.7 (Nunally & Bernstein, 1994).

Table 2

Confirmatory Factor Analysis Fit Statistics

Model	Items	X^2	X^2df	RMSEA [90% CI]	CFI	TLI	WRMR
Initial	46	2119.147	974	.067 [.063, .071]	.90	.89	1.4
Final	42	1552.947	804	.060 [.055, .064]	.93	.92	1.2

Notes. X^2 = chi-square, X^2df = *chi-square* degrees of freedom, RMSEA = root mean square error of approximation, CI = confidence interval, CFI = comparative fit index, TLI = Tucker Lewis index, WRMR = weighted root mean square residual

Table 3

Intercorrelations Among Factors

Factor	1	2	3	4	5	6
1. Therapy Expectations	-					
2. Relationships	.44	-				
3. Psychological	.47	.66	-			
4. Spiritual	.34	.47	.63	-		
5. Physical	.30	.52	.62	.38	-	
6. Work/school	.32	.57	.69	.55	.53	-

Rules of thumb have been established for evaluating goodness of fit (GOF), but the researchers proposing these rules of thumb have cautioned that they should be used as guidelines and not hard and fast cutoffs. With that being said, one of the most common recommendations for evaluating the root mean square error of approximation (RMSEA) is < 0.060 (Hu & Bentler, 1999), which is exactly the obtained value in the final CFA model. The general rule for the comparative fit index (CFI) and the Tucker Lewis index (TLI) has been > 0.90, which indicates "reasonably good fit" (Kline, 2005, p. 140; Raykov & Marcoulides, 2000). Hu and Bentler (1999) suggest a higher standard of "close to" 0.95. The obtained values of CFI = 0.925 and TLI = 0.920 fall between the general and more stringent criteria, which indicates acceptable or good fit for this data set. The obtained weighted root mean square residual (WRMR; 1.2) fell just above the guideline of < .1 proposed by Yu (2002), but the WRMR is still an experimental

statistic that does not yet have a firmly established guideline that is widely accepted and should therefore be interpreted with caution (Byrne, 2012).

Table 4

Cronbach's Alpha for Final Factors

	Factor	α
1	Therapy Expectations Distress	.82
2	Relationship Distress	.83
3	Psychological Distress	.92
4	Spiritual Distress	.83
5	Physical Distress	.78
6	Work/School Distress	.76

Table 5

Confirmatory Factor Analysis: Factor Loadings

	Item	Factors					
		1	2	3	4	5	6
1	I felt anxious about beginning therapy	.86					
2	I had concerns about beginning therapy	.85					
3	I felt uncertain about whether I can be fully honest and open with my therapist	.81					
4	I had doubts about whether my therapist will understand my concerns	.70					
5	I doubted whether therapy will be worth the cost	.46					
6	I felt misunderstood by my loved ones and friends		.78				
7	I felt concerned about my relationships (with your family, partner/spouse, and/or friends)		.74				
8	I had outbursts of anger		.71				
9	I hurt others with my words or actions		.69				
10	I felt sad about how I acted towards my family or friends		.68				
11	I felt irritated and angry towards others		.66				
12	I argued with my loved ones or friends		.61				
13	I felt hurt or disappointed by how my loved ones or friends behaved		.61				
14	I felt worried, agitated, fearful, or tense			.90			
15	I felt worthless or “not good enough”			.85			
16	I thought about past personal failures/mistakes			.84			
17	I felt sad or depressed			.84			
18	I felt stressed out			.82			
19	I felt concerned about distressing thoughts.			.78			
20	I felt powerless or stuck in my problems			.78			
21	I felt physically stressed or worn out			.69			
22	I felt concerned about my self-defeating behaviors:			.66			
23	I had thoughts or images that I couldn't get out of my head			.62			
24	I had difficulty concentrating or remaining focused on a task			.61			

Note. 1 = Therapy Expectations, 2 = Relationships, 3 = Psychological, 4 = Spirituality, 5 = Physical, 6 = Work/School.

Table 5 (continued)

Confirmatory Factor Analysis: Factor Loadings

	Item	Factors					
		1	2	3	4	5	6
25	I felt a loss of inspiration or spiritual direction				.91		
26	I felt distant in my relationship with God or my Higher Power				.84		
27	I felt guilt and regrets over mistakes that were inconsistent with my religious beliefs				.76		
28	I felt concerned about my religious or spiritual life				.68		
29	I experienced unsettling, troubling, or unusual religious thoughts				.58		
30	I felt light headed, weak, or fatigued					.85	
31	I harmed myself (cut, scratched, burned, etc.)					.79	
32	I wondered if I should see my physician					.64	
33	I experienced physical pain or discomfort					.64	
34	I had a stomach ache or other gastrointestinal problems					.62	
35	I experienced medical complications					.58	
36	I felt concerned about my physical health.					.52	
37	I felt over-burdened with too many responsibilities						.74
38	I felt concerned about my work (i.e., employment, school, homemaking responsibilities, volunteer work, etc.):						.73
39	I worried about not meeting expectations or requirements						.68
40	I wanted a change in my responsibilities						.59
41	I felt undervalued and unappreciated						.59
42	I was in trouble for the quality of my performance						.45

Note. 1 = Therapy Expectations, 2 = Relationships, 3 = Psychological, 4 = Spirituality, 5 = Physical, 6 = Work/School.

Discussion

Factor Structure and Multidimensionality

These results provide initial empirical support for the multidimensional structure of the CAMOS, and the majority of the originally hypothesized factors were maintained in the final factor model. It is rare for multidimensional theoretical measures to maintain their dimensions when factor analyzed, and it was not expected that each dimension would be maintained. However, we were pleased to see that five of the original eight dimensions remained largely unchanged, and two of the dimensions combined into one. The end result then shows only a few major changes and very little loss in overall construct coverage. The final factor structure model supported in this study included the following dimensions: (a) therapy expectations, (b) relationship distress, (c) physical health distress, (d) work/school distress, (e) spiritual distress, and (f) psychological distress.

The *therapy expectations* dimension was originally hypothesized as understanding the patient's expectations broadly for therapy, which included concerns as well as hopes. With the exclusion of three positively worded items and one negatively worded item, the therapy expectations dimension changed only slightly by only assessing stress or worry about the their anticipated therapy experience.

The *relationship distress* dimension remained primarily intact and did not change conceptually, except for the fact that it lost the positively worded items. Three positively worded items were excluded, but the remaining items focused on the client's distress around their interactions with others (e.g., arguing, not feeling understood, being sad or disappointed about how they acted or how others acted toward them). This dimension did gain three items from other dimensions, two from *behaviors* and one from *emotions*. Two of these questions had

"others" (referring to other people) in the question wording, which may indicate that while the questions had originally been intended to measure clients' individual behavior or emotions toward others, it seemed to tap more into the quality of their relationships and not specifically into their individual behaviors or emotions.

The *physical health distress* dimension retained six out of the nine items from the hypothesized theoretical factor and therefore remained conceptually the same. This dimension was intended to measure physical symptoms or concerns about their physical health. One question was dropped because it was positively worded and one question was dropped about side effects from medication. The dimension retained two questions that address the client's opinion on seeing a physician and one on experiencing medical complications, and therefore it seems that there is still adequate construct coverage.

The *work/school distress* dimension was originally intended to measure concerns primarily in the area of work and school, but question wording was broad enough to include any type of social responsibility. It is interesting that this dimension emerged as one of the weaker factors, manifest by the fact that it explained the lowest amount of variance and had the lowest internal consistency. However, of the original eight items in the hypothesized dimension only two were eliminated in the final model, and it was because they were positively worded. It seems that the construct remained intact, and further hypotheses that may explain the relatively low performance of this dimension will be discussed later.

The *spiritual distress* dimension was designed to measure the perceived quality of one's spirituality or religiosity. Two of the items that were eliminated from this dimension were positively worded, and therefore the construct shifted slightly to measure only concern or distress

specifically in the area of religion or spirituality. Only one other item was eliminated from this dimension, resulting in five of the eight original items being retained in the final model.

Two of the original dimensions, *emotions* and *thoughts*, combined into one factor, which we named the *psychological distress* dimension. These data indicate that thoughts and emotions are so highly correlated that it is not meaningful or useful to separate them into distinct factors. Another deviation from the obtained factor structure from the original was the absence of a behavior dimension. This may also be evidence for the idea that thoughts, emotions, and behaviors are so intricately connected that it may not be meaningful to try and separate them. As was stated previously, another potential explanation is that behaviors often tie closely to relationships. Three items that were originally designed to measure behavior were subsumed into the relationships factor (i.e., I had outbursts of anger, I hurt others with my words or actions, I felt sad about how I acted toward my family or friends). Each of these items were intended to measure a specific behavior, such as angry outbursts or hurting others with words or actions, but it is also evident that these behaviors often or necessarily play out in the context of relationships. Therefore, it is easy to see how they conceptually fit very well within the relationships dimension.

Method bias. The finding that the majority of positively worded items loaded onto a single factor is interesting but not unusual. Method effects have often shown up in the literature as a significant concern when creating questionnaires (Campbell & Fiske, 1959; Richardson, Simmering, & Sturman, 2009; Williams, Hartman, & Cavazotte, 2010). While having a mix of positively and negatively worded items helps break up "yea-saying" or "nay-saying" response sets, it is important to consider the effect that just a portion of positively worded items can have on the statistical properties of the data set. Possible effects of method bias could be over inflated

reliability and validity estimates (Podsakoff, MacKenzie, & Podsakoff, 2012). If method bias is not addressed then it could show up as a separate factor, which is what seems to have happened in the current study. There seems to be good evidence here to support the idea that method effect was significantly affecting the factor structure of the data set. Therefore the removal of the positively worded items significantly strengthened the factor structure and provides a more accurate understanding of the properties of the data set. However, it is unfortunate that standard psychometric procedures led to the exclusion of items therapist felt would be clinically useful.

Meaningful subscales. The multidimensional structure of the CAMOS allows for sub-scores to be created for each factor. This efficiently organizes and summarizes the data in meaningful ways. These subscales will provide more in-depth information about the client and provides a quick method for obtaining a greater level of information about the client. Instead of just receiving one total score about the client's overall functioning, it is possible with the CAMOS to obtain not only a total score, but also six sub scores that may help clinicians understand specific areas of life in which their clients may be struggling. This concept is supported by Lazarus's BASIC ID theory (1973; 2007) that argues that clinicians often focus too narrowly on psychological symptoms only. Having a valid and reliable multidimensional measure helps clinicians obtain relevant information for a variety of life areas. Richards and Bergin's multidimensional multilevel theory of assessment (2005) builds on this idea by taking into account that once a clinician is aware of potential concerns in one area of life through an outcome measure they can follow up with their client to a greater degree in their therapy session. Having that greater specificity of information can actually help the clinician tune in to certain areas of life that may be of greatest concern to the client. A total score gives only limited information in that the clinician is only made aware of the client's overall response to therapy but

does not give any specific direction as to where to focus their efforts. A multidimensional measure can directly influence treatment planning and treatment goals based on their scores in each dimension. This is helpful on the outset of therapy by helping the clinician to get an idea of what the client may be most concerned about. It is also helpful if it is being used on a session-by-session basis to gauge how the client may be improving or deteriorating in each area of life.

Spiritual Distress Dimension

It is of particular significance to discuss the implications of spirituality emerging as a distinct factor in the multidimensional structure. As was stated earlier, many outcome measures leave out cultural specific questions because it is assumed that if a client is experiencing distress about their culture, religion, or spiritual beliefs then that distress will be manifest on a measure of general psychological distress. However, our study provided evidence to support the idea that spirituality is a distinct area of life that registers as a separate factor empirically. This suggests that other outcome measures are missing significant aspects of client experience if they are neglecting to measure spirituality.

The explicit exclusion of religion and spirituality is nothing new in the field of psychology and is of course not localized only to the field of ROM. The field of ROM is simply on par with the bias in the field of psychology (Richards & Bergin, 2005; Smith et al., 2007). This bias is not just apparent to those in the field of psychology but is also apparent to those who access psychological services. Many people assume that religion or spirituality have to be "left at the door" so to speak. Including a spirituality component into the intake process invites clients from the beginning to know that religious or spiritual concerns can be talked about openly.

Brevity and Flexibility

It is important to highlight that the overall item pool was reduced significantly from 70 items to 42 items. This makes it suitable to be used on a session-by-session basis with clients as

it only takes five to seven minutes to complete. It could also be used simply as a brief intake and discharge measure to document change at the beginning and end of therapy. It is very rare to have such a brief measure cover so much breadth of content. Within a short period of time each week the therapist can get information from seven distinct areas of life and use that to guide their focus in therapy.

Another benefit that comes from this multidimensional structure is that dimensions can be used independently of the other dimensions. This is extremely helpful in shortening the measure if certain dimensions do not seem relevant to a particular client. The CAMOS was created on an Internet-based survey software that allows for quick and easy adjustments even to individual client surveys. An example of this may be that a particular client may not see spirituality as a relevant part of their experience, and therefore the dimension can easily be left out due to the multidimensional structure as well as the computer adaptive capabilities in the software. Another example is that the clinician may only want to their clients to take the psychological distress dimension to get only a total score that measures their overall psychological distress. A measure that is flexible in this way will help to bridge the research–practice gap. This is a measure that is psychometrically supported and yet can be altered to help clinicians respond to "client characteristics, culture, and preferences" (APA, 2006, p. 273) where the measure itself can be tailored. Research data in situations where certain dimensions are left out is of course, inherently limited, but in this system there is great possibility that clinicians will find this method far more practical and useful, which may help increase implementation rates, which has been identified as a major issue in routine outcome monitoring.

Recommendations for Clinical Practice

I have previously discussed the importance of measuring multiple dimensions in one measure and the clinical implications for such a measure. However, it is also important to

highlight the specific utility of each dimension and rationale for inclusion. The dimensions that emerged in the current study are consistent with others studies that have examined what outcomes are important to measure. In a study conducted by Slade (2002), 16 studies were examined that suggested what outcome domains are relevant and important to monitor. Of the emergent domains found in the final factor model of the CAMOS, five out of the six were represented among the emergent factors found in the study conducted by Slade. The only factor of the CAMOS that was not suggested by the Slade study was the spirituality dimension. The general exclusion of spirituality and religion has been discussed previously in this manuscript, and it was with intention that we included this dimension.

In another study done by Miovic and colleagues, a list of relevant and important topics of discussion in psychotherapy was created by a large group of experienced practitioners and researchers (2006). This list was then presented to clients and therapists who were asked to rate the relevance and importance of each item on the list. Findings indicated that while therapists generally rated relevance and importance higher than the clients did, both clients and therapists agreed on the importance of topics relative to the others. That is to say that therapists and clients came up with the same top ten list of important topics. Among the list of top ten important topics, seven are represented among the emergent factors found in the CAMOS. Among the topics discussed that were not covered in the CAMOS are finance, sexuality, and race/ethnicity. It is interesting to note that spiritual or religious beliefs and spiritual or religious activities and practices were ranked fifth and sixth in importance. This lends support again to the inclusion of a spirituality dimension specifically. This also seems to highlight the need for a spirituality measure if clients and therapists deem it important (Miovic et al., 2006), and yet this is something most outcome measures do not include (Slade, 2002). With that being said, these

studies lend support to the importance of the inclusion of the emergent factors found in the current study. In the following section discussion will be included specifically highlighting the importance and utility of each dimension.

The *psychological distress* dimension is measured by questions that focus on cognitive or emotional symptoms of depression or anxiety. This emerged as a strong dimension and, as was stated previously, a combination of the hypothesized emotional and cognitive dimension questions. The ongoing monitoring of these symptoms of distress is of primary importance in the field of psychology, and this principle of accurately identifying client distress is the focus of clinical training. Just as psychologists focus on identifying the proper DSM diagnosis based on clinical distress, the CAMOS also seeks to hone in on the nature and severity of that distress. The *psychological distress* dimension is the backbone of this assessment by providing a measure for the overall level of emotional and cognitive distress. This dimension has the most items (11) and had the highest Cronbach's alpha reliability coefficient at 0.92. This dimension correlated quite strongly with the total score of the OQ-45 ($r = .75$). The OQ-45 is one of the most widely used outcome measures in the nation, and the fact that such a high correlation exists with our 11 item subscale is impressive. Clinicians will be able to get a general sense of client distress by only asking clients to respond to these 11 items.

The *relationship distress* dimension was defined as measuring the distress around clients' interactions with others by tapping into feeling understood, actions or anger toward others, and others' actions toward them. Relationships are often the most important aspect of an individual's life, and interpersonal problems can be the cause of significant personal distress (Davies-Osterkamp, Strauss, & Schmitz, 1996; Horney, 1950). It is no surprise then that issues in relating to others are often the main reason why individuals seek psychotherapy (Horowitz, Rosenberg,

& Bartholomew, 1993). Theory and research also suggests that the presence of interpersonal problems can have an effect on the therapeutic alliance (Renner et al., 2012). There is some evidence to suggest that when the therapeutic alliance is given special attention, these negative effects can be mitigated (Hardy et al., 2001; Howard, Turner, Olkin, & Mohr, 2006). The *relationship distress* dimension can also help the therapist become aware of a poor social support system, in which case the therapist may choose to make that a focus. Many research studies have examined the benefits of a strong social support system (Bankoff, 1996; Bankoff & Howard, 1992; Lara, Leader, & Klein, 1997; Moras & Strupp, 1982), and the therapist may utilize a reported strong support system or seek to strengthen or broaden a weak social support system. Therefore having a *relationship distress* dimension is useful because it is addressing an area of life that is of significant importance, can help therapists assess and address interpersonal issues and understand how they may affect the therapeutic alliance, and help therapists understand the strengths or weaknesses of a client's support system. There is also application of ongoing monitoring of this dimension considering the dynamic nature of relationships.

The *therapy expectations* dimension was defined as measuring a client's anxiety or worry about his or her anticipated therapy experience. It is important to note that at intake the items were worded in such a way as to target client concerns or anxieties for beginning therapy (e.g., I felt anxious about beginning therapy), while in subsequent sessions questions were worded differently to target their current feelings about their therapy progress (e.g., I felt anxious about my therapy progress); the current study only analyzed data from intake sessions. Constantino, Glass, Arnkoff, Ametrano, and Smith researched extensively the importance of client expectations for therapy (2011). They gave several suggestions for addressing clients with low expectations for therapy, including but not limited to explicitly addressing their expectations

generally, hope inspiring statements (Pinel & Constantino, 2003), personalizing statements intended to enhance expectations, and describing successful research findings for the treatment being provided. This highlights just a few clinical implications that could be implemented by attending to a client's expectations for therapy through the *therapy expectations* dimension. They also go on to stress how important it is to "regularly check in on patients' outcome expectations" (Constantino et al., 2011, p. 370), which lends support to the inclusion of a therapy expectations/progress dimension in a routine outcome measure. This can help the therapist know if the process of therapy may need to be addressed directly, whether that be from the very first session or later on if the CAMOS is being used on a weekly basis.

The *spiritual distress* dimension assesses clients' feelings about their relationship with a higher power or acting inconsistently with their spiritual or religious beliefs. This aspect of life can be so diverse and unique to each individual, and it may be a source of help and strength to some clients while others may see it as a source of guilt and shame. If significant distress is indicated in this area it may be appropriate to refer them to their religious leader. In this way the therapist could draw upon this potential support system. However, if the client is experiencing significant distress due to doubts about their religion they may not feel comfortable discussing these concerns with their religious leaders (Richards & Bergin, 2005). Additionally, religious concerns are often intertwined with personal problems (Bergin, Stinchfield, Gaskin, Masters, & Sullivan, 1988; Richards, Smith, & Davis, 1989), and some religious leaders may be ill equipped to provide expertise concerning the complex interplay of psychological concerns and religious concerns. A culturally sensitive therapist can provide the environment where a client can explore those concerns, and monitoring outcomes specific to that domain is helpful.

The *physical health distress* dimension was simply defined as measuring clients' physical symptoms or their level of concern with physical symptoms. There is a large amount of literature that discusses the mind-body connection, positing that physical symptoms can affect psychological symptoms and vise-versa. Practitioners are better able to address client problems holistically when they are aware of clients' physical symptoms. Much research has been done on the detrimental effects that come from fragmenting mental health and physical health (Lurie, Manheim, & Dunlop, 2009) as well as the benefits of integrated care (Kwan & Nease, 2013; Miller, Kessler, Peek, & Kallenberg, 2011). By including a physical health dimension, continuity of care is increased and psychologists are able to become more aware of physical concerns and provide a more holistic approach to treatment. This awareness may facilitate better and more accurate interventions targeting multiple symptoms (mental and physical) and may also improve collaboration with primary care physicians.

The *work/school distress* dimension was designed to measure social role or social responsibility by mainly asking questions targeting work or school but that were intended to apply to other responsibilities such as volunteer work or homemaking. There is an apparent split between career counseling and personal counseling; however, there are those who argue that this is a false dichotomy and that there are great benefits from integrating personal and career counseling (Betz & Corning, 1993; McIlveen, 2015; Manuele-Adkins, 1992; Super, 1993). Robitschek and DeBell suggest that when issues in vocation and career exist they should be considered a primary contextual factor in counseling (2002). Therefore they suggest that career functioning and importance be assessed with every client regardless of their presenting concern. Juntunen also suggests that intake forms should include questions related to work satisfaction and work concerns (2006). Using assessment and discussion of career can facilitate exploration

and understanding of life roles, family roles and responsibilities, self-awareness/self-knowledge, possible selves, and patterns of dealing with stress (Juntunen, 2006). Assessment and monitoring of career/educational concerns facilitates the interventions suggested by these researchers and is in line with several other researchers who extrapolate the importance of integrating career concerns into psychotherapy (Richardson, 2009; Swanson, 2012).

Limitations

One limitation of the present study was the lack of diversity among the client population. There is a need to cross-validate the factor analysis results of the present study with additional client populations that are more religiously, racially, and culturally diverse.

Another limitation was with the CFC sample and the work/school dimension. This dimension was primarily designed to assess work or school concerns, and since the Center for Change is a residential treatment facility, none of the women there were currently employed and only a portion of them were planning on attending school at the center. Even those who were planning on attending school at the center will likely not have started school there yet since the intake administrations happen within the first few days. While the dimension was primarily intended to measure work/school concerns, we did try to word the questions broadly and we even specified to clients that these questions could apply to "volunteer work and homemaking." However, it is interesting to consider how being admitted into a residential treatment facility impacts their view of their social responsibilities. It is important to note however that even with these limitations a distinct factor emerged. This lends credibility to the idea that this is a useful dimension with some valuable questions even if it was one of the weaker dimensions within this data set.

It has been mentioned previously that this measure can be used as a stand-alone outcome measure that could be used on a session-by-session basis. It is important to note, however, that

the main purpose of this study is to understand the properties of this data set and its factor structure. Issues concerning the sensitivity to change have not yet been addressed.

Strengths

Despite the limitations, this study had a number of strengths. First, our samples were drawn from real psychotherapy clients receiving treatment in actual treatment settings. Second, the samples sizes in our two data sets were large and adequate for the factor analytic and item analysis procedures that we used. Third, we used state-of-the-art statistical procedures and decision-making criteria in our use of both exploratory and confirmatory factor analyses. Fourth, although the Idaho sample had limited religious diversity (in terms of affiliation), it did have considerable diversity in regards to client presenting concerns and psychological problems. Although the Utah sample was limited to female eating disorder patients, these patients did have a variety of other co-occurring disorders (e.g., depression, anxiety, substance abuse, personality disorders) and it did provide us with the opportunity to study the CAMOS in an inpatient sample with severe psychiatric disturbances.

Future Research Directions

The findings from this study were originally conceptualized as a first step in a larger project, and while the results of the current study have unique and independent significance, there are many implications for further research. Now that there is support for the multidimensional structure, there is a need for studies that further address the issue of the length of the CAMOS. Although the version of the CAMOS created from the present study is reasonably brief, it would be desirable to create an even shorter form of the CAMOS. This would shorten the length of time even further that clients spend taking the CAMOS on a repeated measures basis. An even shorter form of the CAMOS would also leave more room for additional dimensions that could be added based on the needs of individual treatment sites. In addition to

the standardized dimensions that are applicable across all clients, there will also be other optional dimensions and items that can be individualized to the client. This highlights a crucial benefit of the CAMOS where both assessment approaches (abstractionism and contextualism) are integrated into one assessment and the reporting platform makes it quick and easy to take for the client and to view the results for the therapist. This hybrid approach allows clinicians to be able to more aptly integrate and embody the three decision-making factors outlined by the task force on EBPP. The "best available research" is being represented in the six core dimensions, which will be validated and standardized. Clinical expertise and decision making is enhanced by providing the clinician with a platform to collect information pertaining to specific patient's "characteristics, culture, and preferences" (APA, 2006, p. 273) through the use of a population specific dimension and individual dimension. At the same time we will be addressing clinicians' concerns with local validity, thereby increasing the likelihood of ROM implementation. This will all be provided through an integrated system that effectively combines breadth of assessment with brevity, and abstractionism with contextualism. The task of shortening the standardized portion has been undertaken by Sanders (2015) and his findings are reported in his doctoral dissertation.

An example of this hybrid system is currently underway. At BYU–Hawaii the current model of the CAMOS is being administered at intake along with two new dimensions that were developed in collaboration with BYU–Hawaii Counseling Center and the University of Hawaii Counseling Center. These new dimensions focus on client perception of their therapist's cultural sensitivity in therapy and cultural concerns. At another site, in collaboration with the founders of the CFC, we developed a dimension that specifically assesses concerns with eating disorders. These dimensions may remain at these respective sites only or they may be further analyzed and

developed for use with similar populations at different locations. Due to the multidimensional nature of the CAMOS and the capabilities and flexibility of the survey software it is relatively easy to add dimensions based on relevance to the client population. The results of the current study have established a foundation for standardization while being part of a system that can be expanded based on the needs of the particular site and population.

Given the limitation with the *work/school distress* dimension, it may be useful to separate out the two and, instead of trying to capture social role, perhaps create specific dimensions that measure school distress and work distress separately. This is made possible due to the multidimensional nature of the CAMOS, where different dimensions could be used if they were validated empirically. The Counseling Center Assessment of Psychological Symptoms (CCAPS) is a multidimensional outcome measure that includes a distinct school distress dimension (Locke et al., 2011). This may lend support to the idea that distress concerning school should be measured separately from work or social role concerns. Assessing just one of these domains (i.e., school, work) of course makes it more narrowly applicable. However, with the CAMOS it may be a possibility to create separate work and school dimensions that could be used only when a client is attending school or currently employed. Then the measure could apply to student populations as well as employed populations.

It was unfortunate that the standard psychometric procedures in the current analyses led to the exclusion of all the positively worded items, especially considering the fact that those items were created and included based on the request of the practitioners. Apart from the statistical rationales for inclusion or exclusion of positively worded items, there are also theoretical and practical rationales for their inclusion—one reason being that the absence of negative symptoms does not necessarily imply the presence of positive aspects of life.

Additionally, as these items are meant to apply to a wide variety of therapy orientations, a purely negatively worded survey would go against a therapist practicing positive psychology. Therefore, it would be beneficial to establish future renditions of the CAMOS that were able to include positively worded items while still maintaining strong psychometric properties.

The multidimensional structure of the CAMOS combined with the emergent spirituality factor show support for the idea of creating spirituality dimensions that are denomination specific or non-theistic. The current dimension was designed to be as ecumenical as possible, assuming some form of relationship to God or a higher power. Naturally in making an ecumenical spirituality dimension there is still a great deal of variance that could occur given that clients from different religious and spiritual backgrounds may interpret the questions differently. This spiritual dimension seemed to be interpreted consistently enough as to produce a viable factor in the model. However, this may be due in part to the generally homogeneous religious affiliation reported by the subjects in the samples. More research needs to be done with more religiously diverse populations to answer the question of differences of interpretation among clients of different religions.

The results of this study showed that the CAMOS items about emotions and thoughts were not significantly different from each other to warrant separate dimensions. However, it is possible that the wording of the CAMOS questions overlapped significantly to the degree that the distinction between the two were lost. For example, a question that was designed originally for the thoughts dimension began with the phrase "I felt . . ." which is getting perhaps more at a feeling than a belief or cognition. Additionally, one weakness in the question wording may have been the incorporation of the phrase "I felt sad . . ." Although it was referring to feeling sad about their actions, it may have blurred the lines between a client's appraisal of their behavior

and their feelings about their behavior. It may be possible to more carefully separate cognitions and emotions through clearer question wording, which may in future studies show that there is a significant distinction between the two.

Conclusion

The findings of the present study have laid the foundation for the development of a valuable measure than can be used in the psychotherapy field for assessing the outcomes of treatment. As recommended by the APA task force on evidence-based practice (APA, 2006), the CAMOS provides researchers and therapists with a brief, clinically relevant and adaptable outcome measure that can be used for routine monitoring of client progress. The CAMOS has numerous features that will make it valuable for psychotherapy researchers and practitioners, particularly for those who wish to use practice-based evidence research designs that are sensitive to the needs and challenges of patients and practitioners. The current version of the CAMOS serves as a reliable and valid base that can be added upon in many ways, affording therapists and treatment sites to customize the measure to their needs and research questions. Due to the multidimensional nature of the CAMOS, dimensions are easily added or removed. This format will facilitate a bottom-up approach, which may increase therapist buy-in and usage rates, thus attempting to shrink the research–practice gap. The inclusion of a spirituality dimension in a routine outcome measure is also a significant accomplishment. This provides greater opportunity for spiritual concerns to be assessed in routine practice and may help increase data collection and research in the area of spirituality. All these features and capabilities will allow the CAMOS to not only serve as a valid and reliable outcome measure, but to also serve as a base for further adaptability and research—not just to provide solid statistical information, but to provide opportunities for customized and flexible research that can adapt to changes in technology and practice.

References

Abrahamson, D. J. (1999). Outcomes, guidelines, and manuals: On leading horses to water. *Clinical Psychology: Science and Practice, 6,* 467–471.

Addis, M. E. (2002). Methods for disseminating research products and increased evidence-based practice: Promises obstacles, and future directions. *Clinical Psychology: Science and Practice, 9,* 367–378.

Addis, M. E., & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes towards psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology, 65*, 331–339.

American Psychological Association. (2006). Evidence-based practice in psychology: APA presidential task force on evidence-based practice. *American Psychologist*, *61*, 271–285.

American Psychological Association Division of Clinical Psychology. (1995). Training in and dissemination of empirically validation psychological treatments: Report and recommendations. *The Clinical Psychologist, 48,* 3–27.

Asay, T. P., Lambert, M. J., Gregersen, A. T., & Goates, M. K. (2002). Using patient-focused research in evaluating treatment outcome in private practice. *Journal of Clinical Psychology, 58*(20), 1213–1225.

Bankoff, E. A. (1996). Pre-treatment social support and effective psychotherapeutic process: A panel study. *Psychotherapy, 33,* 51–60.

Bankoff, E. A., & Howard, K. I. (1992). The social network of the psychotherapy patient and effective psychotherapeutic process. *Journal of Psychotherapy Integration, 2*, 273–294.

Barkham, M., Hardy, G. E., & Mellor-Clark, J. (2010). *Developing and delivering practice-based evidence: A guide for the psychological therapies*. Chichester, West Sussex, UK; Malden, MA: Wiley-Blackwell. doi:10.1002/9780470687994

Bergin, A. E., Stinchfield, R. D., Gaskin, T. A., Masters, K. S., & Sullivan, C. E. (1988). Religious life-styles and mental health: An exploratory study. *Journal of Counseling Psychology*, *35*(1), 91–98. doi:10.1037/0022-0167.35.1.91

Betz, N. E., & Corning, A. F. (1993). The inseparability of "career" and "personal" counseling. *Career Development Quarterly, 42*, 137–142. doi:10.1002/j.2161-0045.1993.tb00426.x

Beutler, L. E., Williams, R. E., Wakefield, P. J., & Entwistle, S. R. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist*, *50*, 984–994. doi:10.1037/0003-066X.50.12.984

Bewick, B. M., Trusler, K., Mullin, T., Grant, S., & Mothersole, G. (2006). Routine outcome measurement completion rates in primary care psychological therapies and counselling. *Counselling and Psychotherapy Research, 6,* 33–40.

Bickman, L., Rosof-Williams, J., Salzer, M. S., Summerfelt, W. T., Noser, K., Wilson, S. J., & Karver, M. S. (2000). What information do clinicians value for monitoring adolescent client progress and outcomes? *Professional Psychology: Research and Practice*, *31*, 70–74. doi:10.1037/0735-7028.31.1.70

Brand, N. (2008). Psychologists' approaches to evaluating client progress in psychotherapy practice. *Dissertation Abstracts International, 68,* 5558.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York, NY: Routledge/Taylor & Francis Group.

Callaly, T., & Hallebone, E. (2001). Introducing the routine use of outcomes measurement to mental health services. *Australian Health Review 24*, 43–50.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105. doi:10.1037/h0046016

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 140–161.

Cohen, L. H., Sargent, M. M., & Sechrest, L. B. (1986). Use of psychotherapy research by professional psychologists. *American Psychologist*, *41*, 198–206. doi:10.1037/0003-066X.41.2.198

Constantino, M. J., Glass, C. R., Arnkoff, D. B., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. In J. C. Norcross, & J. C. Norcross (Eds.), *Psychotherapy relationships that work: Evidence-based responsiveness* (2nd ed., pp. 354–376). doi:10.1093/acprof:oso/9780199737208.003.0018

Cook, J. M., Schnurr, P. P., Biyanova, T., & Coyne, J. C. (2009). Apples don't fall far from the tree: Influences on psychotherapists' adoption and sustained use of new therapies. *Psychiatric Services*, *60*, 671–676. doi:10.1176/appi.ps.60.5.671

Davies-Osterkamp, S., Strauss, B. M., & Schmitz, N. (1996). Interpersonal problems as predictors of symptom related treatment outcome in long-term psychotherapy. *Psychotherapy Research*, *6*, 164–176. doi:10.1080/10503309612331331688

Drabick, D. G., & Goldfried, M. R. (2000). Training the scientist–practitioner for the 21st century: Putting the bloom back on the rose. *Journal of Clinical Psychology*, *56*, 327–340. doi:10.1002/(SICI)1097-4679(200003)56:3<327::AID-JCLP9>3.0.CO;2-Y

Garland, A. F., Kruse, M., & Aarons, G. A. (2003). Clinicians and outcome measurement. What's the use? *Journal of Behavioral Health Services & Research, 30*, 393–405.

Gerdes, K., Edmonds, R., Haslam, D., & McCartney, T. (1996). A statewide survey of licensed clinical social workers' use of practice evaluation procedures. *Research on Social Work Practice, 6,* 27–39.

Gilbody, S. M., House, A. O., & Sheldon, T. A. (2002). Psychiatrists in the UK do not use outcome measures. *British Journal of Psychiatry, 180,* 101–103.

Goldfried, M. R., & Wolfe, B. E. (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology, 66*, 143–150.

Hardy, G. E., Cahill, J., Shapiro, D. A., Barkham, M., Rees, A., & Macaskill, N. (2001). Client interpersonal and cognitive styles as predictors of response to time-limited cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, *69*, 841–845. doi:10.1037/0022-006X.69.5.841

Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice, 35,* 485–491. doi:10.1037/0735-7028.35.5.485

Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K., & Tuttle, K. (2004). The effects of providing patient progress information to therapists and patients. *Psychotherapy Research, 31*, 308–327.

Haynes, S. N., Mumma, G. H., & Pinson, C. (2009). Idiographic assessment: Conceptual and psychometric foundations of individualized behavioral assessment. *Clinical Psychology Review*, *29*, 179–191. doi:10.1016/j.cpr.2008.12.003

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*, 191–205. doi:10.1177/1094428104263675

Horn, J. L. (1965). A rationale and a test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185.

Horney, K. (1950). *Neurosis and human growth: The struggle toward self-realization*. Oxford, England: W. W. Norton.

Horowitz, L. M., Rosenberg, S. E., & Bartholomew, K. (1993). Interpersonal problems, attachment styles, and outcome in brief dynamic psychotherapy. *Journal of Consulting and Clinical Psychology*, *61,* 549–560.

Howard, I., Turner, R., Olkin, R., & Mohr, D. C. (2006). Therapeutic alliance mediates the relationship between interpersonal problems and depression outcome in a cohort of multiple sclerosis patients. *Journal of Clinical Psychology*, *62*, 1197–1204. doi:10.1002/jclp.20274

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrics, 23*, 187–200.

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, *63*, 146–159. doi:10.1037/0003-066X.63.3.146

Kettlewell, P. W. (2004). Development, dissemination, implementation of evidenced-based treatments: Commentary. *Clinical Psychology: Science and Practice, 11,* 190–195.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

Kwan, B. M., & Nease, D. (2013). The state of the evidence for integrated behavioral health. In M. R. Talen, & A. B. Valeras (Eds.), *Essentials of integrated care: Connecting systems of care, clinical practice and evidence-based approaches* (pp. 65–98). New York, NY: Springer Science Business Media.

Lakeman, R. (2004). Standardized routine outcome measurement: Pot holes in the road to recovery. *International Journal of Mental Health Nursing, 13,* 210–215.

Lambert, M. J., Bergin, A. E., & Garfield, S. L. (2004). Introduction and historical overview. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.). New York, NY: Wiley.

Lambert, M. J., Bergin, A. E., & Garfield, S. L. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.). New York, NY: Wiley.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 70*, 288–301.

Lara, M. E., Leader, J., & Klein, D. N. (1997). The association between social support and course of depression: Is it confounded with personality? *Journal of Abnormal Psychology, 106,* 478–482.

Lazarus, A. A. (1973). Multimodal behavior therapy: Treating the BASIC ID. *Journal of Nervous and Mental Diseases, 156*, 404–411.

Lazarus, A. A. (2007). Multimodal therapy. In R. J. Corsini, & D. Wedding (Eds.), *Current psychotherapies* (8th ed., pp. 337–371). Belmont, CA: Brooks Cole.

Leon, S. C., Kopta, S. M., Howard, K. I., & Lutz, W. (1999). Predicting patients' responses to psychotherapy: Are some more predictable than others? *Journal of Consulting and Clinical Psychology, 67*, 698–704.

Locke, B. D., Buzolitz, J. S., Lei, P., Boswell, J. F., McAleavey, A. A., Sevig, T. D., . . . Hayes, J. A. (2011). Development of the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62). *Journal of Counseling Psychology*, *58*, 97–109. doi:10.1037/a0021282

Lucock, M., Hall, P., & Noble, R. (2006). A survey of influences on the practice of psychotherapists and clinical psychologists in training in the UK. *Clinical Psychology & Psychotherapy*, *13*, 123–130. doi:10.1002/cpp.483

Lueger, R. J., Howard, K. I., Martinovich, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assessing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology, 69*, 150–158.

Lurie, I. Z., Manheim, L. M., & Dunlop, D. D. (2009). Differences in medical care expenditures for adults with depression compared to adults with major chronic conditions. *Journal of Mental Health Policy and Economics, 12*, 87–95.

Manuele-Adkins, C. (1992). Career counseling is personal counseling. *Career Development Quarterly, 40*, 313–323. doi:10.1002/j.2161-0045.1992.tb00338.x

McIlveen, P. (2015). Psychotherapy, counseling, and career counseling. In P. J. Hartung, M. L. Savickas, W. B. Walsh, (Eds.), *APA handbook of career intervention, Volume 1: Foundations* (pp. 403–417). doi:10.1037/14438-022

Miller, B. F., Kessler, R., Peek, C. J., & Kallenberg, G. A. (2011). *A national research agenda for research in collaborative care: Papers from the collaborative care research network research development conference* (AHRQ Publication No. 11–0067). Rockville, MD: Agency for Healthcare Research and Quality.

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The Partners for Change Outcome Management System. *Journal of Clinical Psychology, 61*, 199–208.

Miovic, M., McCarthy, M., Badaracco, M. A., Greenberg, W., Fitzmaurice, G. M., & Peteet, J. R. (2006). Domains of discussion in psychotherapy: What do patients really want? *American Journal of Psychotherapy*, *60*, 71–86.

Moras, K., & Strupp, H. (1982). Pretherapy interpersonal relations, patients' alliance, and outcome in brief therapy. *Archives of General Psychiatry, 39*, 405–409.

Morrow-Bradley, C., & Elliott, R. (1986). Utilization of psychotherapy research by practicing psychotherapists. *American Psychologist*, *41*, 188–197. doi:10.1037/0003-066X.41.2.188

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments & Computers*, *32*, 396–402. doi:10.3758/BF03200807

Ogilvie, A. J. (2012). The customer is always right: Bridging the research-practice gap by identifying consumer-friendly research methods. *Dissertation Abstracts International*, *72,* 4991.

Patterson, P., Matthey, S., & Baker, M. (2006). Using mental health outcome measures in everyday clinical practice. *Australasion Psychiatry, 14*, 133–136.

Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66,* 126–135.

Pinel, E. C., & Constantino, M. J. (2003). Putting self-psychology to good use: When social psychologists and clinical psychologists unite. *Journal of Psychotherapy Integration, 13*, 9–32.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*539-569. doi:10.1146/annurev-psych-120710-100452

Raykov, T., & Marcoulides, G. A. (2000). *A first course in structural equation modeling.* Mahwah, NJ: Erlbaum.

Renner, F., Jarrett, R. B., Vittengl, J. R., Barrett, M. S., Clark, L. A., & Thase, M. E. (2012). Interpersonal problems as predictors of therapeutic alliance and symptom improvement in cognitive therapy for depression. *Journal of Affective Disorders*, *138*, 458–467. doi:10.1016/j.jad.2011.12.044

Robitschek, C., & DeBell, C. (2002). The reintegration of vocational psychology and counseling psychology: Training issues for a paradigm shift. *The Counseling Psychologist, 30*, 801–814. doi:10.1177/00110000 2237755

Roth, A., & Fonagy, P. (2005). *What works for whom: A critical review of psychotherapy research* (2nd ed.). New York, NY: Guilford Publications.

Richards, P. S., & Bergin, A. E. (2005). *A spiritual strategy for counseling and psychotherapy* (2nd ed.). Washington, DC: American Psychological Association. doi:10.1037/11214-000

Richards, P. S., Smith, S. A., & Davis, L. F. (1989). Healthy and unhealthy forms of religiousness manifested by psychotherapy clients: An empirical investigation. *Journal of Research in Personality*, *23*, 506–524. doi:10.1016/0092-6566(89)90017-2

Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods*, *12*, 762–800. doi:10.1177/1094428109332834

Richardson, M. S. (2009). Another way to think about the work we do: Counselling for work and relationship. *International Journal for Educational and Vocational Guidance, 9,* 75–84. doi:10.1007/s10775-009-9154-3

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29*, 304–321. doi:10.1177/0734282911406653

Slade, K., Lambert, M. J., Harmon, S. C., Smart D. W, & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Journal of Clinical Psychology and Psychotherapy, 15*, 287–303.

Slade, M. (2002). What outcomes to measure in routine mental health services, and how to assess them: A systematic review. *Australian and New Zealand Journal of Psychiatry*, *36*, 743–753. doi:10.1046/j.1440-1614.2002.01099.x

Slife, B. D. (2004). Theoretical challenges to therapy practice and research: The constraint of naturalism. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 44–83). New York, NY: Wiley.

Smith, T. B., Bartz, J., & Richards, P. S. (2007). Outcomes of religious and spiritual adaptations to psychotherapy: A meta-analytic review. *Psychotherapy Research*, *17*, 643–655. doi:10.1080/10503300701250347

Sonnanburg, K. (1996). Meaningful measurement in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, *33*, 160–170. doi:10.1037/0033-3204.33.2.160

Stewart, R. E., & Chambless, D. L. (2007). Does psychotherapy research inform treatment decisions in private practice? *Journal of Clinical Psychology 63*, 267-281. doi:10.1002/jclp.20347.

Streiner, D. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80*, 99–103. doi:10.1207/S15327752JPA8001_18

Super, D. E. (1993). The two faces of counseling: Or is it three? *Career Development Quarterly, 42,* 132–136. doi:10.1002/j.2161-0045.1993.tb00425.x

Swanson, J. L. (2012). Work and psychological health. In N. A. Fouad, J. A. Carter, L. M. Subich, N. A. Fouad, J. A. Carter, & L. M. Subich (Eds.), *APA handbook of counseling psychology: Practice, interventions, and applications* (Vol. 2, pp. 3–27). doi:10.1037/13755-001

Trauer, T. (Ed.). (2010). *Outcome measurement in mental health: Theory and practice.* Retrieved from http://www.ebrary.com

Ventimiglia, J. A., Marschke, J., Carmichael, P., & Loew, R. (2000). How do clinicians evaluate their practice effectiveness? A survey of clinical social workers. *Smith College of Studies in Social Work, 70*, 287–306.

Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment failure and problem solving strategies in routine practice. *Journal of Counseling Psychology, 58,* 59–68.

Williams, L. J., Hartman, N., & Cavazotte, F. (2010). Method variance and marker variables: A review and comprehensive CFA marker technique. *Organizational Research Methods*, *13*, 477–514. doi:10.1177/1094428110366036

Zimmerman, M., & McGilchey, J. B. (2008). Why don't psychiatrists use scales to measure outcome when treating depressed patients? *Journal of Clinical Psychiatry, 69,* 1917–1919.

APPENDIX A: REVIEW OF THE LITERATURE

During the past 50 years, a great deal of research has been done in the field of psychotherapy (Lambert, Bergin, & Garfield, 2004). This research has provided much insight into the processes and outcomes of psychotherapy. It has also provided support for the effectiveness of psychotherapy in general, and for many specific types of psychotherapy, such as behavioral, cognitive, humanistic, family, and group psychotherapy approaches (Lambert, Bergin, & Garfield, 2013).

Despite the large amount of research supporting the effectiveness of various types of psychotherapy, relatively little is known about what types of psychotherapy approaches are most effective for what types of clients and what types of clinical issues (Roth & Fonagy, 2005). The lack of specificity in the knowledge base about what approaches are most effective for specific types of clients and clinical issues has contributed to the infamous research–practice gap in the field of psychotherapy (Drabick & Goldfried, 2000; Kazdin, 2008; Ogilvie, 2012). The research–practice gap refers in part to the fact that empirical research has little influence on most practitioners. Treatment interventions that practitioners select and use with clients are based more on theoretical rationales and clinical experience than on the findings of psychotherapy research (Beutler, Williams, Wakefield, & Entwistle, 1995; Cohen, Sargent, & Sechrest, 1986; Cook, Schnurr, Biyanova, & Coyne, 2009; Lucock, Hall, & Noble, 2006; Morrow-Bradley & Elliot, 1986; Stewart & Chambless, 2007). The research–practice gap also refers to the lack of communication between researchers and practitioners. Researchers often fail to include practitioners in the process of designing their studies, which leads to research questions, methods, and findings that practitioners perceive as irrelevant to practice-related needs.

In response to concerns about the research–practice gap, and because of frustration that research findings were having so little influence on psychotherapy practice, an empirically supported treatment movement began within Division 12 of the American Psychological Association during the 1990s. Division 12 created a task force that developed a list of psychotherapy approaches that they considered "empirically supported treatments" (APA, 1995). The list of empirically supported treatments, and the criteria used by the task force to select these treatments, bred substantial controversy and criticism (Addis, 2002; Addis & Krasnow, 2000; Goldfried & Wolfe, 1998; Kettlewell, 2004; Persons & Silberschatz, 1998).

Seeing the need for revision of how best practice was defined in psychology, the APA assembled a presidential task force that established a definition of for evidence-based practice in psychology (EBPP) that focused on principles rather than specific treatments. This report emphasized the importance of using multiple methods for building an evidence base. This is a substantial shift from ESTs where treatments could only be validated by the use of rigorous RCTs. There is no doubt that RCTs play an important role in establishing best practice, but there are certainly other methods of establishing evidence for effective treatment and outcome. Specifically they mention clinical observation, qualitative research, systematic case studies, single-case experimental designs, public health and ethnographic research, process/outcome studies, effectiveness research, and meta-analyses (APA, 2006). Clearly the APA's current stance is that a wide variety of research methods should be used in establishing a robust and diverse evidence base in psychology. It is interesting to note that among the list of research designs that the task force article highlighted, most fall under the category of practice-based evidence (PBE; e.g., effectiveness research, public health and ethnographic research), and most

of these designs require repeated measurement during treatment (e.g., process/outcome research, single case experimental designs).

Practice-Based Evidence

Practice-based evidence is one form of research that is significantly different from the traditional RCT methodology, but in many ways it is complementary (Barkham, Hardy, & Mellor-Clark, 2010). One of the main differences from RCTs is that practice-based evidence focuses on collecting data in a naturalistic setting, which significantly increases its generalizability which is a significant limitation of RCTs. Conducting research in a community setting helps to show that treatment works in a real life setting in the presence of uncontrolled variables. While this makes it harder to make causal conclusions or make direct links to what variable influenced the outcome most strongly, it does make the results more practical and applicable.

In PBE studies it is important that therapists continue their treatment as usual and that they do not use a treatment manual approach. This allows therapists to use theories and interventions that they are most comfortable with and which they would most likely end up using in their work anyway (Beutler et al., 1995; Cohen et al., 1986; Cook et al., 2009; Lucock et al., 2006; Morrow-Bradley & Elliot, 1986; Stewart & Chambless, 2007). Additionally, it is not necessary to monitor the implementation of that treatment to ensure that it meets a certain criteria or protocol. Not only do these guidelines make results more generalizable and applicable, but they also make it more feasible for studies to be implemented more often and for longer duration. There is no need for lengthy training or treatment fidelity assurances that can be time consuming and costly. Therapists are free to choose a variety of treatment approaches and are not limited to a single approach.

Another feature of PBE is that all subjects presenting for treatment are included in the study, which creates heterogeneity with regard to personal characteristics as well as presenting problems. This not only allows for quicker data collection, but also directly results in greater applicability of findings. A drawback to RCTs is that they often only apply to a specific population and/or a specific disorder, which means that the obtained findings are severely limited and are only applicable to clients from that identified population or that have that targeted disorder. Findings from PBE research can be applied more generally across populations and presenting concerns because the samples reflect that wide range of subjects and problems.

Ownership by practitioners is an important feature of PBE. Practitioners and site managers are the driving force behind the research questions being addressed, and they are the ones that have ownership of the data. This is in stark contrast to RCTs, where researchers designate every element of the study. This idea also applies specifically to the research questions and designs of the study. Practitioners help create the research questions beings asked and the method for obtaining the proper data to answer those questions.

While many of the rigorous criteria of RCTs have been relaxed, it is important to note that there is still a strong emphasis on using reliable and valid measures. However, their application takes a more practical focus, accommodating for the site and existing procedures, rather than focusing on experimental control. Therefore, reliable and valid routine outcome measures are essential to most types of PBE designs. It is through these measures that data is collected in the naturalistic setting from every client.

Routine Outcome Measurement

The APA presidential task force defined evidence-based practice in psychology as “the integration of the best available research with clinical expertise in the context of patient

characteristics, culture, and preferences" (APA, 2006, p. 273). With the current definition of EBPP the burden falls on practitioners to show they are incorporating research with clinical expertise while taking into account specific client characteristics. The report from the task force also stated that "ongoing monitoring of patient progress and adjustment of treatment as needed are essential to EBPP" (APA, 2006, p. 280). In order to monitor the progress of patients, it is essential to engage in routine outcome measurement (ROM). ROM simply means that practitioners and researchers collect quantitative and qualitative measurements of patient processes and outcomes during the course of treatment.

The APA taskforce statement specifically includes recommendations that there is a need for "developing well-normed measures that clinicians can use to quantify their diagnostic judgments, measure therapeutic progress over time, and assess the therapeutic process" (APA, 2006, p. 278). ROM can been seen as trying to address the three major decision-making factors outlined by the EBPP task force with varying degrees of success. ROMs incorporate research as far as the ROM itself is research based, reliable, and valid. Analyzing the aggregate results of the ROM used perpetuates the research process as well. ROM can also enhance clinical judgment by providing the clinician with information about patient progress, allowing them increased responsiveness to potential problems. In this way the ROM allows the therapists access to up-to-date client characteristics that can inform the therapeutic process.

Routine treatment monitoring is a necessary part of establishing effective quality treatments. There have been many systems created to routinely assess client progress, but most of these systems share common theoretical and statistical methods for documenting client improvement. Some of these methods include a concept known as the reliable change index (RCI; Jacobson & Truax, 1991) and expected treatment response (ETR; Howard, Moras, Brill,

Martinovich, & Lutz, 1996) or benchmarking. These methods are used as a foundation for establishing estimates of client improvement in the COMPASS Treatment Assessment System (Howard, Brill, Lueger, & O'Mahoney, 1992), Outcome Questionnaire-45 (OQ-45; Lambert, Burlingame et al., 1996), Treatment Outcome Package (TOP; Kraus, Seligman, & Jordan, 2005), and the Partners for Change Outcome Management System (PCOMS; Miller, Duncan, Sorrell, & Brown, 2005) to name just a few of the more prominent outcome measures in the field. These commonly used methods for establishing client change will be discussed.

Reliable change index. Jacobson and Truax established a statistical approach to clinical significance (1991). They first discuss a method for identifying clinically relevant change within an individual client. This method hinges on the idea that a when a client enters therapy their overall functioning is inhibited, and therapy is intended to return them to normal functioning. They propose three methods for operationalizing when a client has moved from a dysfunctional population to a functional population. These include, (a) when a client's score on an outcome measure has moved two standard deviations from the mean of the dysfunctional population, (b) when the client's score enters two standard deviations below the mean of the functional population, and (c) when the client's score is closer to the functional population than the dysfunctional population. The authors suggest that when these two distributions are overlapping and normal distribution is assumed, then the third option is the most desirable, because this takes into account the variance of each distribution. This is a very practical model that uses the client's final score to categorize them into functional or dysfunctional populations and thereby establishes clinically relevant change.

This then raises the question, "How do we know if that movement from dysfunctional to functional is significant or reliable?" It just might be the case that the client's score's already

falls very near the midpoint between the dysfunctional and dysfunctional populations and any improvement would put their score closer to the functional population, thereby meeting the criteria for change, even though their improvement may just be due to chance or measurement error. Jacobson and Truax (1991) account for this problem by using statistical techniques to establish a reliable change index. The RCI is equal to difference between pre- and post-test scores divided by the standard error of the difference between the two test scores. The standard error of the difference "describes the spread of the distribution of change scores that would be expected if no actual change had occurred" (Jacobson & Truax, 1991, p. 637). Therefore, an RCI larger than 1.96 would occur less than 5% of the time without actual change taking place. This approach has been used in traditional pre- and post-outcome studies as well as studies where outcomes are measured routinely.

Expected treatment response or benchmarking. Another commonly used method in identifying client change involves comparing an individual client to a population of clients that have experienced treatment and experienced some form of improvement. Basically, this answers the question, "Is this client responding to treatment in the expected pattern of improvement?" (Trauer, 2010, p. 6). To implement this method for establishing change, the outcome measure used would necessarily need normative data to establish a typical response pattern for clients (Lueger et al., 2001). Individual clients' response patterns can then be compared to the normed sample and conclusions can be made on the individual client's progress.

This simple method has been improved upon in several ways. Lambert, Hansen, and colleagues (1996) rely heavily upon the client intake scores to assign recovery curves to clients. The creators of the COMPASS take into account client intake score as well as expected change per session (Howard et al., 1996; Lutz, Martinovich, & Howard, 1999). Other methods of

calculating treatment response also include identifying characteristics at intake that significantly affect rate of change, including, diagnoses, previous treatment, age, race, and gender (Howard et al., 1996; Kraus et al., 2005; Lutz, Martinovich, & Howard, 1999).

Benefits of ROM. These methods provide several benefits to clinical practice. ROM can help establish the effectiveness of treatment, document progress, and serve as an aid in decision making for therapists. Howard and colleagues (1996) introduced a patient-focused method for predicting outcome that uses the ETR method. Much of their research uncovered the complexity of attempting to measure client outcomes, especially tracking those who do not respond favorably to treatment (Leon, Kopta, Howard, and Lutz, 1999). Their calculations took into account several client characteristics and were integrated into a phase theory that postulates different recovery curves based on patient characteristics and phase of treatment (Howard, Lueger, Maling, & Martinovich, 1993). It was shown that this system could assess clients' outcomes against their ETR, which would support clinical decision making (Lueger et al., 2001).

Lambert, Hansen, and colleagues (1996) simplified this process by simply calculating patient recovery curves based on severity of symptom distress as measured by their intake score on the Outcome Questionnaire-45 (OQ-45). The OQ-45 is a 45-item routine outcome measure using a five-point Likert scale. While the measure contains three dimensions (i.e., social role performance, subjective discomfort, and interpersonal relationships) the total score is primarily used for analysis and interpretation. There is also evidence to show adequate validity and reliability of the OQ-45 (Lambert, Hansen, et al., 1996; Lambert, Burlingame, et al., 1996). Much research has been done to show the practical applications of this method of outcome monitoring. In a study conducted by Asay, Lambert, Gregersen, and Goates (2002) at a private practice, the OQ-45 was administered to all the clients of a PhD level therapist Through this

study the researchers illustrated several benefits to the private practitioner. With the data collected practitioners could compare progress and outcomes of clients in their own practice with patients from larger national samples. This information could also be used to satisfy third party requirements for treatment (e.g., insurance companies). The practitioner could also use the information gathered to create a database that could be used for comparison of future clients. The OQ-45 also can alert therapists to clients that are at risk for leaving therapy with a negative outcome. Another outcome study was conducted by Hawkins, Lambert, Vermeersch, Slade, and Tuttle (2004) using the OQ-45 where three treatment conditions were used, including treatment as usual, therapist feedback, and therapist and client feedback. The results of the study showed significantly greater improvement for clients in the feedback conditions. This supports the argument that using feedback from a ROM can enhance treatment. These findings were corroborated by a meta-analytic review of three studies conducted by Lambert and colleagues (2003). In addition to findings that supported treatment enhancement for all clients, it was also found that in the feedback condition using the OQ-45 the percentage of clients that left treatment with a negative outcome was significantly less (5%) than the percentage of clients that left treatment with a negative outcome in the no feedback condition (9%). In addition to the 45-item measure, they have also created several other measures to meet other practical or clinical needs. There is a short version containing only 30 items and a very short version that only contains 10 items. Burlingame and colleagues (2001) created a Youth Outcome Questionnaire (YOQ) for child and adolescent populations and Carey (2001) created the Severe Outcome Questionnaire (SOQ) for populations with severe mental illness.

The creators of the CORE-OM took a very similar approach. The CORE-OM (Barkham et al., 2010) was designed from the beginning with flexibility in mind. It consists of a base of 34

questions that assess four dimensions including well-being, problems, functioning, and risk. While 34 items is typically considered a reasonable length for a routine outcome measure, the researchers also provide subsets of the original 34 items to form smaller versions that are even more brief and manageable for routine assessment. They created a CORE-NR if the clinician wants to exclude the risk items. There is a CORE-10 if the clinician only wants a quick screening tool including risk. There is even a five-item version (CORE-5) to be extremely brief and feasible for routine assessment. Another option they provide is to split up the 34 items of the CORE-OM into two 18 item measures, Form A and Form B Short forms, to be administered every other week. This allows the client to answer the whole CORE-OM questions every two sessions while only requiring the client to take an 18-question measure each week instead of a 34-item measure. Additionally there is a form created for a non-clinical population (GP-CORE), which contains only 14 items. Finally, there are also two other measures designed for specific populations including those with learning disabilities (LD-CORE) and young people ages 11–16 (YP-CORE). The unique element of the CORE is its flexibility that allows the clinician to decide what version is best for their style and clientele based on length, level of risk, or demographic characteristics. There have been several studies documenting the benefits of using the CORE-OM in practice. In a study done by Lucock and colleagues, (2003) researchers administered the CORE-OM at five points before and after therapy. They also found that three quarters of clients fell within the clinical range at all three pre-therapy administrations of the measure. Post-therapy administrations, which occurred at discharge and a six-month follow up, showed that fewer than half of the clients fell within the clinical range. These results were also statistically significant between the average of the three pre-therapy measures and the discharge and follow up measures. Lucock and colleagues (2003) also analyzed this data using Jacobson and Truax's RCI

(1991) and found that "a total of 42% of clients show reliable and clinical improvement and none showed reliable and clinical deterioration, while 58% show statistically reliable improvement and 3% showed statistically reliable deterioration" (pp. 393-394). Additionally the study was able to show a steady baseline with the three administrations before therapy (i.e., referral, assessment, and pre-therapy), which indicates that improvement was not seen before the services began. The study also showed that significant improvement was maintained for six months after the cessation of treatment. In addition to aggregate analyses done on large data sets, single case studies can be conducted when the measure is used on a session-by-session basis. Lucock and colleagues (2003) also demonstrated how an analysis of major life events correlated with one individual's scores throughout the course of treatment. This information allows the therapist to use the scores from the CORE-OM to better understand each individual client and the experience and to what degree that is affecting their life.

In another study (Stiles et al., 2003) researchers used the CORE-OM to support a finding from an earlier study (Tang & DeRubeis, 1999) that reported early and sudden gains within a single between-sessions interval. The CORE-OM was used to further explore this phenomenon. Stiles and colleagues (2003) also found further evidence to support the sudden-gains phenomenon. The major finding from this phenomenon is gaining a greater understanding of the process of improvement and healing in psychotherapy (2003). This gives evidence to the idea that progress in therapy need not always be slow and gradual, but can occur suddenly over a relatively short period of time. This information can be used toward treatment planning and treatment adjustment by better understanding clients' change process while it is happening instead of just after the fact. Both the OQ and the CORE-OM created a base assessment as well as several other versions to allow flexibility in the balance of relevance, depth, and length.

Miller, Duncan, Sorrel, and Brown (2005) created an outcome measure based partially on the OQ subscales with the vision of making it the ultimate brief and feasible measure, based on the theory that clinicians will only use a measure that is quick and easy to use. They created a four-item measure called the Outcome Rating Scale (ORS) that assesses four dimensions: (a) symptom distress, (b) relationship distress, (c) social relationships outside the home, and (d) a general sense of well-being (Duncan, 2012). The ORS was found to have adequate validity and reliability (Miller, Duncan, Brown, Sparks, & Claud, 2003). Concurrent validity of the ORS has been examined across three studies (Bringhurst, Watson, Miller, & Duncan, 2006; Campbell & Hemsley, 2009; Miller et al., 2003) and was found to have moderately strong concurrent validity (r = .59) with the OQ-45 (Campbell & Hemsley, 2009; Duncan, 2012). In addition to the ORS, they also created another brief measure that measures the therapeutic alliance called the Session Rating Scale (SRS; Miller et al., 2005). This measure is only four items as well and was also found to have adequate validity and reliability (Duncan et al., 2003).

With each measure only four items apiece it becomes feasible to utilize both measures on a weekly basis and still have significantly shorter assessment times than most outcome measures. The use of these measures has been shown to improve both alliance and outcome (Duncan, Miller, Wampold, & Hubble, 2010). The SRS particularly is a unique feature that is not typically included in routine assessment. The authors of this measure stress the importance of the therapeutic alliance in client progress. They argue that it is beneficial to monitor that alliance on a session-by-session basis. Doing this allows the therapist to quickly identify and address issues arising in the relationship before they become problematic. Other outcome measures often require so many items that it is not feasible to also measure the therapeutic alliance. The brevity of the ORS and SRS allow both measures to be feasibly used together every session. Not only

does it become more feasible because of brevity, but it also adds simplicity. Other measures are both longer and more complicated. The Working Alliance Inventory, for example, was designed for research purposes and therefore has a level of complexity that is difficult to translate into routine practice (Duncan et al., 2003).

The creators of the Treatment Outcome Package (TOP) took a different approach to outcome assessment by creating a multidimensional measure that captures several domains of client symptoms and functioning. Though the development of the TOP began with 250 questions, it has been shortened significantly (Kraus et al., 2005). The current version contains 58 items spanning 12 dimensions that cover diagnostic symptoms and functioning areas (Youn, Kraus, & Castonguay, 2012). There is evidence to support good model fit and test-retest reliability, as well as convergent, concurrent, and discriminant validity (Kraus et al., 2005). Naturally with 12 dimensions being assessed the measure is significantly longer than the most routine outcome measures and takes about eight minutes to administer (Youn et al., 2012).

This is a stark contrast to the philosophy behind the ORS and SRS, which take less than a minute apiece to administer. However, the philosophy behind the TOP is that by measuring multiple dimensions of functioning the therapist can get a more complete picture of the client. The therapist has access to more detailed information to formulate treatment plans or adjust treatment based on the breadth of information provided (Youn et al., 2012). The argument then is that the therapist will have information that is more practical and detailed instead of just a single overall symptom distress score. Another argument put forth by the authors is that with multiple dimensions of functioning being assessed a more complete picture of therapists' effectiveness is painted. For example, in one study they found that 92% of clients showed statistically and clinically significant change in at least one dimension on the TOP (Kraus et al., 2005). This

statistic not only speaks to client improvement, but also therapist effectiveness. We can also conclude that 91% of the time therapists were effective in at least one dimension. Another study (Kraus, Castonguay, Boswell, Nordberg, & Hayes, 2011) analyzed which therapists were effective for which diagnostic or functional cluster. It was found that 96% of therapists were reliably effective in treating at least one of those clusters. Had only one or two dimensions been assessed in these studies, the results may have indicated that far fewer clients reliably improved and far fewer therapists were reliably effective. This means that if only one dimension is being used to evaluate effectiveness, a very narrow understanding of effectiveness is achieved. This type of multidimensional research helps to answer questions beyond what treatment is effective and begins to ask in what way it is effective. The research produced from the TOP suggests that it may be less important to come to a consensus about the most important outcome than it is to assess several important aspects of outcome in an effort to capture the complexity of this phenomenon.

This literature review shows just a small sample of the major outcome measures in the field, but represents the major differences and theories among them. Each method has their benefits and weaknesses. One of the primary challenges of a routine outcome measure is the balance between relevance and feasibility. The information needs to be detailed enough to tap into relevant issues but brief enough to be feasibly administered on a routine basis. The ORS, for example, represents the extreme in brevity and feasibility whereas the TOP represents the extreme in detailed and relevant content.

Limitations of ROM. Many research studies have been conducted to create and validate ROMs. However, the implementation rates by practitioners of ROMs remain relatively small (Bewick, Trusler, Mullin, Grant, & Mothersole, 2006; Bickman et al., 2000; Brand, 2008;

Garland, Kruse, & Aarons, 2003; Gerdes, Edmonds, Haslam, & McCartney, 1996; Gilbody, House, & Sheldon, 2002; Hatfield & Ogles, 2004; Patterson, Matthey, & Baker, 2006; Ventimiglia, Marschke, Carmichael, & Loew, 2000; Zimmerman & McGilchey, 2008). Implementation rates are so low, in fact, that there are research studies and articles dedicated to investigating the reasons behind the lack of implementation of these empirically validated ROMs (Abrahamson, 1999; Garland et al., 2003; Gilbody et al., 2002; Lakeman, 2004; Roth & Fonagy, 2005; Trauer, 2010; Zimmerman & McGilchey, 2008). This research has yielded several reasons for lack of implementation including concerns with lack of resources such as time and money, feasibility, and local validity.

Garland and colleagues (2003) randomly selected 117 clinicians from a pool of 358 eligible clinicians from San Diego County. In the end, 50 clinicians participated in the study and were interviewed in both individual and group formats, and they were all given a self-report questionnaire. The data were analyzed qualitatively, and among the differing opinions expressed there was general agreement on issues with feasibility. Clinicians expressed concerns specifically with "time burdens" and "clients' literacy challenges" (Garland et al., 2003, p. 404). Additionally, many clinicians expressed conceptual concerns with the measures. Approximately 25% "were ideologically strongly opposed to quantifying the complexity and nuance of human change in psychotherapy" (p. 398). Another 25% believed that the measurement of psychotherapeutic outcomes is "virtually impossible" (p. 398). Many clinicians also expressed concern that they were not able to take part in the decision-making process for the design and implementation of the outcome measure and the study as a whole.

Meehan and colleagues (2006) conducted focus groups with 324 mental health staff from a large mental health organization in Australia. The authors noted that clinician reactions to

outcome measures are frequently ignored by researchers, and they therefore extoled the need for research in this area. Some of the logistical concerns that were frequently mentioned include competing work demands on clinicians' time, not having access to the computers to administer the measures, slowness of the computer networks, lack of computer skills of the staff, forgetting passwords, and not understanding the summary graphs. These are all of course valid concerns that take up the clinicians' time and can add substantial frustration to the process of implementing routine outcome measures. Other concerns had a more fundamentally theoretical foundation. One of the concerns raised in these focus groups was the validity of the measure. A quote that characterized and summarized this concern was that they felt that the measure was "too brief and broad to be useful" (Meehan et al., 2006, p. 583). Here the clinicians are asking for more specific, detailed questions that would be more applicable to their particular clients. Another concern was that the clinicians felt that the measures were "reductionist in that they reduced a consumer's life experiences to a single score/number" (p. 583). This has often been a criticism: that the outcome measures don't offer enough actionable data.

In a study done by Hatfield and Ogles (2004), researchers surveyed approximately 600 practitioners in the United States. Of the survey questions provided to the practitioners, the most highly endorsed concerns included paperwork burden, time burden, and burden on clients. Again, these practitioners bring up logistical concerns for why they do not engage in ROM. The fourth highest endorsed reason for not using an outcome measure was that they felt like it would not be helpful or relevant to practice.

Lakeman (2004) wrote about this concern as well stating that "a serious flaw of ROM is its failure to capture the subtlety of individual differences" (p. 211). He goes on to say that "routine outcome measurement in its present form cannot capture individual differences with any

clinically useful sensitivity" (p. 214). This is a bold claim and a valid concern that needs to be addressed to achieve widespread acceptance among clinicians. These same concerns were echoed by Happell (2008), who critiqued several outcome measures that were being used in Victoria, Australia, following the research of Smith, Manderscheid, and Flynn's (1997) 12 principles of routine outcome measurement. She concluded that the current outcome measures being used in Victoria do "not produce information of value to clinicians or to the consumers of mental health services themselves" (p. 323). Happell goes on to say that ROMs currently being utilized in Victoria "do not reflect the areas considered of most relevance and importance to mental health consumers" (p. 317). This is not only a concern in the area of psychotherapy but also in other related fields like psychiatry. Gilbody and colleagues analyzed questionnaires from 340 psychiatrists mostly working in hospital settings in the United Kingdom (2002). One of the main concerns put forth by these psychiatrists was that ROMs don't "capture the subtlety of multi-faceted outcome and describe the individual patient" (p. 102). There is also evidence to suggest that consumers of services also hold these concerns (Graham et al, 2001; Miller, Siggins, Kavanagh, & Donald, 2003). Two studies in Australia surveyed consumers of mental health services and both obtained similar findings. The consumers reported that while the existing measures did collect some valuable information they were severely lacking in many other aspects of outcomes that they considered important (Graham et al., 2001; Miller et al., 2003).

One of the most common issues brought up by clinicians is the burden of extra time necessitated by ROM (Hatfield & Ogles, 2004; Meehan et al., 2006). Many developers of ROM have developed online administrations that require little involvement from the clinician in the actual administration of the test, especially at centers where secretaries administer the tests. The time burden has also been addressed by making scoring and reporting electronic; these results are

easily and immediately accessible after survey completion (Slade, Lambert, Harmon, Smart, & Bailey, 2008).

While much research has accurately identified and sought to address concerns of practicality (e.g., paperwork, time burden, insufficient resources, burden on client), the concern of relevance is the more difficult and less often addressed concern. In fact, many ROM researchers will simply avoid or dismiss these concerns. In the study done by Meehan and colleagues (2006), after citing clinicians concerns with validity and reductionism, the authors dismiss these concerns by saying "these concerns regarding validity and scope of the instruments highlighted a lack of understanding of the measures" (p. 583). Instead of considering options for addressing clinician concerns, the researchers dismiss these concerns and attribute it to a "lack of understanding" on the clinicians' part. In the 2004 study by Hatfield and Ogles, they give short shrift to the concerns of relevance and utility by simply addressing that some resistance from practitioners comes from insight-oriented practitioners. They dismiss this concern by simply by saying that there is a need for "various outcome measures" based on the different needs of theoretical orientation. In their follow-up study they state that issues with relevance are best addressed by the administrators instead of the researchers (Hatfield & Ogles, 2007).

While many researchers fail to address these questions of relevance seriously, many others have made progress in increasing relevance through various methods. The main method for addressing relevance has been to make different versions of tests become more relevant for certain populations. The Outcome Questionnaire 45 is a good example of addressing this concern by making population-specific adaptation of the original OQ. Burlingame and colleagues (2001) developed the Youth Outcome Questionnaire (YOQ) for a child and adolescent population and Carey (2001) tested the psychometric properties of the Severe Outcome Questionnaire (SOQ),

which was created for a population with severe mental illness. While this does increase relevance among certain populations, the argument still stands that within each population, the questionnaires do not address clinicians' concerns about lack of relevance for their adult outpatient clients. Furthermore, it follows the suggestion made by Hatfield and Ogles, who alluded to the idea that different outcome measures should be made for population and therapeutic approaches (2004)—a task that would lose standardization as well as feasibility.

Part of practitioners' relevancy argument is that unidimensional measures only give one total score and focus only on alerting therapists of a client's overall level of distress and documenting if that overall level of distress decreases over treatment. This single score can then be used to determine if the client is on track for a positive outcome or not. This is a very blunt tool that only informs the clinician of two possible scenarios: either patients are on track or off track. If they are on track then the clinician continues treatment as usual. If they are off track then the therapist considers reassessing the treatment plan. In short, the tool is relevant to the degree of saying, "keep going" or "make a change." With this limited use, it is easy to see how clinicians cite lack of relevance or clinical utility as a primary concern for ROM implementation. Therefore one tactic that ROM researchers have used has been to create measures with multiple dimensions, thus providing more detailed information. The CCAPS addresses issues of relevance by expanding the measurement of general distress to a multidimensional model that assesses relevant dimensions of client functioning. This broadens the scope of assessment and makes it more likely to tap into something relevant for the client. However, the CCAPS is still only narrowly applicable to the college student population. The TOP is more broadly applicable to an adult population, giving the therapist summaries of different clusters of symptoms or areas of functioning. This information can then be used to guide therapy toward certain topics based on

the dimension's scores instead of just knowing one score that would indicate, "keep going" or "reevaluate your approach." While each of these approaches has perhaps increased relevance and eased practitioners' concerns to a degree, there are deeper and more complex reasons why the argument of relevance and local validity are so poignant. These reasons have not yet been adequately addressed, and it is important to do so to increase the acceptance of ROM among practitioners.

In addition to the practical concerns with relevance, there are also theoretical reasons why clinicians and researchers clash on issues of relevance. This is a profound concern that challenges the adequacy of the principles of standardization and generalizability in psychological assessment and outcome measurement. These principles are grounded in the philosophy of abstractionism, which is one of the naturalistic assumptions that often goes unexamined in psychology and science (Slife, 2004). When grounded in the philosophy of abstractionism, the very nature of psychological measurement is to overlook the nuances and complexities of each individual in favor of standardization and generalizability (Sonnanburg, 1996).

ROMs follow these assumptions and principles, which is demonstrated by the fact that they are developed and designed explicitly to apply to all persons, contexts, races, and religions; every socioeconomic status; and every presenting concern. This methodology has its place and yields valuable information, but it should not be seen as the *only* way to measure outcome. This is a glaring example of how defining outcome through a solely universalistic paradigm provides a one-sided view of outcome and consequently an impoverished understanding of assessment and treatment.

The assumption of abstractionism is of obvious concern to the practitioners in the studies cited earlier as evidenced by their complaints of local validity and the overall lack of

implementation of these ROM, but it should also pose significant concern to the APA considering the theoretical conflicts inherent in statistical standardization. This solely universalistic approach to ROM is in direct conflict with the definition of EBPP put forth by the APA that states that "psychological services are most effective when responsive to patient's specific problems, strengths, personality, sociocultural context, and preferences" (APA, 2006, p. 284). If researchers were to include these characteristics into their ROM, then practitioners would be more likely to see their value and utility, which would result in higher implementation rates and the aims of EBPP would be better realized.

Another issue that arises when discussing the assessment of population specific or individualized assessment is feasibility. It is virtually impossible to create a measure that takes into account all possible population-specific characteristics, not to mention individual characteristics (e.g., culture, preference). A possible solution to this problem has been to use several different measures that have been designed for specific populations or problems. This does address concerns of local validity and begins to step away from a solely universalistic approach, yet this becomes cumbersome to administer and can be quite lengthy if a client is taking multiple measures every week. Thus, the problem remains that researchers seek to develop standardized measures that meet criteria for reliability and validity (abstractionism), but practitioners feel the outcomes being measured are not relevant or specific enough (contextualism) to be useful. A potential solution for this dilemma is to create a hybrid assessment system that maintains universality while also incorporating a greater level of contextual applicability.

The Clinically Adaptive Multidimensional Outcome Survey

The Clinically Adaptive Multidimensional Outcome Survey (CAMOS) was designed for this purpose. The CAMOS builds upon two theoretical ideas from two different paradigms. The first feature of the CAMOS is designed to incorporate traditional standardized statistical assessment. The eight dimensions of the CAMOS globally assess eight aspects of human experience (therapy progress, relationships, distressing behaviors, distressing thoughts, distressing emotions, spirituality, work, and physical health). This multisystem structure was created using a theoretical rational (Lazarus, 1973, 2007; Richards & Bergin, 2005) and is also generally supported empirically (Miovic et al., 2006; Slade, 2002). This multidimensional structure builds on the idea that by providing dimension scores therapists will have a broader range of useful and actionable information on which they can base decisions. For instance, a dimension score indicating high distress may prompt the therapist to make the topic of that dimension the focus of the current session. This multidimensional method provides a more complete picture of clients' concerns and helps them make decisions about treatment focus. It is important to note the inclusion of a dimension that assesses spirituality, which is another way in which the CAMOS increases relevance and local validity. There is evidence to suggest that most people are religious (Newport, 2014) and that therefore assessing spirituality would be relevant to the great majority of clients. By assessing spirituality, the CAMOS taps into an aspect of life relevant to treatment that most outcome measures ignore. In this way even the "standardized" portion of the CAMOS will have greater applicability and breadth than most outcome measures by taking into account this aspect of a clients' "characteristics, culture, and preferences" (APA, 2006, p. 273). In a survey of clients conducted by Miovic and colleagues (2006), they found that the fifth and sixth most relevant topics in treatment were spiritual or religious beliefs and

spiritual or religious practices; yet this outcome is not included in mainstream routine outcome monitoring.

Tracking concerns in the areas of emotions, thoughts, and behaviors is nothing new and is widely accepted in the literature. The therapy expectations dimension was intended to measure the clients' anxiety or concern for entering therapy and closely relates to the therapeutic alliance literature that documents the importance of the therapeutic relationship and its effect on outcomes (Horvath & Symonds, 1991). The importance of relationships is clearly addressed in the literature as an important aspect of mental health and is often referred to as the client's social support system (Bankoff, 1996; Bankoff & Howard, 1992; Lara, Leader, & Klein, 1997; Moras & Strupp, 1982). There is also support for the idea of including the assessment of physical concerns as well as psychological (Kwan & Nease, 2013; Miller, Kessler, Peek, & Kallenberg, 2011). Work, school, or social responsibility in general also has significant research backing its potential applicability to psychotherapy (Betz & Corning, 1993; Manuele-Adkins, 1992; McIlveen, 2015; Super, 1993).

Historically there has been an explicit and pervasive bias in the field of psychology toward religion and spirituality (Slife & Reber, 2009). There have been several researchers that have seen a need to assess spiritual and religious outcomes of mental health (P. Hill, & Pargament, 2003; Richards & Bergin, 2005; Worthington, Kurusu, McCollough, & Sanders, 1996). However, there is still a great lack of outcome research when it comes to religion and spirituality. There are several reasons to assess spiritual and religious outcomes, and one in particular addresses the historical bias toward religious populations. Many religious people fear that psychotherapy will undermine their religious beliefs (Richards & Bergin, 2014). Therefore it is important to conduct research in this area to show if harm is being done in this way. Including

religion and spirituality also shows multicultural sensitivity. Recent polls suggest that most people claim some religious affiliation, and it would be culturally insensitive to ignore this prominent area of life that has potential to have a significant impact on people's lives and their mental health (Newport, 2014). Additionally, many have advocated a multifaceted approach to outcome assessment (Kazdin, 1994; Lambert & Hill, 1994). There are many multifaceted theories that include spirituality: it stands to reason that there is now a need to assess spiritual and religious outcomes. In the past decade there have been several measures developed that assess religious and spiritual outcomes. However, their widespread implementation is limited due in part to the lack of integration into mainstream ROM. The field of ROM has maintained the tradition of leaving out assessment of religion and spirituality. If religion and spirituality are not integrated into ROM, then their use as a stand-alone measure is severely limited. Implementation rates of a single routine outcome measure are low, let alone multiple routine outcome measures. Asking the therapist and client to complete multiple surveys every session is burdensome. Implementing two separate outcome measures running on two different scales with two different software programs with two different modes of viewing and interpreting data makes it extremely unlikely that they will be used on a routine basis. Therefore, if religion and spirituality are to be included in mainstream routine practice then it must be integrated into a comprehensive measurement system, which is one of the current purposes of the CAMOS.

As was mentioned previously, assessing for religious or spiritual concerns is a part of multicultural competence and is in line with the definition of EBPP, which says that it is important to take into account "client characteristics, culture, and preferences (APA, 2006, p. 273)." Assessing spirituality also addresses in part practitioners' concerns about relevance and local validity. Assessing for religious and spiritual concerns is an important way to assess the

cultural concerns, and when done in an ecumenical way can still apply broadly to most clients. Including spirituality into a ROM can be an important method for more fully incorporating client culture into the best practice of psychotherapy. Religious and spiritual topics and concerns can be diverse and personal, and the spirituality dimension was not created to completely encapsulate a client's religious experience, but it is a starting point to help the therapist be aware of specific concerns in that area of life. By including spirituality into the standardized portion of the CAMOS we are expanding even the abstractionist point of view by including a relevant aspect of life that has been previously left out.

The second feature of the CAMOS provides opportunities for the therapist to assess clients through a contextual and relational approach. The CAMOS's global eight-dimensional structure can be expanded to include population-specific dimensions (e.g., eating disorder dimension, cultural sensitivity dimension). It also allows therapists and clients to tailor the assessment to their unique context and relationship by creating unique items that apply only to that client and their work in therapy, an idea that is based in the area of individualized assessment (Haynes, Mumma, & Pinson, 2009).

The CAMOS seeks to combine these two very different paradigms into an integrated system that combines standardization (facilitating comparability) with customization (facilitating local relevance). This highlights a crucial benefit of the CAMOS where both assessment approaches (abstractionism and contextualism) are integrated into one assessment and reporting platform, making it quick and easy to take for the client and to view the results for the therapist. This hybrid approach allows clinicians to be able to more aptly integrate and embody the three decision-making factors outlined by the task force on EBPP. The best available research is being represented in the core eight dimensions, which will be validated and standardized. Clinical

expertise and decision making is enhanced by providing the clinician with a platform to collect information pertaining to specific "patient characteristics, culture, and preferences" (APA, 2006, p. 273) through the use of a population-specific dimension and an individual dimension. At the same time we will be addressing clinicians' concerns with local validity, thereby increasing the likelihood of ROM implementation. This will all be provided through an integrated system that effectively combines breadth of assessment with brevity and abstractionism with contextualism. The first step in this process is to establish a multidimensional standardized base that will allow for flexibility and adaptability.

References

Abrahamson, D. J. (1999). Outcomes, guidelines, and manuals: On leading horses to water. *Clinical Psychology: Science and Practice, 6,* 467–471.

Addis, M. E. (2002). Methods for disseminating research products and increased evidence-based practice: Promises obstacles, and future directions. *Clinical Psychology: Science and Practice, 9,* 367–378.

Addis, M. E., & Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes towards psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology, 65*, 331–339.

American Psychological Association. (2006). Evidence-based practice in psychology: APA presidential task force on evidence-based practice. *American Psychologist*, *61*, 271–285.

American Psychological Association Division of Clinical Psychology. (1995). Training in and dissemination of empirically validation psychological treatments: Report and Recommendations. *The Clinical Psychologist, 48,* 3–27.

Asay, T. P., Lambert, M. J., Gregersen, A. T., & Goates, M. K. (2002). Using patient-focused research in evaluating treatment outcome in private practice. *Journal of Clinical Psychology, 58*, 1213–1225.

Barkham, M., Hardy, G. E., & Mellor-Clark, J. (2010). *Developing and delivering practice-based evidence: A guide for the psychological therapies*. doi:10.1002/9780470687994

Barkham, M., Mellor-Clark, J., Connell, J., Evans, C., Evans, R., & Margison, F. (2010). Clinical outcomes in routine evaluation (CORE)—The CORE measures and system: Measuring, monitoring and managing quality evaluation in the psychological therapies. In M. Barkham, G. E. Hardy, & J. Mellor-Clark, (Eds.), *Developing and delivering practice-*

based evidence: A guide for the psychological therapies (pp. 175–219). doi:10.1002/9780470687994.ch8

Betz, N. E., & Corning, A. F. (1993). The inseparability of "career" and "personal" counseling. *Career Development Quarterly, 42*, 137–142. doi:10.1002/j.2161-0045.1993.tb00426.x

Beutler, L. E., Williams, R. E., Wakefield, P. J., & Entwistle, S. R. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist*, *50*, 984–994. doi:10.1037/0003-066X.50.12.984

Bewick, B. M., Trusler, K., Mullin, T., Grant, S., & Mothersole, G. (2006). Routine outcome measurement completion rates in primary care psychological therapies and counselling. *Counselling and Psychotherapy Research, 6,* 33–40.

Bickman, L., Rosof-Williams, J., Salzer, M. S., Summerfelt, W. T., Noser, K., Wilson, S. J., & Karver, M. S. (2000). What information do clinicians value for monitoring adolescent client progress and outcomes? *Professional Psychology: Research and Practice*, *31*, 70–74. doi:10.1037/0735-7028.31.1.70

Brand, N. (2008). Psychologists' approaches to evaluating client progress in psychotherapy practice. *Dissertation Abstracts International, 68,* 5558.

Bringhurst, D. L., Watson, C. W., Miller, S. D., & Duncan, B. L. (2006). The reliability and validity of the Outcome Rating Scale: A replication study of a brief clinical measure. *Journal of Brief Therapy, 5,* 23–30.

Burlingame, G. M., Mosier, J. I., Wells, M. G., Atkin, Q. G., Lambert, M. J., Whoolery, M., & Latkowski, M. (2001). Tracking the influence of mental health treatment: The development of the Youth Outcome Questionnaire. *Clinical Psychology & Psychotherapy*, *8*(5), 361–379. doi:10.1002/cpp.315

Campbell, A., & Hemsley, S. (2009). Outcome Rating Scale and Session Rating Scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist, 13*, 1–9.

Carey, J. A. (2001, April). The Severe Outcome Questionnaire: A preliminary study of reliability and validity. *Dissertation Abstracts International*, *61*, 5554.

Cohen, L. H., Sargent, M. M., & Sechrest, L. B. (1986). Use of psychotherapy research by professional psychologists. *American Psychologist*, *41*, 198–206. doi:10.1037/0003-066X.41.2.198

Cook, J. M., Schnurr, P. P., Biyanova, T., & Coyne, J. C. (2009). Apples don't fall far from the tree: Influences on psychotherapists' adoption and sustained use of new therapies. *Psychiatric Services*, *60*, 671–676. doi:10.1176/appi.ps.60.5.671

Drabick, D. G., & Goldfried, M. R. (2000). Training the scientist–practitioner for the 21st century: Putting the bloom back on the rose. *Journal of Clinical Psychology*, *56*, 327–340. doi:10.1002/(SICI)1097-4679(200003)56:3<327::AID-JCLP9>3.0.CO;2-Y

Duncan, B. L. (2012). The Partners for Change Outcome Management System (PCOMS): The Heart and Soul of Change Project. *Canadian Psychology/Psychologie Canadienne*, *53*, 93–104. doi:10.1037/a0027762

Duncan, B. L., Miller, S. D., Sparks, J. A., Claud, D. A., Reynolds, L. R., & Brown, J. (2003). The Session Rating Scale: Preliminary psychometric properties of a working alliance measure. *Journal of Brief Therapy, 3,* 3.

Duncan, B. L., Miller, S. D., Wampold, B. E., & Hubble, M. A. (2010). *The heart and soul of change: Delivering what works in therapy* (2nd ed.). Washington, DC: American Psychological Association. doi:10.1037/12075-000

Garland, A. F., Kruse, M., & Aarons, G.A. (2003). Clinicians and outcome measurement. What's the use? *Journal of Behavioral health Services & Research, 30*, 393–405.

Gerdes, K., Edmonds, R., Haslam, D., & McCartney, T. (1996). A statewide survey of licensed clinical social workers' use of practice evaluation procedures. *Research on Social Work Practice, 6,* 27–39.

Gilbody, S. M., House, A. O., & Sheldon, T. A. (2002). Psychiatrists in the UK do not use outcome measures. *British Journal of Psychiatry, 180,* 101–103.

Goldfried, M. R., & Wolfe, B. E. (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology, 66,* 143–150.

Graham, C., Coombs, T., Buckingham, W., Eagar, K., Trauer, T., & Callaly, T. (2001) The Victorian mental health outcomes measurement strategy: Consumer perspectives on future directions for outcome self-assessment. *Report of the Consumer Consultation Project.* Retrieved from http://www.health.vic.gov.au/mentalhealth/ outcomes/concare.htm

Happell, B. (2008). The value of routine outcome measurement for consumers of mental health services: Master or servant? *International Journal of Social Psychiatry*, *54*, 317–327. doi:10.1177/0020764008090285

Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice, 35*, 485–491. doi:10.1037/0735-7028.35.5.485

Hatfield, D. R., & Ogles, B. M. (2007). Why some clinicians use outcome measures and others do not. *Administration and Policy in Mental Health and Mental Health Services Research*, *34*, 283–291. doi:10.1007/s10488-006-0110-y

Hawkins, E. J., Lambert, M. J., Vermeersch, D. A., Slade, K., & Tuttle, K. (2004). The effects of providing patient progress information to therapists and patients. *Psychotherapy Research, 31*, 308–327.

Haynes, S. N., Mumma, G. H., & Pinson, C. (2009). Idiographic assessment: Conceptual and psychometric foundations of individualized behavioral assessment. *Clinical Psychology Review*, *29*, 179–191. doi:10.1016/j.cpr.2008.12.003

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, *38*, 139–149. doi:10.1037/0022-0167.38.2.139

Howard, K. I., Brill, P. L., Lueger, R. J., & O'Mahoney, M. (1992). Integra Outpatient Tracking Assessment. Radnor, PA: Integra, Inc.

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy: Causal mediation of outcome. *Journal of Consulting and Clinical Psychology, 61,* 678–685.

Howard, K. I., Moras, K., Brill, P., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51,* 1059–1064.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59,* 12–19.

Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin, & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.). New York, NY: Wiley.

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, *63*, 146–159. doi:10.1037/0003-066X.63.3.146

Kettlewell, P. W. (2004). Development, dissemination, implementation of evidenced-based treatments: Commentary. *Clinical Psychology: Science and Practice, 11,* 190–195.

Kraus, D. R., Castonguay, L. G., Boswell, J. F., Nordberg, S. S., & Hayes, J. A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research, 21*, 267–276. doi:10.1080/ 10503307.2011.563249

Kraus, D. R., Seligman, D., & Jordan, J. R. (2005). Validation of a behavioral health treatment outcome and assessment tool designed for naturalistic settings: The treatment outcome package. *Journal of Clinical Psychology, 61,* 285–314. doi:10.1002/jclp.20084

Kwan, B. M., & Nease, D. (2013). The state of the evidence for integrated behavioral health. In M. R. Talen, & A. B. Valeras (Eds.), *Essentials of integrated care: Connecting systems of care, clinical practice and evidence-based approaches* (pp. 65–98). New York, NY: Springer Science Business Media.

Lakeman, R. (2004). Standardized routine outcome measurement: Pot holes in the road to recovery. *International Journal of Mental Health Nursing, 13,* 210–215.

Lambert, M. J., Burlingame, G. M., Umphress, V. J., Hansen, N. B., Vermeersch, D., Clouse, G., & Yanchar, S. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy, 3,* 106–116.

Lambert, M. J., Bergin, A. E., & Garfield, S. L. (2004). Introduction and historical overview. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed.). New York, NY: Wiley.

Lambert, M. J., Bergin, A. E., & Garfield, S. L. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.). New York, NY: Wiley.

Lambert, M. J., Hansen, N. B., Umphress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., Heufner, J. C., & Reisinger, C. R. (1996). Administration and scoring manual for the Outcome Questionnaire (OQ-45.2). Wilmington, DE: American Professional Credentialing Services LLC.

Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In A. E. Bergin, S. L. Garfield, A. E. Bergin, & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 72–113). Oxford, England: John Wiley & Sons.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 70*, 288–301.

Lara, M. E., Leader, J., & Klein, D. N. (1997). The association between social support and course of depression: Is it confounded with personality? *Journal of Abnormal Psychology, 106,* 478–482.

Lazarus, A. A. (1973). Multimodal behavior therapy: Treating the BASIC ID. *Journal of Nervous and Mental Diseases, 156*, 404–411.

Lazarus, A. A. (2007). Multimodal therapy. In R. J. Corsini, & D. Wedding (Eds.), *Current psychotherapies* (8th ed., pp. 337–371). Belmont, CA: Brooks Cole.

Leon, S. C., Kopta, S. M., Howard, K. I., & Lutz, W. (1999). Predicting patients' responses to psychotherapy: Are some more predictable than others? *Journal of Consulting and Clinical Psychology, 67*, 698–704.

Lucock, M., Hall, P., & Noble, R. (2006). A survey of influences on the practice of psychotherapists and clinical psychologists in training in the UK. *Clinical Psychology & Psychotherapy*, *13*, 123–130. doi:10.1002/cpp.483

Lucock, M., Leach, C., Iveson, S., Lynch, K., Horsefield, C., & Hall, P. (2003). A systematic approach to practice-based evidence in a psychological therapies service. *Clinical Psychology & Psychotherapy*, *10*, 389–399. doi:10.1002/cpp.385

Lueger, R. J., Howard, K. I., Martinovich, Z., Lutz, W., Anderson, E. E., & Grissom, G. (2001). Assessing treatment progress of individual patients using expected treatment response models. *Journal of Consulting and Clinical Psychology, 69*, 150–158.

Manuele-Adkins, C. (1992). Career counseling is personal counseling. *Career Development Quarterly, 40*, 313–323. doi:10.1002/j.2161-0045.1992.tb00338.x

McIlveen, P. (2015). Psychotherapy, counseling, and career counseling. In P. J. Hartung, M. L. Savickas, & W. B. Walsh, (Eds.), *APA handbook of career intervention: Foundations* (Vol. 1, pp. 403–417). doi:10.1037/14438-022

Miller, B. F., Kessler, R., Peek, C. J., & Kallenberg, G. A. (2011). *A national research agenda for research in collaborative care: Papers from the collaborative care research network research development conference* (AHRQ Publication No. 11–0067). Rockville, MD: Agency for Healthcare Research and Quality.

Miller, M., Siggins, I., Kavanagh, D., & Donald, M. (2003) *Consumer self-rated outcome measures in mental health*. Melbourne, Australia: Mental Health Branch, Department of Human Services.

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J., & Claud, D. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*, 91–100.

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The Partners for Change Outcome Management System. *Journal of Clinical Psychology, 61*, 199–208.

Miovic, M., McCarthy, M., Badaracco, M. A., Greenberg, W., Fitzmaurice, G. M., & Peteet, J. R. (2006). Domains of discussion in psychotherapy: What do patients really want? *American Journal of Psychotherapy*, *60*, 71–86.

Morrow-Bradley, C., & Elliott, R. (1986). Utilization of psychotherapy research by practicing psychotherapists. *American Psychologist*, *41*, 188–197. doi:10.1037/0003-066X.41.2.188

Newport, F. (2014, July 28). Religion remains a strong marker of political identity in U.S. *Gallup.* Retrieved from http://www.gallup.com/poll/174134/religion-remains-strong-marker-political-identity.aspx

Ogilvie, A. J. (2012). The customer is always right: Bridging the research-practice gap by identifying consumer-friendly research methods. *Dissertation Abstracts International*, *72,* 4991.

Patterson, P., Matthey, S., & Baker, M. (2006). Using mental health outcome measures in everyday clinical practice. *Australasion Psychiatry, 14*, 133–136.

Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66*, 126–135.

Roth, A. & Fonagy, P. (2005). *What works for whom: A critical review of psychotherapy research* (2nd ed.). New York, NY: Guilford Publications.

Richards, P. S., & Bergin, A. E. (2005). *A spiritual strategy for counseling and psychotherapy,* (2nd ed.). doi:10.1037/11214-000

Richards, P. S., & Bergin, A. E. (2014). *Handbook of psychotherapy and religious diversity* (2nd ed.). doi:10.1037/14371-000

Slade, K., Lambert, M. J., Harmon, S. C., Smart D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Journal of Clinical Psychology and Psychotherapy, 15*, 287–303.

Slade, M. (2002). What outcomes to measure in routine mental health services, and how to assess them: A systematic review. *Australian and New Zealand Journal of Psychiatry*, *36*, 743–753. doi:10.1046/j.1440-1614.2002.01099.x

Slife, B. D. (2004). Theoretical challenges to therapy practice and research: The constraint of naturalism. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 44–83). New York, NY: Wiley.

Slife, B. D., & Reber, J. S. (2009). Is there a pervasive implicit bias against theism in psychology? *Journal of Theoretical and Philosophical Psychology*, *29*, 63–79. doi:10.1037/a0016985

Smith, G. R., Manderscheid, R. W., Flynn, L. M., & Steinwachs, D. M. (1997). Principles for assessment of patient outcomes in mental health care. *Psychiatric Services, 48,* 1003–1036.

Sonnanburg, K. (1996). Meaningful measurement in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, *33*, 160–170. doi:10.1037/0033-3204.33.2.160

Stewart, R. E., & Chambless, D. L. (2007). Does psychotherapy research inform treatment decisions in private practice? *Journal of Clinical Psychology 63*, 267-281. doi:10.1002/jclp.20347.

Stiles, W. B., Leach, C., Barkham, M., Lucock, M., Iveson, S., Shapiro, D. A., Iveson, M., & Hardy, G. E. (2003). Early sudden gains in psychotherapy under routine clinic conditions: Practice-based evidence. *Journal of Consulting and Clinical Psychology*, *71*, 14–21. doi:10.1037/0022-006X.71.1.14

Super, D. E. (1993). The two faces of counseling: Or is it three? *Career Development Quarterly, 42,* 132–136. doi:10.1002/j.2161-0045.1993.tb00425.x

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67,* 894–904.

Trauer, T. (Ed.). (2010). Outcome measurement in mental health: Theory and practice. Cambridge, UK: Cambridge University Press.

Ventimiglia, J. A., Marschke, J., Carmichael, P., & Loew, R. (2000). How do clinicians evaluate their practice effectiveness? A survey of clinical social workers. *Smith College of Studies in Social Work, 70*, 287–306.

Youn, S. J., Kraus, D. R., & Castonguay, L. G. (2012). The treatment outcome package: Facilitating practice and clinically relevant research. *Psychotherapy*, *49*, 115–122. doi:10.1037/a0027932

Zimmerman, M., & McGilchey, J. B. (2008). Why don't psychiatrists use scales to measure outcome when treating depressed patients? *Journal of Clinical Psychiatry, 69,* 1917–1919.

APPENDIX B: SUPPLEMENTARY TABLES AND FIGURES

Table B1

Dimension comparisons

CAMOS	Lazarus	Slade	Miovic	Current CAMOS
Physical	Sensation	Physical health	Physical fitness	Physical Health Concerns
	Drugs			
Relationships	Interpersonal Relationships	Interpersonal	Relationships with family	Relationship Concerns
			Relationships with friends	
Work/School			Personal educational pursuits	Work/School Concerns
			Work	
Emotion	Affect	Cognition/Emotion		Psychological Distress
Cognition	Cognition			
	Imagery			
		Wellbeing		
Behavior	Behavior	Behavior		
Therapy Expectations		Services		Therapy Expectations
Spiritual			Spiritual or religious beliefs	Spirituality Concerns
			Spiritual or religious activities and practices	
		Society		
			Sexuality	
			Finance	
			Ethnicity and race	

Table B2
Corrected Item-Total Correlations

Therapy Expectation	Corrected Item-Total Correlation
During the past week, I had concerns about beginning therapy	.668
I felt anxious about beginning therapy	.610
I felt confident that therapy will help me improve my life	.601
I had doubts about whether my therapist will understand my concerns	.562
I felt uncertain about whether I can be fully honest and open with my therapist	.530
I felt hopeful that I will learn more about myself in therapy	.488
I doubted whether therapy will be worth the cost	.472
Relationships	
I felt misunderstood by my loved ones and friends	.716
I argued with my loved ones or friends	.690
I felt irritated and angry towards others	.663
I felt sad about how I acted towards my family or friends	.631
I felt hurt or disappointed by how my loved ones or friends behaved	.617
I felt concerned about my relationships (with your family, partner/spouse, and/or friends)	.603
I had outbursts of anger	.570
I felt accepted by my friends and loved ones	.545
I hurt others with my words or actions	.508
I thought about harming others	**.331**
Psychological	
I felt worthless or "not good enough"	.805
I felt concerned about distressing emotions:	.732
I felt concerned about distressing thoughts.	.729
I felt worried, agitated, fearful, or tense	.728
I felt sad or depressed	.727
I felt powerless or stuck in my problems	.716
I thought about past personal failures/mistakes	.700
I had difficulty concentrating or remaining focused on a task	.699
I worried about what other people thought about me	.695
I felt physically stressed or worn out	.688
I struggled with perfectionism or constant self-criticism	.681
I felt stressed out	.680
I was preoccupied with my body size or shape	.599
I felt peaceful and calm	.585
I felt concerned about my self-defeating behaviors:	.582
I had thoughts or images that I couldn't get out of my head	.520
I felt unsure about what I need to change or how to do it	**.378**

Table B2 (continued)
Corrected Item-Total Correlations

Behaviors	Corrected Item-Total Correlation
I lied or kept secrets from my loved ones or close friends	.602
I thought about harming myself	.594
I acted in impulsive, risky, or thoughtless ways	.580
I gave up on my goals and plans for the future	.570
I was honest with myself and others	.490
I harmed myself (cut, scratched, burned etc.)	.485
I used harmful substances (ex. alcohol, drugs, tobacco etc.)	**.359**
Positive	
I felt hopeful or excited for each new day	.619
I have been kind and accepting towards myself	.582
I felt confidence and faith that God will help me resolve my difficulties	.557
I was motivated and productive	.534
I worked well with others	.514
I felt positive and loving feelings for my loved ones or friends	.478
I felt accepted in my religious community	.437
I confided in or shared feelings with a loved one or friend	.422
I felt motivated to learn all that I can from therapy and put it into practice	.413
Spiritual	
I felt a loss of inspiration or spiritual direction	.739
I felt distant in my relationship with God or my Higher Power	.698
I felt guilt and regrets over mistakes that were inconsistent with my religious beliefs	.636
I felt concerned about my religious or spiritual life	.607
I experienced unsettling, troubling, or unusual religious thoughts	.498
Physical	
I experienced physical pain or discomfort	.745
I felt light headed, weak, or fatigued	.661
I had a stomach ache or other gastro-intestinal problems	.636
I felt physically well and healthy	.606
I wondered if I should see my physician	.595
I felt concerned about my physical health	.587
I experienced medical complications	.541
I worried about something bad happening to myself or loved ones	.440
I had side effects from my prescribed medications	**.336**
Work	
I felt over-burdened with too many responsibilities	.656
I worried about not meeting expectations or requirements	.591
I felt concerned about my work (i.e., employment, school, homemaking responsibilities, volunteer work, etc.):	.566
I wanted a change in my responsibilities	.564
I was in trouble for the quality of my performance	.501
I felt under-valued and unappreciated	.497

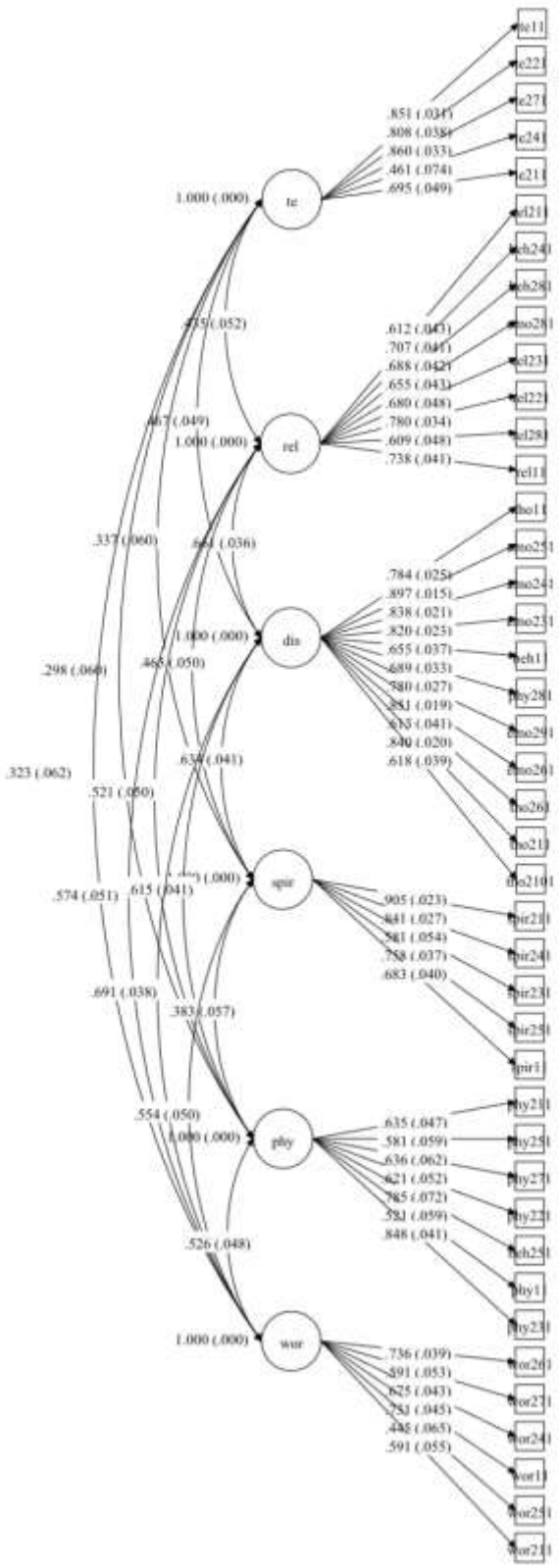

Figure B1. Path diagram for the final confirmatory factor analysis