THE USE OF CLINICAL OUTCOMES TO MEASURE WITHIN-GROUP DIFFERENCES IN COUNSELORS IN TRAINING

A Dissertation

by

KAITLYN E. STEIN

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Daniel F. Brossart
Committee Members,	Jay Ganz
	Myeongsun Yoon
	Arnold LeUnes
Head of Department,	Fuhui Tong

August 2021

Major Subject: Counseling Psychology

ABSTRACT

Understanding the development and acquisition of professional competencies of counselors in training is crucial to the field of psychology. Many assume that increased training is related to greater development of these professional competencies and improved therapeutic outcomes. Previous literature has shown inconsistency regarding the relationship between training level and clinical outcomes. Some studies claim that there are differences in outcomes among training levels (Budge et al., 2013; Driscoll et al., 2003) and some show a lack of differences among counselors with varying levels of experience (Lambert et al., 2003; Michael, Huelsman, & Crowley, 2005; Nyman et al., 2010; Propst, Paris, & Rosberger, 1994; Stein & Lambert, 1995). Few studies have focused on understanding therapist effects in outcome research for counselors in training. This study was conducted to explore within-group differences in counselors in training in their ability to facilitate clinical outcomes. For this study, Outcome Questionnaire (OQ-45.2) data from Texas A&M University's Counseling and Assessment Clinic was analyzed to better understand differences among trainees in a counseling psychology doctoral program. Results indicated that there was not a statistically significant relationship between counselor education level and client outcome. Despite this, within-group differences were observed among the highest and lowest performing post bachelor's and post master's counselors. By better understanding how counselors in training differ in their ability to generate clinical outcomes, counselor training programs might be better equipped to train and provide appropriate intervention for future cohorts of psychologists in training.

ACKNOWLEDGEMENTS

I cannot imagine having written this dissertation without the love, support, and encouragement from those close to me. To my husband, Dale, thank you for being my greatest supporter throughout this process. Words cannot express the love and appreciation that I hold for you. Thank you for encouraging me to persist in the face of challenges. To those who have supported me, I want to say thank you for your friendship, understanding, and patience over these past four years. To my family, thank you for teaching me to value education, advocacy, and helping others. I am also incredibly thankful for the mentors and those who have supported me professionally at Texas A&M University. Thank you to my clinical supervisors, chair, and committee members for sharing your wisdom and expertise.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professors Daniel Brossart, Jay Ganz, and Myeongsun Yoon of the Department of Educational Psychology and Professor Arnold LeUnes of the Department of Psychological and Brain Sciences. The data analyzed for Chapter 3 was provided from the Counseling and Assessment Clinic (CAC) by Professor Krystal Simmons, the clinic executive director of the CAC. The work conducted for this dissertation was completed by the student independently.

Funding Sources

No funding was provided for the research study described in this dissertation defense.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Counselors in Training and Therapeutic Outcomes

Measuring competency of counselors in training is an important part of the gatekeeping role that professional psychology programs hold (Lumadue & Duffey, 1999). Moreover, previous research has highlighted a need for psychology training programs to provide data showing "(a) whether the skills they teach relate directly to the year-to-year increases in the successful number or quality of therapy outcomes among the patients of trainees or (b) that specific didactic or practicum experiences affect dropout rates over time" (Stein & Lambert, 1995, p. 193). This increased understanding of what experiences impact retention and outcomes could lead to more effective training for counselors in training.

There are few measures that have been created for the purposes of assessing student growth (Spurgeon et al., 2012). Previous studies have highlighted the need for more empirical methods of assessing clinical skills in counselor training programs (Alberts & Edelstein, 1990; Ford, 1979). Currently, the American Psychological Association (APA) recommends graduate training programs use a "benchmarks evaluations system", which includes the use of a rating form to assess competence areas (2012). This rating form asks the rater "how characteristic of the trainee's behavior is this competency description?"; raters can score the trainee from 0-4 (APA, 2012). On this form 0 indicates "Not at All/Slightly", 1 indicates "Somewhat", 2 indicates "Moderately", 3 indicates "Mostly", and 4 indicates "Very" (APA, 2012). The form places little focus on the assessment of actual clinical skills. This can be shown by the functional competencies section, under application 10B "displays clinical skills" and 10C "implements

evidence-based interventions" (APA, 2012, p. 10). This rating form provides little data regarding specific clinical skills or interventions and indicates only that clinical skills and evidence-based interventions were exhibited by the trainee. These items highlight the lack of focus on direct assessment of clinical skills and outcomes in graduate training programs.

Despite the utility of outcome measures in psychotherapy research, these outcome measures are rarely used to measure a trainee's counseling skills (Budge et al., 2013; Lambert & Hawkins, 2001). It has been argued that the use of clinical outcome measures to assess counselor skills and growth would provide "a clear barometer for how trainees are progressing in assisting their clients" (Budge et al., 2013, p. 151). Several previous researchers have used client outcome measures to evaluate the performance of counselors in training (e.g. Budge et al., 2013; Iberg, 1991), but these studies have not focused on comparing therapeutic growth within student cohorts. Previous studies have focused on the differences between counselors in training, interns, and licensed professionals, but less research has addressed within-group differences among counselors in training.

Despite the assumption that increased training should yield increased psychotherapy outcomes, the current literature has yielded conflicting results about training and its relationship to client outcomes (Budge et al., 2013). One study comparing client outcomes between practicum students, predoctoral interns, and licensed professionals determined that there were no significant differences in client outcomes or premature termination between these three groups (Nyman et al., 2010). These researchers hypothesized that more advanced counselors would experience lower rates of premature termination among their clients (Nyman et al., 2010). Other studies have also shown that therapist experience does not necessarily lead to increased client outcomes (Budge et al., 2013; Lambert et al., 2003; Michael et al., 2005; Nyman et al., 2010;

Propst et al., 1994). In research by Budge et al. (2013) client outcomes were analyzed between four training levels of counselors. These four groups were 1) beginning practicum, 2) advanced practicum, 3) intern, and 4) psychologist (Budge et al., 2013). This study showed that the intern group, and not the hypothesized psychologist group, actually yielded the greatest amount of change in the clients they worked with in terms of increased "life functioning" and reduction of symptoms (Budge et al., 2013, p. 150). These findings conflict with other studies that found that licensed providers are likely to obtain higher rates of client improvement than professionals in training (Driscoll et al., 2003). Studies have also shown that training level impacts the development of conceptualization skills (Eells et al., 2005; Mayfield et al, 1999) and attrition rates (Stein & Lambert, 1995). Overall, there are few studies that directly look at training level, experience, and client outcome (Lambert, 2005). The mixed results regarding competence level and clinical outcomes warrant continued exploration of how we understand counselor competence and whether or not the competence we measure is indicative of a counselor's ability to achieve results in therapy.

Studies have identified a disconnect between actual competence level and perceived competence level in counselors (Duncan & Reese, 2016). Previous research has shown "that individuals see themselves as more able than statistically probable" (Walfish et al., 2012, p. 639). Walfish et al. (2012) explains that this is due to self-assessment bias, which can be conceptualized as "an overly positive assessment of personal performance" and impacts individuals in many professional contexts, including psychotherapy (p. 639). Research by Reese et al. (2009a) demonstrated that counselors who received feedback on their performance received increased clinical outcomes, when compared with a no feedback condition. In this same study, the levels of self-efficacy for both the feedback and no feedback groups increased significantly,

despite the significant differences in the clinical outcomes their clients experienced (Reese et al., 2009a). This study calls into question the idea that increased counselor self-efficacy leads to increases in client gains.

This has numerous implications for the field of psychotherapy (Walfish et al., 2012). Clinicians often rely on clinical judgement and experience to determine if deterioration or progress has occurred, and there is often a disconnect between clinical experience and ability to make predictive judgements about clinical treatment progress (Hannan et al., 2005). This research highlights the importance of the use of other standardized measures to address this disconnect between perceived ability to facilitate change, ability to accurately predict treatment progression or deterioration, and clinical judgement/experience.

Counselor Training and Evaluation

There are many different methods that can be used to assess the abilities of a counselor in training's skill level. Common evaluation methods for trainee growth include practicum evaluations from supervisors, competence rating forms, and annual portfolios. Some counselor training programs evaluate their student performance through the assessment of dispositions that are relevant for success as a counselor, such as "commitment, openness, respect, integrity, and self-awareness" (Spurgeon et al., 2012, p. 103). Common methods of evaluation often involve the counselor reflecting about their counseling skills and emotional and cognitive responses to their clients but fail to adequately measure actual counseling microskills (Budge et al., 2013). Evaluations are often completed by a single clinical supervisor, and research has highlighted some of the issues associated with this method of evaluation. Ladany and Melincoff (1999) showed that the majority of supervisors, and also supervisees, refrain from expressing certain types of constructive feedback. In this study the feedback that was withheld was often related to

"negative reactions to the trainee's counseling and professional performance, supervisor personal issues, and negative reactions to the trainee's supervision performance" (Ladany & Melincoff, 1999, p. 167). This tendency for feedback to be withheld is a significant problem associated with evaluations from a single supervisor.

There are other methods, aside from supervisory evaluations, that assess competence. Counselor competence has been measured in previous studies through the use of self-efficacy measurement (Budge et al., 2013). Previous research by Kozina et al. (2010) identified these competence domains as "micro skills, process, handling difficult client behaviors, cultural competence, and awareness of values" (2010, p. 117). The aforementioned study focused on understanding these competence domains through the measurement of self-efficacy beliefs (Kozina et al., 2010). Research in multiple contexts has shown a link between self-efficacy beliefs and performance in a given area (Kozina et al., 2010). There are many self-report instruments that have been created to assess self-efficacy measures in counselors in training (Kozina et al., 2010). A downside to this method of evaluation is the inherent potential for bias that self-report measures are prone to (Kozina et al., 2010). Despite this, Kozina et al. (2010), argue for the utility of these self-efficacy measures on a regular basis to provide empirical data that can be used to measure counselor trainee development. This focus fails to take into consideration the self-assessment bias that might impact a counselor's ability to identify therapeutic gains and deteriorations in their clients (Walfish et al., 2012).

Current methods of evaluations rarely involve the direct assessment of client outcomes as a measure of counselor progress. Using client outcome as a way to measure or improve supervisory effectiveness or counselor effectiveness is rarely used by supervisors (Lambert & Hawkins, 2001). It is surprising that clinical outcomes are not commonly used to assess

counselor competence due to the common use of clinical outcomes to measure effectiveness in general psychotherapy research (Lambert & Hawkins, 2001). This is also surprising given the "outcome-oriented evaluation focus" of the APA Guidelines and Principles for Accreditation of Programs in Professional Psychology (APA, 2009a, p. 4). These guidelines state that there will be "great emphasis on the outcomes or products of a program's training efforts" (APA, 2009a, p. 4). Despite this, Fauth et al. (2007) argue that "most current psychotherapy training methods do not durably improve the effectiveness of psychotherapy as delivered by trainees" (p. 388). As a profession, it is imperative that we are able to adequately assess the effectiveness of psychotherapy training programs and intervene when necessary to improve the quality of training.

A handful of previous studies have used clinical outcomes to compare performance among counselor trainees (eg: Iberg, 1991). Clinical outcomes can be measured using a number of psychometrically sound measures (eg: OQ-45.2, ORS). Client outcomes can be tracked using assessment tools such as the Outcome Questionnaire 45 (OQ-45.2), Session Rating Scale (SRS), Outcome Rating Scale (ORS), and Patient Health Questionnaire (PHQ full). Client outcomes from the OQ-45.2 were analyzed for the present study. The OQ-45.2 is a self-report measure that contains 45 items related to psychological distress in the areas of "symptomatic distress, interpersonal problems, and social role dysfunction" (Nyman et al., 2010, p. 205). The measure also produces a global score of overall functioning (Nyman et al., 2010). This assessment tool is commonly used to measure therapeutic changes over time and has established acceptable levels of reliability and validity (Nyman et al., 2010). Previous research looking at the use of the signal-alarm codes on the OQ-45.2 has shown clinically significant differences in treatment outcomes when a counselor is given feedback that their client is deteriorating (Lambert et al., 2003).

Moreover, the self-assessment bias implies that "psychotherapists will likely overestimate their skill and positive client outcomes and underestimate client deterioration rates" (Walfish et al., 2012, p. 643). This brings up the importance of the use of empirical assessment measures to guide the treatment process (Walfish et al., 2012).

One common therapeutic alliance and outcome measuring system, called the Partners for Change Outcome Management System (PCOMS), involves the use of the SRS and ORS (Reese et al., 2009b). The ORS includes four items that are used to measure clinical outcomes; the SRS includes four items used to measure the therapeutic alliance (Reese et al., 2009b). These measures are administered each session and are brief in nature (Reese et al., 2009b). The ORS was created by modifying items from the OQ-45.2 (Reese et al., 2009b). Research on the PCOMS system has indicated increased clinical outcomes and a reduction in premature terminations of therapy (Reese et al., 2009b). The APA Division 29 Task Force on Empirically Supported Relationships speaks of the importance of regular measurement of "patients' responses to the therapy relationship and ongoing treatment" because of "increased opportunities to repair alliance ruptures, to improve the relationship, to modify technical strategies, and to avoid premature termination" (Ackerman et al., 2001, p. 496).

Purpose of Study

The purpose of this study was to add to the preexisting body of literature about the use of clinical outcomes as a way of measuring progress and skill differences in developing counselor trainees. I hoped to increase awareness of the field's lack of usage of clinical outcomes for training purposes and to add to the literature base about the connection between clinical outcomes and training/competency level. I hypothesized that counselors in training would experience gains in their ability to facilitate therapeutic outcomes as their education level

increased, master's counselors would have higher percentages of reliable improvement among their clients when compared with bachelor's counselors, and that trainees would vary in their percentages of reliable improvement among the clients they served. This study examined the rates of growth in counselor competence among early training clinicians and how these rates varied among training clinicians.

Previous research (Reese et al., 2009; Walfish et al., 2012) has yielded concerning results about the ability of counselor professionals to determine if clients are progressing or deteriorating in therapy without the use of assessment measures. This has implications for the effectiveness of psychotherapy that should be addressed by counselor training programs. This research should be used to inform future studies addressing a counselor's ability to perform the required duties of a counselor professional. The purpose of this study was to improve our understanding of the rates of change among counselors in training with the use of routine outcome monitoring. This could lead to a call for changes in how we supervisee trainees and the importance of incorporating routine outcome monitoring in clinical settings.

Significance to the Field

This dissertation provided additional information about the development of counselors in training across time. There is a lack of research directly assessing counselor in training progression across education levels and compared to peers. This study has filled gaps in knowledge related to the variance in counselor competency level from student to student in the same year in a program. This study was significant to the field of counseling psychology, because it increased our understanding of counselor development and how clinical outcome measures can be used to measure clinician professional development over time. The current body of research has yielded conflicting results related to performance of interns versus psychologists

and practicum students when these groups are compared (eg: Budge et al., 2013; Lambert et al., 2003; Michael et al., 2005; Nyman et al., 2010; Propst et al., 1994). The assumption that increased training leads to increases in client outcomes needs greater evidence/support and was assessed through this study.

CHAPTER II

LITERATURE REVIEW

Introduction

The purpose of this study was to better understand the ability of counselors in training (CIT) to facilitate growth and change in the clients they work with across their graduate school careers. Few studies have directly looked at the development of CIT by comparing the psychotherapy outcome measures of the clients they see over time. Clinical outcomes can be conceptualized as a "valued and changed state", such as decreases in anxiety levels or increases in social support in clinical populations (Howard et al., 1996, p. 1059). These outcomes act as a measure of gains clients experience in treatment. The present study analyzed how change happens over time for CIT early on in their doctoral training. Due to the few studies on this topic, I examined patterns of change and how students' progress over time compared with peer progress.

This literature review provides a summary of the body of relevant literature related to how counselor clinical competency is measured, developmental models of supervision, how psychotherapy outcomes are measured in doctoral trainees, the current climate of counselor training and how trainee progress is measured, theoretical models of supervision, and areas not addressed in previous literature.

Professional Competence

It is imperative that counselors in training develop competence in delivering clinical interventions and assessments throughout their training programs. The American Psychological Association has identified six categories of competency and sixteen specific benchmark areas

within those categories that indicate professional competence in the field of psychology. These six categories include: 1) Professionalism, 2) Relational, 3) Science, 4) Application, 5) Education, and 6) Systems (American Psychological Association [APA], 2011). The specific benchmark areas included in these categories are: 1) Professional Values and Attitudes, 2) Individual and Cultural Diversity, 3) Ethical Legal Standards and Policy, 4) Reflective Practice/Self-Assessment/Self-Care, 5) Relationships, 6) Scientific Knowledge and Methods, 7) Research/Evaluation, 8) Evidence-Based Practice, 9) Assessment, 10) Intervention, 11) Consultation, 12) Teaching, 13) Supervision, 14) Interdisciplinary Systems, 15) Management-Administration, and 16) Advocacy (APA, 2011).

It is expected and assumed that graduate trainees will work towards professional development in each of these areas as they progress through their doctoral program. Despite this assumption, there has been little "research demonstrating an association between within-program training procedures and the subsequent quality of therapy outcomes" (Stein & Lambert, 1995, p. 183). It is difficult to obtain data on how training programs ensure that these quality standards are met within their trainees.

Professional Development Models

Professional development and growth can be measured in a number of ways and vary across disciplines. One common method of assessing professional development in counselor training programs is the use of portfolio evaluations, where a trainee provides examples and descriptions of their professional work (Baltimore et al., 1996). This method is used in other professional disciplines as well. Portfolios act as a barometer of development across time (Baltimore et al., 1996). This way of measuring professional growth allows students to reflect on one's performance, which is a crucial component of professional education (Jensen & Saylor,

1994). Other measures of professional growth include supervisory evaluations, class performance, and involvement in professional organizations/groups/research teams.

Developmental Models

Researchers have come to acknowledge a developmental approach to the training of future psychologists for professional practice (Stoltenberg, 2005). Training should be appropriately attuned to the trainees' developmental level (Stoltenberg, 2005). This approach indicates that professionals in training grow in competence as they progress in their professional development as counselors. Stoltenberg et al. (1998) developed a model describing three levels of professional development called the Integrated Developmental Model of supervision (IDM). The first level is reserved for beginning counselors and is characterized by anxiety, negative beliefs about one's ability to facilitate change in clinical populations, and a focus on self (Stoltenberg, 2005). This self-focus involves heavy consideration of the trainees "own behavior (trying to implement skills), thoughts (figuring out what to do next, trying to understand the client), and emotions (anxiety, frustration, hopefulness)" (Stoltenberg, 2005, p. 859). The second level is associated with more advanced training clinicians and involves a greater focus on client and client experiences (Stoltenberg, 2005). Stoltenberg argues that trainees at this level can experience positive or negative effects related to this shift of focus from self to client. This can lead to increased ability to conceptualize and apply appropriate interventions in some cases, or "in confusion and negative emotions with diminished effectiveness and negatively affecting motivation and autonomy" for others (Stoltenberg, 2005, p. 859). The third level of the IDM of supervision is associated with a balance between the focus on self and the focus on client, which allows the counselor to adequately monitor client's experiences, as well as their own experiences, in the therapy room (Stoltenberg, 2005). Level three is characterized by greater

independence, confidence, knowledge, and awareness (Stoltenberg, 2005). Developmental growth is characterized by the acquisition of technical counseling skills, professional knowledge and understanding, critical thinking, and the social skills needed to perform the duties expected of a psychotherapist.

Personal Growth Models

Other models of trainee development describe personal and professional characteristics and competencies that counselors in training should possess (Bland, 2018). There has been an increased focus in recent years on "developing the helper as a person in addition to honing technical proficiency" (Bland, 2018, p. 6). Despite this, there is still little research regarding how to concretely evaluate these descriptive models that detail the skills and personal characteristics that counseling professionals should have (Bland, 2018). Assessment measures, like Hart and Hart's *Spiritual Assessment Matrix* (SAM), have been used for evaluating components of personal development (Bland, 2018). Hart and Hart's model is based upon the "*Four Virtues",* which are presence, heart, wisdom, and creation (Hart, 2014). Research by Bland (2018) involved using the SAM as a pre and posttest to measure professional growth over time.

Clinical Outcomes and Training Level

In professional disciples, individuals often assume that increased training leads to increased job performance. In clinical work, supervisee development is exhibited through "quantitative changes within each developmental level and reflecting qualitative changes between levels" (Stoltenberg, 2005, p. 859). There is evidence to support that trainees differ based on developmental level (e.g.: Stoltenberg et al., 1994), but there has also been conflicting literature that only found differences between beginning and internship trainees (e.g.: Holloway,

1992). Other studies have shown differences between mental health providers with differing licenses types and education levels. The results from Seligman (1995) indicated that psychologists, psychiatrists, and social workers had clients with significantly more improvements than did marriage therapists. Although the common assumption is that counselors with more training should yield greater clinical outcomes, the literature does not fully support this notion (Budge et al., 2013). Budge et al. (2013) argue that there is "no consistent evidence that training leads to better client outcomes" (p. 151). Some studies have shown that there are differences among outcomes of counselors based upon training level (Budge et al., 2013; Driscoll et al., 2003), whereas other studies have shown no differences or only a modest effect size among outcomes from counselors of varying training level (Lambert et al., 2003; Michael et al., 2005; Nyman et al., 2010; Propst et al., 1994; Stein & Lambert, 1995). Research has also shown differences in outcomes that are contradictory to the assumption that increased experience leads to increases in clinical outcomes (Daniel, 2006).

Previous outcome studies have placed heavy focus on the differences between counselors of varying training levels, but few studies have looked at within-group differences among counselors in training. Other components of treatment related to outcomes, such as conceptualization skills (Eells et al., 2005; Mayfield et al., 1999) and attrition rates (Stein & Lambert, 1995) have been associated with training level. These differences indicate a lack of clarity in the notion that increased training in fact leads to increases in clinical outcomes. The conflicting research in this area highlights a need for increased understanding of the development of counselors in training. The aforementioned study indicates that additional research is needed in this area.

Trainee Evaluation Methods

There are many methods that clinical and counseling psychology doctoral programs use to evaluate clinical competence and the growth areas of trainees. Clinical competence can also be assessed through case conceptualization presentations, live supervision, self-efficacy questionnaires, and assessment measures. Often these methods involve the evaluation of one individual clinical supervisor that might only meet with the supervisee for one hour per week or a group supervisor who might meet with the clinician in a group setting. Clinical supervisors might focus more on trainee development versus clinical outcomes (Freitas, 2002). A less commonly used method of evaluating counselor performance or competence is through the use of clinical outcome measures. A handful of studies have looked at trainees' ability to produce clinical outcomes (Budge et al., 2013; Nyman et al., 2010). Despite the benefits that could be obtained through regularly assessing client outcomes in supervision or training, there has been greater focus placed "on the acquisition of both interpersonal and technical skills in the developing therapist, the therapist-supervisor relationship, and models of supervisory development" instead of on the actual "consequences of practice on patient outcome" (Lambert, & Hawkins, 2001, p. 131). This lack of focus on clinical outcomes among trainees has implications for clinical practice. Providing relief or decreases in symptomology is a cornerstone of clinical work, but current evaluation methods for clinicians in training and relevant research does not heavily focus on clinical outcomes.

Researchers have highlighted an important discrepancy between psychotherapy literature and psychotherapy training literature. Psychotherapy literature determines the competency level of the therapist by "directly assessing the degree to which their clients improved" e.g. outcomes, whereas psychotherapy training literature assesses trainee competence using constructs such as

self-reflections of counseling experiences or assessments of beginning counseling skills (Budge et al., 2013, p. 151). These methods fail to measure true performance as a counselor (Budge et al., 2013).

A literature review was conducted by Freitas (2002) to explore the impact of supervision on clinical outcomes. This literature review discusses the research done by Dodenhoff (1981), which highlights that supervisors might be more focused on supervisee growth and performance versus client change (Freitas, 2002). This provides more data supporting the lack of focus on actual client change, which is ultimately one of the key purposes of clinical work. In contrast Freitas (2002) provides a summary of research by Iberg (1991) that provides a unique outcome-oriented process, involving a statistical procedure, to inform supervision. In Iberg's study, the clinical outcomes of the clients of graduate trainees were compared.

Measuring Clinical Outcomes

Tracking client outcomes for CITs might provide a more direct observation of clinical growth and competence than by supervisory feedback or self-reflection alone. Outcome measurements involve systematically measuring client functioning on a regular basis using measurement tools. The therapy outcome is the amount of change a client experienced while receiving therapy services. Client outcomes can be tracked using assessment tools such as the Outcome Questionnaire 45 (OQ-45.2), Session Rating Scale (SRS), Outcome Rating Scale (ORS), and Patient Health Questionnaire (PHQ full). In the present study I examined client outcomes from the OQ-45.2. The OQ-45.2 is a self-report measure that contains 45 items related to psychological distress in the areas of "symptomatic distress, interpersonal problems, and social role dysfunction" (Nyman et al., 2010, p. 205). The measure also produces a global score of overall functioning (Nyman et al., 2010). This assessment tool is commonly used to measure

therapeutic changes over time and has established acceptable levels of reliability and validity (Nyman et al., 2010). Few studies have looked at a counselor's training status, experience, and the outcomes that their clients experience through therapy (Lambert, 2005). Studies looking at these variables have also used poor research methods and design, had low internal and external validity, and failed to address some confounding variables when addressing their research questions (Lambert, 2005).

Variance in Outcomes

There are downsides to using therapeutic outcomes as a professional development measure because clinical outcomes and the therapy process are highly impacted by client factors (Gordon, 2012). Client factors can include severity of symptoms (Daniel, 2006), what stage of change the client is at, access to resources, or commitment to treatment (Wampold & Budge, 2012). Maslow's hierarchy of needs (1943) provides a theoretical framework to conceptualize client factors. Maslow's theory hypothesized that there are four different levels of needs for an individual and that the individual must accomplish a level of need before moving on to the higher need state. The needs described in this theory are first physiological needs, followed by safety, love, esteem, and lastly self-actualization. James (2016) assessed the relationship between client barriers to treatment and Maslow's hierarchy of needs. Findings from this study showed that 72.27% of clients reported that one or more of their barriers to treatment was related to a basic need (James, 2016).

Client factors such as severity of distress and type of symptom presentation also mediate clinical outcomes and duration of therapy. Facets of interpersonal functioning such as "capacity to form attachments" and "openness to therapeutic interventions" influence a client's ability to experience therapeutic growth, with more interpersonal distress leading to difficulty achieving

outcomes (Steenbarger, 1994, p. 113). Psychological distress has been shown to have a differing relationship with outcome. Mohr et al. (1990) identified clients with higher levels of psychological distress as having more motivation to engage in treatment. In this study distress had a linear relationship with clinical outcomes, with higher levels of distress being associated with higher outcome ratings (Mohr et al., 1990).

Client factors were also included in Lambert's common factors of psychotherapy. Common factors are the "dimensions of the treatment setting (therapist, therapy, client) that are not specific to any particular technique" (Lambert, 2005, p. 856). Lambert broke down the common factors of counseling to include 40% of therapeutic outcomes relating to extra therapeutic client factors (1986). Lambert (1992) estimated that 30% of the variance in outcomes was due to quality of the therapeutic relationship, 15% to technical factors, and 15% to expectancy. Common factors often impact outcomes and can include components of treatment such as the therapeutic alliance, positive regard, empathy, and authenticity from the therapist to client (Wampold & Budge, 2012). In previous research there have been studies that look at the common factors and what specific treatment ingredients lead to change among clinical populations (Wampold & Budge, 2012). Common factors have been conceptualized through different theoretical models, including the contextual model. The contextual model proposes that there are three different paths for therapeutic healing to occur. These pathways include the therapeutic relationship, expectations for treatment, and specific treatment ingredients (Wampold, 2015).

Despite the role of common factors in achieving outcomes, we also know that therapist effects exist in clinical work. Therapist effects occur when certain "therapists consistently achieve better outcomes with their patients than other therapists, regardless of the nature of the

patients or the treatment delivered" (Wampold, 2015, p. 274). Modest effect sizes have been found for therapist effects in clinical outcomes (Wampold, 2015). Meta-analyses have shown therapist differences in outcome to be at a medium effect size (Crits-Cristoph et al., 1991). Moreover, therapist effects are larger in real world settings when compared to clinical trials (Wampold, 2015). A meta-analysis conducted by Crits-Christoph et al. (1991) indicated that experience and the use of a treatment manual accounted for many of these therapist effects.

Other therapist variables that might impact treatment outcomes include therapist well-being, therapist burnout, interpersonal style, or attention to the therapeutic alliance (Kraus et al., 2011). Kraus et al. (2011) found that effect sizes for therapists in their study of treatment in naturalistic studies varied from small to large. For certain therapists, their effect sizes for treatment were large, despite treating a wide variety of presenting concerns, and other therapists had consistently negative treatment effect sizes, across a variety of presenting concerns. Kraus et al. (2011) argue for weekly monitoring of client outcomes among graduate trainees to facilitate professional development.

As a discipline, it is important that psychological researchers are addressing how therapist effects might impact the outcomes, especially in randomized control trials (Crits-Cristoph et al., 1991). Differences among therapists' effectiveness might be particularly salient among outcome studies that use graduate trainee therapists because of the variability in skill level in this group (Crits-Cristoph et al., 1991). Therapy approach and length of treatment are other components that can impact variability in clinical outcomes (Crits-Cristoph et al., 1991). The more structured the approach is, the less room for deviation in treatment. This difference in structured versus less structured approaches could mean additions or subtractions to the treatment being provided (Crits-Cristoph et al., 1991). Another important consideration for

outcome studies is client diagnosis, due to the variation in likelihood of receiving improvement from treatment due to severity of diagnosis (Crits-Cristoph et al., 1991). These studies provide further data to support the importance of understanding the effectiveness of therapists in training and providing interventions, when necessary, to help improve clinical performance in student trainees. Having the expertise and ability to facilitate growth and change within clinical populations is central to clinical practice. Despite this, there has been little application of clinical outcomes to addressing counselor development.

Clinical Significance and Reliable Change

Clinical significance provides a way of assessing the level of growth and change that has occurred for clients in treatment. In research, clinically significant change is indicated by two benchmarks: 1) a client must transition from a dysfunctional to functional score range and 2) the level of change "observed must meet a statistically significance criterion by being reliable, that is, greater than the measurement error of the outcome instrument used" (Hansen et al., 2002, p. 330). In order for a treatment to cause clinically significant change, the treatment modality must "meet standards of efficacy set by consumers, clinicians, and researchers" (Jacobson & Truax, 1991, p. 12). This conceptualization of clinical outcomes allows us to better understand the utility and significance of therapeutic change.

It is important to understand whether change has occurred in treatment, as well as how much change the client has experienced (Jacobson & Truax, 1991). This can be better understood through the reliable change index (RCI), proposed by Jacobson and Truax (1991). The RCI has a cutoff score of 1.96, meaning that if the RC score is above a 1.96 clinical change is likely to have occurred. The RCI is useful in determining whether true change occurred or if measurement instruments were not reliable (Jacobson & Truax, 1991). Reliable change happens when there are

differences in observed scores that are greater than differences that would be expected based on the standard error of measurement (Erekson et al., 2015). On the OQ-45.2, there has to be a 14-point observed score difference for reliable change to have occurred (Erekson et al., 2015). The RCI is a less stringent method of assessing clinical change than clinical significance, as there is not a requirement that the score starts in a dysfunctional score range and moves to a functional range.

Previous studies have provided varying estimates of clients who experienced reliable change. Seligman (1995) discusses results of a study in which individuals who had participated in therapy provided responses to items in the three subscales of "Specific improvement", "Satisfaction", and "Global improvement" (p. 6). 54% of these participants responded that their treatment "made things a lot better" (Seligman, 1995, p. 9). Daniel (2006) reported that 57% of clients achieved reliable improvement in the sample of counseling professionals from different training levels used, but percentages of reliable change varied based upon training level. Some studies involving trainees or training clinics have somewhat lower percentages of reliably improved clients. Samstag and Norlander (2019) reported that 40% of clients experienced reliable improvement while working with a counselor in training, Carr et al. (2017) indicated that 12.2% of clients achieved reliable improvement, and 29% achieved clinically significant change. The proposed study was conducted to provide greater clarity regarding the expectations for levels of change in clients served by counselors in training at the doctoral level.

Models of Clinical Change

In practice, clinical change often does not occur linearly (Kadera et al., 1996). Because of this, for the present study I have analyzed session by session data for clients to assess at what point reliable change in a positive direction (reliable improvement) has occurred. There are two

commonly referred to models of understanding change in clinical populations: the Good Enough Level (GEL) model and Dose-Effect model. There is significant literature on the dose-effect relationship of therapy. Through this model, change can be measured by looking at the differences in functioning across a total number of sessions attended or by assessing session by session increases in wellbeing and functioning. The dose-effect model of counseling relates to the amount of change clients experience per session or per "dose" of treatment (Baldwin et al., 2009). Wolgast et al. (2003) found that "24% of clients can be expected to achieve RC by 4 sessions, 51% will achieve RC by 10 sessions, and 75% will achieve RC after 24 sessions" (p. 22). Other dose-effect research has shown that approximately 50% of clients should experience symptom improvement between session 8 and 13 (Howard et al., 1986). Clients with more complex presentations require more doses of treatment in order to experience improvements (Howard et al., 1986). This study was conducted to provide further data regarding the number of sessions until reliable improvement occurs in a sample of counselors in training.

Another proposed model for understanding client change, the GEL model, assumes that clients experience different rates of change (Barkham et al., 2006). The number of sessions until the good enough level of change occurs varies based on presenting concern, treatment type, and client factors (Barkham et al., 2006). This good enough level is not tied to reliable or clinically significant improvement but is instead associated with a decision to discontinue treatment (Barkham et al., 2006). Some individuals might opt out of treatment once they have experienced minor improvements, whereas some might continue until accomplishing their treatment goals. This model suggests that clients experience therapeutic improvements in a linear direction from session to session until reaching the GEL (Barkham et al., 2006). Once reaching this level, the

course of treatment is likely to be modified to focus on a new presenting concern, or the therapy relationship will be terminated (Barkham et al., 2006).

The timeliness in which clients experience therapeutic change is important, due to the broader implications for organizations and overall public health (Erekson, 2015). When clients experience therapeutic growth more efficiently, there are more resources at an agency to be distributed to other clients. Research has shown "that higher session frequency increases the efficiency of psychotherapy in clinically significant ways, decreasing length of patient suffering and possibly requiring fewer institutional resources" (Erekson, 2015).

Summary of Research

As discussed in previous research, there has been a general lack of focus on the ability of counselors in training to facilitate growth and change in the clients they work with. This has significant implications for the field of counseling psychology, as providing treatment and facilitating outcomes are some of the more crucial components of this discipline. Previous research and training has also been based on the assumption that increased training leads to increases in competence and ability to facilitate therapeutic outcomes, which has not fully been supported by literature. The present study was completed in the hopes of adding to the preexisting literature regarding counselors in training and ability to facilitate growth and change in clinical populations.

CHAPTER III

METHODS

Procedures

The participants in this study included clients who received services from the Counseling and Assessment Clinic (CAC) in Bryan, Texas. The CAC was a doctoral training clinic associated with the Counseling Psychology and School Psychology doctoral programs at Texas A&M University (TAMU). This community mental health clinic served clients from diverse cultural/socioeconomic/regional backgrounds with varying presenting concerns. The CAC provided counseling and assessment services and charges based on a sliding fee scale. Clients included in this study paid between $6.00 and $20.00 per session for counseling. The guidelines for the sliding fee scale at the CAC was based on the Federal poverty guidelines. According to this sliding fee scale, a family of 4 with an annual income of $98,400 would pay $20.00 per session, whereas a family of 4 with an annual income of $24,600 would pay $6.00 (Kleine-Kracht, 2019). Clients served at this clinic were referred from other community agencies in the area, such as HealthPoint, by previous clients, or from other community providers.

In order to receive services at this clinic, potential clients were required to complete a phone interview with one of the clinic's service coordinators. These coordinators were doctoral level graduate students. They conducted brief interviews to determine whether an individual was appropriate to receive services at the clinic. This clinic would deny services to clients who were referred for court mandated therapy, were involved in custody proceedings, or in a Child Protective Services (CPS) case. Clients were screened for psychotic symptoms (unmanaged by medications), significant substance abuse, and/or high-risk suicidal ideation. If clients denied the previously listed symptomology, then the service coordinators would assign the client to a clinic

waiting list that was processed by clinical supervisors at the agency. Clinical supervisors determined which clients would be assigned to which practicum counselors. Clients at the CAC were seen for an initial two-hour long intake appointment with their assigned counselor, where a detailed clinical history was gathered. Practicum counselors then met with each client on a weekly basis. Length of treatment varied for clients at the CAC based upon client progress and severity of presenting concerns.

Practicum counselors at the CAC were first year doctoral students trained through the scientist-practitioner model, which emphasizes the importance of both clinical and research skills. This included both post bachelor's and post master's students. Many of the post bachelor's students had no previous experience providing counseling services. These clinicians received weekly individual and group supervision by a licensed psychologist affiliated with the counseling psychology doctoral program at TAMU. Clinicians at this site were often in the early stages of developing their therapeutic framework and provided treatment utilizing techniques from many different theoretical orientations.

A number of measures were used to assess the therapeutic process of clients at this clinic. These include the Outcome Questionnaire 45.2 (OQ-45.2; Lambert et al., 1996), Session Rating Scale (SRS; Miller & Duncan, 2000), and Outcome Rating Scale (ORS; Miller & Duncan, 2000). The OQ-45.2 was administered to clients on a tablet each week before their session. The average OQ-45.2 total score for the intake appointment at the CAC was 75.83, and the average OQ-45 total score for the last session at the CAC was 68.81. Scores of 64 and above on the OQ-45.2 indicated functioning that is in the clinical range. In this sample, current scores ranged from 0 to 166. OQ-45.2 scores for clients who attended only one session were not included in the average intake score. Practicum counselors had the opportunity to look at a printout of their client's

scores before they started each session. The SRS and ORS were both administered at the end of each session. These measures were routinely administered and collected. Student clinicians also completed their clinical notes using the Titanium note-taking/scheduling software.

Measures

The OQ-45.2 is a brief self-report measure that assesses levels of functioning and distress in clinical populations (Erekson et al., 2015). Clients are often administered the assessment prior to each session as a way of measuring therapeutic growth/progress. This measure assesses "four domains of functioning: symptoms of psychological disturbance (mainly depression and anxiety); interpersonal problems; social role functioning (e.g., problems at work or school); and quality of life (positive aspects of life satisfaction)" (Lambert, 2007, p. 2). Subscales for each of the four domains, as well as a "general mental health factor", which is the total score, are provided (Lambert, 2007, p. 2). The OQ-45.2 consists of 45 items in a 5-point Likert-type format. Total scores can range from 0 to 180 (Erekson et al., 2015). Higher scores indicate greater levels of distress, while lower scores indicate lower levels of distress and healthier functioning. The measure has proven reliability, validity, and the ability to measure change over time (Erekson et al., 2015, Lambert et al., 2001). Erekson et al. (2015) reported an internal consistency of .93. The test-retest reliability of the measure was .84, for a three-week interval (Erekson et al., 2015). It has also been normed on a variety of clinical populations (Erekson et al., 2015).

The OQ-45.2 provides alarm-signal color codes that are displayed at the top of the protocol (Lambert, 2007). These color codes provide clinicians with additional data about the client's potential for successful or unsuccessful treatment (Lambert, 2007). Alarm-signal color codes include red, blue, and yellow. Blue codes represent that a client is experiencing an increase

in functioning that indicates they "will leave treatment with a positive outcome and maintain gains for at least 1 year"; whereas yellow and red codes indicate that "therapy may be heading toward an unsuccessful conclusion" and different treatments should be considered (Lambert, 2007, p. 5). Research has shown that routinely monitoring client outcomes leads to increased quality of client care and decreased likelihood of client deterioration (Boswell et al., 2015). This has been indicated through increased effect sizes of therapeutic change (Kraus et al., 2011). Moreover, providing feedback to clients based upon outcome monitoring has yielded greater increases in client outcomes than treatment without feedback (Anker et al., 2009).

Reliable Change Index and Clinically Significant Change

The OQ-45.2 can be used to assess clinically significant change in client presentations, using the reliable change index (RCI) (Erekson et al., 2015). Requirements for both reliable and clinically significant change have been provided from previous research. The RCI (14 points or more on the OQ-45.2) was determined based upon the Jacobson and Truax (1991) model of statistically operationalized clinically significant change (Erekson et al., 2015). The RCI is used to determine whether the change that occurred in functioning is more significant than changes/fluctuations in symptom severity that might be indicative of poor measurement tools (Jacobson & Truax, 1991). Jacobson and Truax (1991) determined that when the "RC is greater than 1.96, it is unlikely that the posttest score is not reflecting real change." (p. 14). Clinically significant change occurs when there are changes in functioning that are greater than the RCI score, and the new score is in the normal functioning range. For the OQ-45.2, scores less than or equal to 63 are in the "normal" or nonclinical range of functioning, and scores greater than 63 fall in the clinical range of functioning (Erekson et al., 2015). In order for clinically significant change to occur, a client must have initial scores that are in the clinical range and ending scores

in the non-clinical range, as well as a reliable change score that is greater than the reliable change cutoff point for the measure.

Inclusion/Exclusion Criteria

For this study, I included clients who attended at least two sessions of therapy in my data analyses. Counselors at this clinic focused on gathering information to plan treatment during their first appointment and did not heavily focus on clinical interventions until the second session. Clients experience improvement on a continuum, dependent on the severity of their presenting concern. Studies have shown that clients with more severe presentations needed more intensive, longer lasting, or more frequently occurring treatment in order to experience clinically significant change (Kadera et al., 1996). Kadera et al. (1996) showed a range of 2 to 13 sessions were required for therapeutic improvements, with the average number of sessions attended to reach therapeutic change being 7. Other studies have assessed therapeutic outcome using a minimum of 3 sessions (e.g. Baldwin et al., 2009).

Participants

Clients

This study included all clients that were previously seen at the CAC and all previous counselors at the CAC, since 2007 when the clinic started administering the OQ-45.2. The dataset consisted of approximately 579 clients and 94 counselors. Client and counselor demographics can be found in Tables 1 and 2 respectively. Clients represented various diverse racial groups. Client age ranged from 18 to 94 years old. There was no documentation of employment status, education level, or socioeconomic status for the sample. Some clients were missing demographic information, and this was reflected in the unknown column in Table 1.

Clients had varying diagnoses and symptom severity, leading to differences in numbers of therapy sessions attended.

Clients attended an average of 16.2 total sessions (SD = 24.2), and 71.5% had one counselor while receiving treatment at the CAC (see Table 3). Treatment sessions were broken down into separate segments for clients who were seen by multiple counselors. There were a total of 579 participants in this study, and treatment was separated into a total of 859 segments. Most clients were only seen by one counselor, but some clients were transferred to new counselors as their previous counselors left the clinic. Each counselor saw a differing number of total clients while at the CAC. For post bachelor's counselors, the number of total clients ranged from 1-21, with an average of 10 total clients seen. For post master's counselors, the number of total clients ranged from 1-16, with an average of 8 total clients seen. Information regarding total clients seen per counselor can be found in Table 4.

Counselors

Counselors in this study had a variety of educational and training experiences. The CAC was the first practicum site for many counselors whose data was included in this research study. Some counselors came in with a master's degree and had significantly more counseling experience. For the purposes of this study, counselors were separated into post bachelor's and post master's level training groups. A breakdown of counselor participants' education level and gender can be found in Table 2. Unfortunately, aside from gender, additional counselor demographic information was not available.

Institutional IRB approval was obtained prior to conducting the present study. Two datasets were used for the purposes of this research project, the first was the outcome data set with OQ-45.2 data, and the second one was the titanium dataset that included client

demographics and transfer information. These two databases were merged for the subsequent analyses.

Table 1.

Demographic Characteristics of Client Participants

Demographic Information	Total Clients	
	n	%
Gender		
Male	174	30.1
Female	338	58.4
Transgender	2	0.3
Unknown	65	11.2
Age		
18 - 25	82	14.2
26 - 35	197	34.0
36 - 45	130	22.5
46 - 55	82	14.2
56 - 65	57	9.8
> 65	29	5.0
Unknown	2	0.3
Race		
American Indian	1	0.2
Asian	14	2.4
Biracial	15	2.6
Black	47	8.1
Hispanic	119	20.6
White	311	53.7
Pacific Islander	1	0.2
Unknown	71	12.3
Marital Status		
Divorced	62	10.7
Divorced & Living Together	1	0.2
Living Together	13	2.2
Married	129	22.2
Separated	32	5.5
Single	236	40.8
Unknown	96	16.6
Widowed	10	2.0

Note. N = 579. The average age of participants was 39.2 years old (SD = 13.6).

Table 2.

Demographic Characteristics of Counselor Participants

Demographic Information	Total Counselors	
	n	%
Gender		
Male	23	25.5
Female	69	74.5
Transgender	0	0.0
Education Level		
Bachelor's Degree	59	64.0
Master's Degree	33	36.0

Note. $N = 92$.

Table 3.

Session Information

Number of Counselors throughout Duration of Treatment	Number of Clients	
	n	%
1	414	71.5
2	107	18.5
3	32	5.5
4	15	2.6
5	2	0.4
6	6	1.0
8	1	0.2
10	1	0.2
11	1	0.2

Table 4.

Total Number of Clients per Counselor

Education Level	*M*	*Mdn*	Mode	Range	
				Lowest	Highest
Bachelor's Counselors	10	10	9	1	21
Master's Counselors	8	9	1	1	16

Research Questions/ Hypothesis

This study was conducted to answer the following research questions:

i. What is the average number of sessions until reliable improvement was reached on the OQ-45.2 for the first time for the entire sample and for each counselor?

   Hypothesis: The average number of sessions until reliable improvement was reached for the first time on the OQ-45.2 will be lower for post master's than post bachelor's counselors.

ii. What is the average percentage of clients per counselor and in the entire sample who achieve reliable improvement?

   Hypothesis: The average percentage of clients who achieve reliable improvement on the OQ-45.2 will be higher for post master's than post bachelor's counselors.

iii. Is there a significant relationship between counselor education level and client's reliable change status (reliably improved, unchanged, or deteriorated) when initial OQ-45.2 score is controlled?

1. Are there counselors that struggle or have lower rates of reliable change (reliable improvement) or many sessions before reliable change happens?
2. Are there counselors that are higher performing and have higher rates of reliable change (reliable improvement) or less sessions until reliable change occurs?

   Hypothesis: There is a significant relationship between education level and change status, with higher education level associated with positive change status.

Possible Follow Up Questions

iv. Does doctoral training level (semester at CAC-1st, 2nd, or 3rd) and number of sessions (amount of time) have an influence on whether or not a client experiences reliable change?

v. Are doctoral trainees improving in their ability to facilitate clinically significant change (based on RCI) during their first three semesters of doctoral practicum experience?

vi. What variables account for these differences in clients achieving reliable change? Does gender of counselor, counselor training level (semester at CAC/number of clinical hours), age of client/counselor, and number of

therapy sessions influence the probability of whether or not reliable change occurs?

- Gender of counselor
- Semester at the CAC/ number of clinical hours as a measure of counselor experience
- Number of therapy sessions until reliable change

Statistical Analysis

Descriptive data, including demographic variables for counselors and clients was gathered for each of the participants. This included information such as race, gender, marital status, and client age. OQ-45.2 data was analyzed for each client participant. The number of sessions until reliable improvement occurred was calculated for each client, as was the average number of sessions until reliable improvement occurred for each counselor participant. Descriptive information such as, mean, median, mode, and range of sessions until reliable change occurred, was provided for each counselor in the study. The average of this number was calculated as well to provide the average number of sessions until reliable improvement occurred for the entire sample. The sample was divided into post bachelor's and post master's level counselor groups to see if there were any differences in the average number of sessions it took for a client to receive reliable change based on counselor training level. The mean, median, mode, and range for the number of sessions attended to reach reliable change was calculated for each counselor group (post bachelor's and post master's) to examine possible group differences. These results provided crucial information about the nature and quality of trainee data, in addition to whether there are differences in counseling outcomes based on trainee education level.

Treatment was broken down into multiple treatment segments for clients that were transferred to multiple counselors. For clients that were transferred during treatment, their session numbers were started over again at 1 for each new counselor that they were assigned to. This was done to maintain the integrity of treatment segments for newly assigned transfer counselors. To address the second research question, the average percentage of clients per counselor who achieved reliable improvement, no reliable change, or deterioration was calculated as well as the average for the entire sample.

Lastly, to address the third research question, a chi-square test of independence was calculated to assess whether there was a significant association between counselor education level and change status. For this research question, the independent variable was counselor education level, and the dependent variable was reliable change status. Counselor education was a binary categorical variable and consisted of post bachelor's and post master's students. Reliable change status was a three-level categorical variable and included the levels of reliably improved, unchanged, and deteriorated. The null hypothesis stated that education level was not associated with reliable change status, and the alternative hypothesis stated that education level was associated with reliable change status.

Two robust ordered logistic regression models were ran to determine if counselor education level or initial OQ-45.2 score was significantly related to change status. The dependent variable for both models was change status, which was broken down into the three levels of reliably improved, unchanged, and deteriorated. For the first model, the independent control variable was initial OQ-45.2 score, and for the second model the independent variable of counselor education level was added. It was hypothesized that education level was statistically significantly related to change status, with post master's counselors having higher rates of

reliable change, when compared with post bachelor's counselors. It was also hypothesized that clients with lower initial OQ-45.2 scores were more likely to achieve reliable improvement. The results of this study could have strong implications for clinical training and counselor education.

CHAPTER IV

RESULTS

Preliminary Analyses

For this study, treatment was broken down into multiple segments for clients seen by multiple counselors. Because of this, there was a total of 833 treatment segments. A breakdown of baseline score categories and education level can be found in Table 5.

Table 5.

Breakdown of Baseline OQ-45.2 Scores for Client Participants

Education Level	Functional Range		Clinical Range		Total
	n	%	n	%	
Bachelor's Counselors	144	25.1	429	74.9	573
Master's Counselors	75	28.9	185	71.2	260
Total Sample	219		614		833

An independent samples t-test was conducted to ensure that there were no significant differences in baseline scores among clients assigned to post bachelor's and post master's counselors. Each of the assumptions associated with an independent samples t-test were examined. The first assumption of an independent samples t-test is that the dependent variable is measured at the continuous level. The second is that the independent variable should be made up of two separate categorical groups. The third assumption is that there is independence of observations. The fourth assumption is that there were not any significant outliers. There were only two outliers in the dataset for initial intake score. A Shapiro-Wilk test of normality was conducted to test the fifth assumption for normal distribution. This test indicated that intake

scores were not perfectly normally distributed (p= .007). Despite this, skewness of the dataset was close to 0 (-0.1), and kurtosis was close to 3 (2.6). A normal distribution has a skewness of 0 and kurtosis of 3, so this dataset was considered to be normal (Curran et al., 1996). Lastly Levene's test of homogeneity of variances was conducted in Stata to determine if the variances in intake score was equal among post bachelor's and post master's counselors. Levene's test of homogeneity of variances indicated that there was not a statistically significant difference *(*p= .6) in variance of intake score between post bachelor's and post master's counselors. A median test was also conducted to assess for differences among the median scores for post bachelor's and post master's counselors. The results showed that there were no significant differences between initial OQ-45.2 intake scores for post bachelor's and post master's counselors (p= .33).

An independent samples t-test was completed to assess differences in means among intake scores between post bachelor's and post master's counselors. Intake/baseline scores were the dependent variable for this test, and education level was the independent variable. The post bachelor's group consisted of 573 observations, and the post master's group consisted of 260 observations. Results of the t-test showed that the intake scores did not significantly differ across the two counselor groups, $t(831) = 1.4$, $p = .15$.

Statistical Analyses

Session Number Results

Research Question 1. What is the average number of sessions until reliable improvement was reached on the OQ-45.2 for the first time for the entire sample and for each counselor?

The average number of sessions until reliable improvement was reached on the OQ-45.2 for the first time for the entire sample was 5 sessions. Descriptive statistics, including the mean, median, mode, and range of sessions is provided for the entire sample in Table 6.

Table 6.

Descriptive Statistics of Number of Sessions to Achieve Reliable Change for the Entire Sample

Averages	*M*	*Mdn*	Mode	Range	
				Lowest	Highest
Combined Average	5.0	4.4	3.4	2	51
Post Bachelor's Average	4.9	4.2	3.0	2	34
Post Master's Average	5.2	4.8	4.3	2	51

Percentage of Reliable Improvement Results

Research Question 2. What is the average percentage of clients per counselor and in the entire sample who achieved reliable improvement?

The average percentage of clients who achieved reliable improvement, no change, or deterioration was calculated per counselor and for the entire sample. The average percentage of clients in the entire sample who achieved reliable improvement was calculated to be 50.1% (see Figure 1). Results broken down by education level are provided in Figures 2 and 3. Change status data broken down by counselor for the post bachelor's cohort was split in half, in the order of most to least reliably improved clients and is displayed in Figures 4 and 5. Data for the post master's cohort is displayed in order of most reliably improved clients to least in Figure 6. Table 7 shows the total number of reliably improved, unchanged, and deteriorated clients for the entire sample.

Figure 1.

Percentage Change Status for Entire Sample

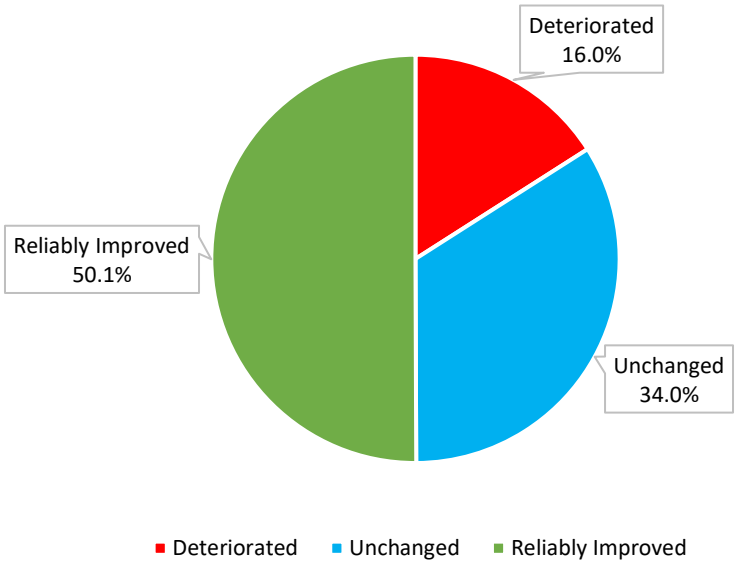

Figure 2.

Percentage Change Status for Post Bachelor's Counselors

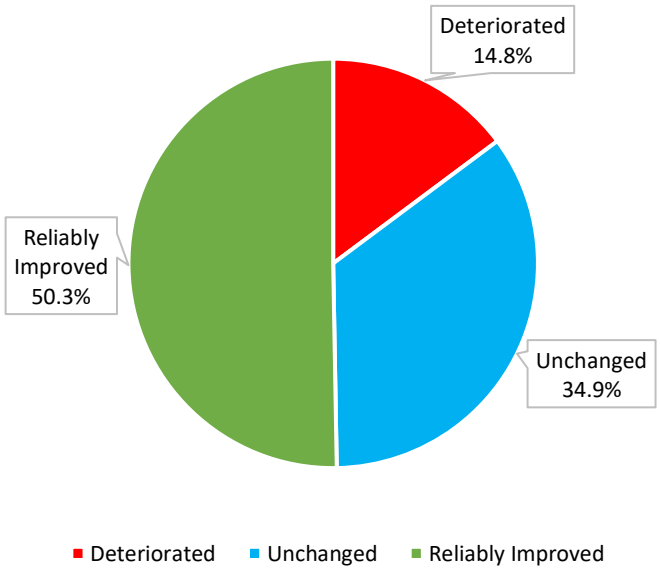

Figure 3.

Percentage Change Status for Post Master's Counselors

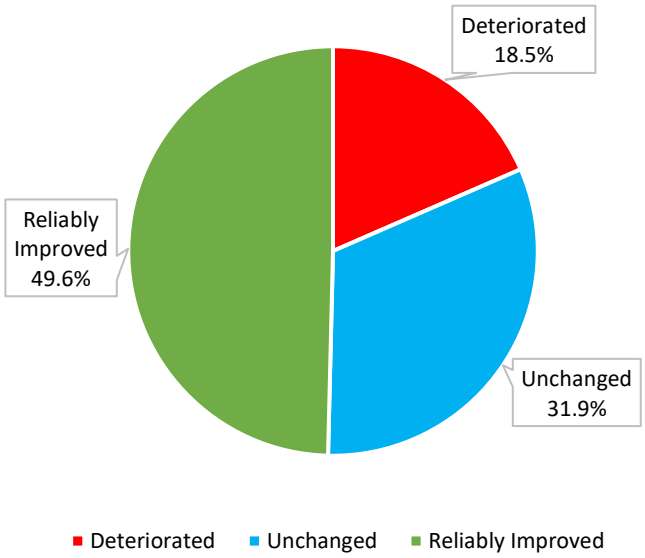

Figure 4.

Change Status for Highest Performing Post Bachelor's Counselors

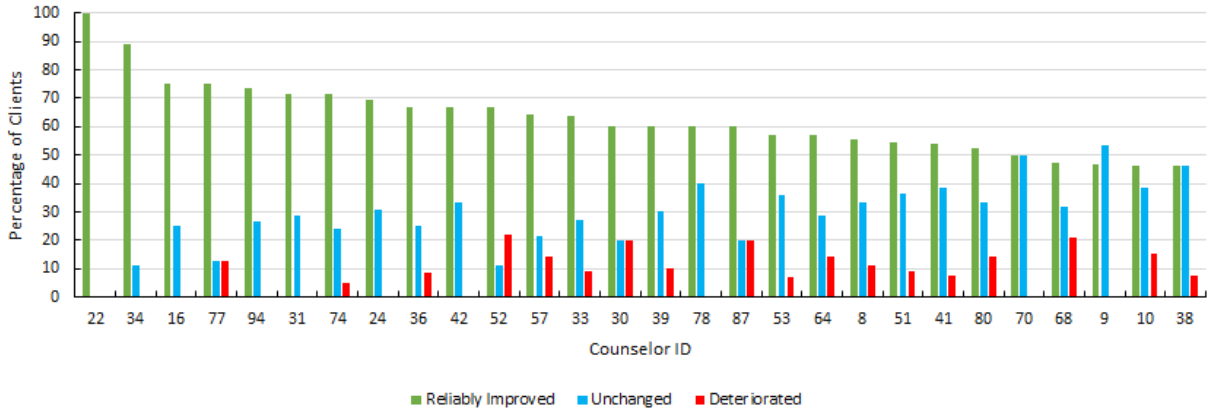

Figure 5.

Change Status for Lowest Performing Post Bachelor's Counselors

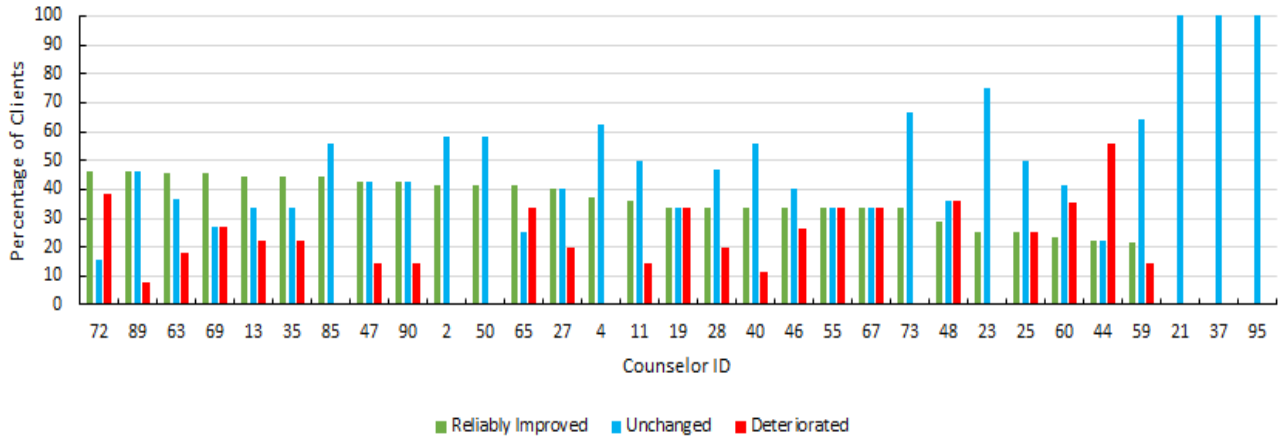

Figure 6.

Change status for All Post Master's Counselors

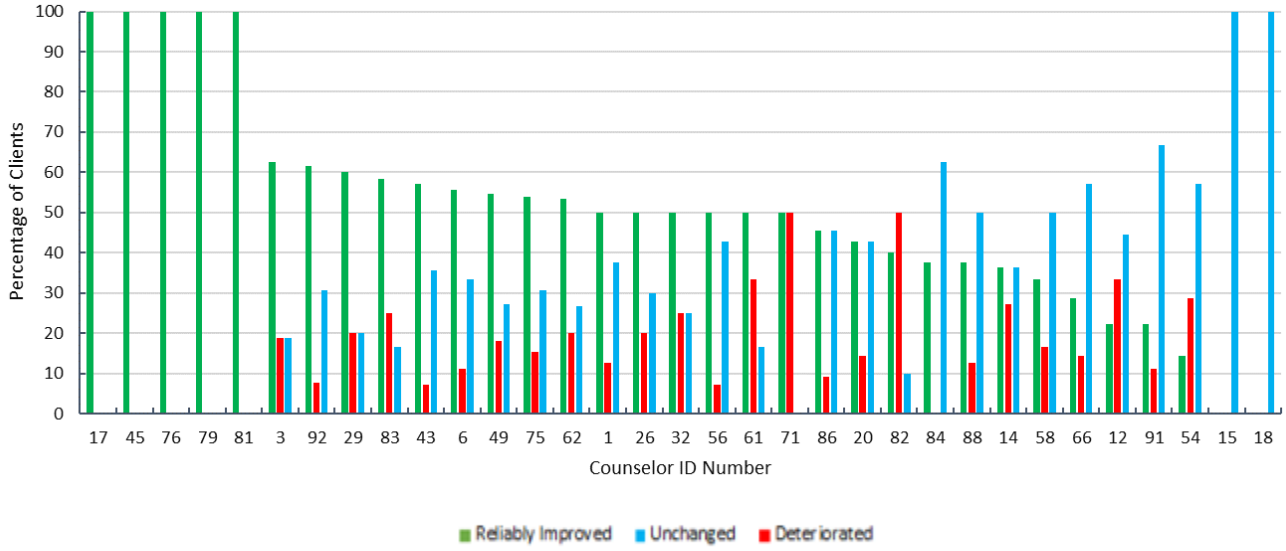

Table 7.

Total Number of Clients by Change Status

Education Level	Reliably Improved		Unchanged		Deteriorated	
	n	%	n	%	n	%
Post Bachelor's Counselors	288	50.3	200	34.9	85	14.8
Post Master's Counselors	129	49.6	83	31.9	48	18.5
Entire Sample Total	417	50.1	283	34.0	133	16.0

Education Level and Change Status Results

Research Question 3. Is there a significant relationship between counselor education level and client's reliable change status (reliably improved, unchanged, or deteriorated) when initial OQ-45.2 score is controlled?

a. Are there counselors that struggle or have lower rates of reliable change (reliable improvement) or many sessions before reliable change happens?
b. Do we have counselors that are higher performing and have higher rates of reliable change (reliable improvement) or less sessions until reliable change occurs?

To address the last research question, a Chi-square test of independence was used to determine whether a significant association existed between the variables of education level and change status. A Chi-square test is a non-parametric statistic that is used with nominal or ordinal variables (McHugh, 2013). The assumptions of a Chi-square test are as follows: 1. data is in the form of frequencies, 2. levels of each variable are mutually exclusive, 3. each subject is represented once and not over time, 4. there are independent groups, 5. there are two categorical variables, and 6. "the value of the cell *expecteds* should be 5 or more in at least 80% of the cells, and no cell should have an expected of less than one" (McHugh, 2013, p. 144). Expected values

were calculated by hand and can be found in Table 8. None of the assumptions of the Chi-square test were violated.

Table 8.

Cell Expected Values (Cell Chi-square Values)

Change Status	Education Level	
	Bachelor's	Master's
Reliably Improved	286.8 (<.01)	130.2 (<.1)
Unchanged	194.7 (0.2)	88.3 (0.3)
Deteriorated	91.5 (0.5)	41.5 (1.0)

It was hypothesized that education level would have a significant relationship with change status so that post master's counselors would have higher rates of reliably improved clients than post bachelor's counselors. There was not a significant relationship between education level and change status X^2 (2, N= 833) = 1.96, p = .376. Therefore, the null hypothesis was accepted. Because these data require a 2 x 3 table, it was not appropriate to report a Fisher's exact test (McHugh, 2013). A maximum likelihood ratio Chi-square test was also not reported, due to the sample size assumption being met for the Chi-square test of independence (McHugh, 2013). Despite the results from the Chi-Square Test of Independence, there were counselors with higher rates of reliable change and those with lower rates of reliable change. The highest and lowest performing counselors' data, broken down by education level, are reported in Tables 11, 12, 13, 14, and 15.

To control for baseline OQ-45.2 score, an ordered logistic regression analysis was conducted to determine whether a significant association existed between the independent variables of education level and baseline OQ-45.2 score and the dependent variable of change status. Dummy variables were created for education level and change status, as they were

categorical variables with multiple levels. The assumptions required for this regression were examined. An ordered logistic regression analysis can be conducted if the following assumptions are met: there must be only one dependent variable, the independent variables must be either continuous, categorical, or ordinal, there must be proportional odds/parallel lines (Liu, 2009), and there must be no multicollinearity. Each of these assumptions were tested, and none were violated.

A robust ordered logistic regression model, followed by the odds ratio, was then estimated in Stata using the dependent variable of change status and the independent control variable of baseline score. For the dependent variable, deterioration was coded as 0, unchanged status as 1, and reliably improved as 2. The Wald chi-square test, Wald $\chi^2(2) = 41.67$, $p < .001$, indicated that the model was statistically significant when compared to a null model with no predictors. This model had better fit than a model without any predictors in regards to predicting change status. According to this model, baseline OQ-45.2 score was significant ($p< .001$). This indicates that baseline OQ-45.2 score had a positive effect on predicting change status (OR=1.02). Based on the proportional odds ratio, if there was a one unit increase in baseline OQ-45.2 score, then the odds of a client being in a higher change status (meaning reliably improved) are 1.02 times greater than the combined middle and low categories (unchanged and deteriorated). Moreover, for a one unit increase in baseline OQ-45.2 score, the odds of a client being in the combined middle and high categories (unchanged and reliably improved) was 1.02 times higher than for the low category (deterioration). This was expected given that the other variables stay constant in this model.

Another robust ordered logistic regression model, followed by the odds ratio, was then estimated in Stata using the dependent variable of change status and the independent control

variable of baseline score and the independent variable of education level. For this second model, the Wald chi-square test, Wald $\chi^2(2) = 41.71$, $p < .001$, indicated that the model was statistically significant when compared to a null model with no predictors. This model had better fit than a model without any predictors when predicting change status. According to this model, baseline OQ-45.2 score was significant ($p< .001$), but education level ($p= .765$) was not. Baseline OQ-45.2 score had a significant positive effect on predicting change status (OR=1.02). Education level had neither a positive nor negative effect on predicting change status (OR=0.96). According to the proportional odds ratio, if there was a one unit increase in baseline OQ-45.2 score, then the odds of a client being in a higher change status (meaning reliably improved) were 1.02 times greater than the combined middle and low categories of unchanged and deteriorated. Moreover, for a one unit increase in baseline OQ-45.2 score, the odds of a client being in the combined middle and high categories (unchanged and reliably improved) was 1.02 times higher than for the low category (deterioration). This was expected given that the other variables stay constant in this model. These findings indicated that clients with higher OQ-45.2 scores were statistically significantly more likely to experience reliable improvement than individuals with lower baseline OQ-45.2 scores.

Table 9.

Data for the Top 10 Highest Performing Post Bachelor's Counselors

Counselor ID	Total Number of Clients	Percentage of Reliably Improved Clients	Average Number of Sessions until Reliable Improvement
22	1	100.0	4.0
34	9	88.9	4.9
16	4	75.0	7.7
77	8	75.0	5.0
94	15	73.3	7.3
31	7	71.4	5.0
74	21	71.4	3.8
24	13	69.2	6.9
36	12	66.7	4.5
42	9	66.7	4.0
Averages	9.9	75.8	5.3

Table 10.

Data for Bottom 10 Lowest Performing Post Bachelor's Counselors

Counselor ID	Total Number of Clients	Percentage of Reliably Improved Clients	Average Number of Sessions until Reliable Improvement
73	3	33.3	2.0
48	14	28.6	3.5
23	4	25.0	3.0
25	4	25.0	2.0
60	17	23.5	2.8
44	9	22.2	5.0
59	14	21.4	2.3
21	1	0	-
37	1	0	-
95	2	0	-
Averages	6.9	18.1	2.9

Note. Dashes (-) were used for counselors who had 0 clients who reached reliable improvement.

Table 11.

Data for the Top 10 Highest Performing Post Master's Counselors

Counselor ID	Total Number of Clients	Percentage of Reliably Improved Clients	Average Number of Sessions until Reliable Improvement
17	1	100.0	6.0
45	1	100.0	7.0
76	1	100.0	2.0
79	1	100.0	19.0
81	1	100.0	2.0
3	16	62.5	3.2
92	13	61.5	4.9
29	5	60.0	2.7
83	12	58.3	2.9
43	14	57.1	13.0
Averages	6.5	80.0	6.3

Table 12.

Data for Bottom 10 Lowest Performing Post Master's Counselors

Counselor ID	Total Number of Clients	Percentage of Reliably Improved Clients	Average Number of Sessions until Reliable Improvement
84	8	37.5	10.3
88	8	37.5	5.7
14	11	36.4	3.3
58	6	33.3	4.5
66	7	28.6	8.5
12	9	22.2	5
91	9	22.2	2.5
54	7	14.3	3
15	1	0	-
18	1	0	-
Averages	6.7	23.2	5.3

Note. Dashes (-) were used for counselors who had 0 clients who reached reliable improvement.

Table 13.

Data for Highest and Lowest Performing Counselors

Counselor Education	Average Number of Clients	Average % of Reliably Improved Clients	Average Number of Sessions until Reliable Improvement
Top 10 Post Bachelor's	9.9	75.8	5.3
Bottom 10 Post Bachelor's	6.9	18.1	2.9
Top 10 Post Master's	6.5	80.0	6.3
Bottom 10 Post Master's	6.7	23.2	5.3

CHAPTER V

DISCUSSION AND CONCLUSION

The purpose of this study was to add to the current body of literature regarding within-group differences among counselors in training in their ability to facilitate clinical outcomes. Few studies have looked at the differences in client outcomes among counseling students at different developmental levels. Although some studies have found differences in clinical outcomes among trainee or professional developmental level (e.g.: Budge et al., 2013; Driscoll et al., 2003, Holloway, 1992, Stoltenberg et al., 1994), other studies have failed to replicate similar findings (Lambert et al., 2003; Michael et al., 2005; Nyman et al., 2010; Propst et al., 1994; Stein & Lambert, 1995). Overall, the results of this study indicated that there was not a significant relationship between counselor education level and therapeutic outcomes of clients seen in a sample of doctoral student counselors. These findings provided greater evidence that contradicts the common assumption that increased training leads to increased clinical outcomes. A discussion of the findings broken down by research question has been provided below.

Session Number Results

The average number of sessions attended until reliable improvement occurred was 5 for the entire sample. For bachelor's counselors the average number of sessions attended until reliable improvement occurred was 4.9 and for master's counselors it was 5.2. These numbers were consistent with previous dose-effect model literature claiming that approximately 50% of clients experience symptom improvement by session 8 to 13 (Howard et al., 1986). In other samples, research has shown that after about 4 sessions a quarter of clients will experience reliable improvement, and after 10 sessions half of clients will experience reliable improvement

(Wolgast et al., 2003). Symptom improvement and therapeutic change can be impacted by extra therapeutic client factors as well which were not thoroughly assessed in this study.

Percentage of Reliable Improvement Results

The findings of this study indicated that for the entire sample 50.1% of clients achieved reliable improvement, 34.0% achieved neither reliable improvement nor reliable deterioration, and 16.0% experienced reliable deterioration. A reliable improvement rate of 50.1% was comparable to other reliable improvement rates found in literature. In another study where the sample including both licensed professionals as well as trainees, 57% of the sample met criteria for reliable improvement (Daniel, 2006). Other studies assessing reliable improvement among trainees showed that 40% (Samstag & Norlander, 2019) and 12.2% (Carr et al., 2017) experienced reliable improvement, while 29% achieved clinically significant improvement (Carr et al., 2017). These percentages were comparable to the level of reliable improvement observed in this study.

Education Level and Change Status Results

There is a body of literature that has shown both differences (e.g.: Budge et al., 2013; Driscoll et al., 2003; Holloway, 1992, Stoltenberg et al., 1994) and a lack of differences (e.g.: Lambert et al., 2003; Michael et al., 2005; Nyman et al., 2010; Propst et al., 1994; Stein & Lambert, 1995) in clinical outcomes between professional mental health providers with unique levels of education and training. This study has provided further data that supports the claim that education level does not have a significant relationship with clinical outcomes or change status (Reliably Improved, Unchanged, or Deteriorated) among early clinicians in training with either a bachelor's or a master's degree. Moreover, these findings also showed a statistically significant

higher likelihood of experiencing reliable change among clients with higher initial OQ-45.2 scores than clients with lower baseline OQ-45.2 scores. This was commensurate with other studies that analyzed client distress level and response to treatment.

The literature has shown that level of distress and symptom presentation can mediate therapeutic outcomes (Mohr et al., 1990). Acute symptom presentations are associated with quicker responses to treatment when compared with longer standing characterological symptoms (Barkham et al., 1996; Kopta et al., 1994). Uckelstam et al. (2019) analyzed client distress and improvement rates to further explore this association. They found that clients with higher levels of distress and lower levels of suicidal ideation or self-harming behaviors at intake experienced higher rates of change when compared with clients with very high distress levels who did experience suicidal ideation or self-harming behaviors (Uckelstam et al., 2019). In this study, both the high and very high distress groups experienced rates of change that were steeper than the low or average distress groups (Uckelstam et al., 2019). This association between severity of distress and steepness of outcomes could be due to increased motivation to engage in treatment, due to high symptom severity (Mohr et al., 1990). In contrast, higher levels of interpersonal distress can lead to difficulty achieving therapeutic outcomes (Steenbarger, 1994). This is likely due to decreased ability to build rapport with one's counselor and less openness to clinical interventions (Steenbarger, 1994).

Clients with higher levels of distress could be more likely to experience quicker rates of reliable improvement, due to RC having less stringent requirements than clinically significant change. The RCI for the OQ-45.2 requires a 14-point decrease in scores from one time point to another time point later in order to meet criteria for reliable improvement. Despite the lack of significant relationship between counselor education level and reliable change status, there were

patterns that existed among counselor data. Data from the 10 highest performing and 10 lowest performing counselors from the post bachelor's and post master's counseling groups were placed into Tables 9, 10, 11, 12, and 13.

Tables 9, 10, 11, 12, and 13 have highlighted differences in performance among participants in the sample. Although the average number of clients seen by the master's counselor group (8) and bachelor's counselors group (10) were comparable, we can see from Tables 9 and 11 that the highest performing counselors from the two education levels had a variety of total number of clients seen. For the master's counseling group, the top five highest performing counselors only had one client included in this study, whereas only one of the top five highest performing bachelor's counselors had one client. Having one client included in the sample for these post master's counselors likely inflated their rates of reliable change.

The average number of clients seen and average number of sessions until reliable change occurred also highlight a slight difference among the top 10 highest performing post bachelor's and post master's counselors. The post master's cohort saw an average of 6.5 clients and took 6.3 sessions on average to achieve reliable change. For the post bachelor's cohort there was an average of 9.9 clients per counselor with an average of 5.3 sessions until reliable change occurred. These results were contrary to the hypothesis. The top 10 highest performing post bachelor's counselors had more clients on their average caseload and were able to elicit reliable improvement in a shorter period of time. Although these differences were not statistically significant, it indicates a trend in the data.

It is also worth noting that the percentage of reliably deteriorated clients was higher for the post master's cohort (18.5%) versus the post bachelor's cohort (14.8%). One could speculate that this might mean the top performing bachelor's counselors outperformed the highest

performing master's counselors. These findings were surprising, due to the hypothesis that master's counselors would outperform bachelor's counselors with regard to clinical outcomes. This study provided further data to support that therapists vary in their ability to elicit change in the clients they serve. The findings showed differences among trainees that are similar to what has been reported in therapist effects literature showing effect sizes ranging from small to large in naturalistic studies (Kraus et al., 2011).

Areas such as well-being, burnout, and attention to therapeutic alliance are variables that could also impact treatment outcomes (Kraus et al., 2011). Counselors in training might experience higher levels of burnout, lower well-being, and might have less experience, leading to less focus on the quality of the therapeutic alliance. Variability in these individual domains could have impacted the results of this study. Certain counselors might have more support or resources for navigating the increased stress associated with graduate training. It would be worthwhile to explore variables related to counselor well-being in future studies. These findings have highlighted the importance of continued exploration of within-group differences among counselors in training to better understand how clinical outcomes and clinical training can be improved.

Limitations and Future Directions

Despite the focus on within-group differences of clinicians in this study, it is important to also acknowledge the limitations of using outcome measures to assess clinician performance. Approximately 40% of the variance in clinical outcomes is believed to be due to client factors, which are unrelated to the therapy process or therapist (Lambert, 1986). Lambert (1992) estimated that another 30% of variance in outcomes is due to therapeutic relationship, 15% to technical factors, and 15% due to client expectancy.

Using education as a measure of clinical competency was another limitation of this study. Doctoral students in this sample had varying levels of clinical experience, and students with bachelor's degrees might not have had significantly different levels of previous clinical experiences when compared to master's level students who entered into the same program. For this dataset, it was only known whether the counselor participants had a master's degree or bachelor's degree. Information about the type or amount of actual clinical experiences associated with the degree was unknown. For example, bachelor's level students who start a doctoral program years after finishing their bachelor's degree could have had significantly more clinical experiences than a post master's student who entered into a doctoral program right after the termination of their master's program. In future research it would be helpful to assess competence through actual number of clinical hours instead of degrees held. Future studies could also include a measure of supervisory evaluation to compare with a counselor's percentage of clinical improvement to further assess the relationship between clinical competence and outcomes.

Using the reliable change index (RCI) versus a clinically significant change measure was another limitation of this study. The RCI does not require that a participant's scores start in a dysfunctional score range and move to a functional score range in order to be included in the study. Despite this, a reliable improvement in outcome score might indicate noteworthy changes in an individual's quality of life. The OQ-45.2 is also a self-report measure, and self-report measures have shown inconsistency when compared with clinician reports of therapeutic progress. Cuijpers et al. (2010) showed statistically significant differences between clinician ratings and self-report ratings of therapeutic progress in clients with depression. In the aforementioned study, these differences were hypothesized to be due to client self-report

outcome assessment being more conservative than clinician ratings and clinician ratings being more sensitive to change than client ratings (Cuijpers et al., 2010). Employing a clinician rating or rating from a close family member or friend could provide a more powerful understanding of therapeutic progress.

Another important limitation to consider was that the population of clients served in this dataset were highly marginalized and had low access to resources, low socioeconomic status, and poor community and social supports. We know based on Maslow's hierarchy of needs (1943) that individuals must have their basic needs met prior to engaging in other higher level needs, such as self-actualization. Clients in this dataset could have been less engaged in treatment, due to the systemic barriers they experienced in accessing mental healthcare. Future studies should include a measure that assesses client social support, access to resources, and severity of presenting concerns to allow these variables to be considered when analyzing a counselor's percentage of clients with reliable improvement. The results of this study might have looked different if another clinical population with higher access to resources and increased social and community supports was analyzed, such as clients from a University Counseling Center (UCC). This study also consisted of an unequal sample size of each cohort of bachelor's (N= 59) and master's counselors (N=33) who were also not selected at random to participate. Ideally, a sample of counselor participants who were randomly selected from other universities and settings would have provided results with greater generalizability.

Due to the population served in this study, it would have also been interesting to assess clients' overall trend of OQ-45.2 scores. This could have provided a deeper understanding of subtle differences that might exist in post bachelor's and post master's counselors in training. Further research should directly analyze the clinical outcomes of counselors in training across

time from semester to semester throughout their graduate training. With this knowledge, counselor training programs might be better equipped to identify which counselors in training need additional support or intervention.

Conclusions

In summary, this study provided no support for the common assumption that increases in training and education lead to increases in a helping professional's ability to elicit therapeutic changes in the clients they serve. Post bachelor's and post master's counselors had comparable rates of reliably improved, unchanged, and deteriorated clients in this sample. Despite these similarities, the study also highlighted some within-group and between-group differences among the samples of post bachelor's and post master's counselors. Some noteworthy, yet statistically nonsignificant, differences included the reliable deterioration rate of 18.5% in the post master's cohort, versus a rate of 14.8% in the post bachelor's cohort, the smaller average number of clients seen in the post master's versus post bachelor's cohorts, and the higher average number of sessions until reliable change occurred for the post master's versus post bachelor's cohorts Future research should continue to explore how to strengthen our understanding of therapist effects so that poorly performing counselors in training can be identified and provided with effective interventions to boost their professional skill acquisition if needed. Despite the importance of facilitating growth and change within clinical populations, there has been little integration of clinical outcomes in understanding counselor development. Increased incorporation of clinical outcome measures into supervision and clinical training could be an avenue to increase the efficacy of clinical work.

REFERENCES

Ackerman, S. J., Benjamin, L. S., Beutler, L. E., Gelso, C. J., Goldfried, M. R., Hill, C., Lambert, M. J., Norcross, J. C., Orlinsky, D. E., & Rainer, J. (2001). Empirically supported therapy relationships: Conclusions and recommendations for the Division 29 Task Force. *Psychotherapy: Theory, Research, Practice, Training*, *38*(4), 495–497. https://doi.org/10.1037/0033-3204.38.4.495

Alberts, G., & Edelstein, B. (1990). Therapist training: A critical review of skill training studies. *Clinical Psychology Review*, *10*(5), 497–511. https://doi.org/10.1016/0272-7358(90)90094-Q

American Psychological Association. (2009a) Guidelines and Principles for the Accreditation of Programs in Professional Psychology. American Psychological Association.

American Psychological Association. (2011). Revised Competency Benchmarks for Professional Psychology. https://www.apa.org/ed/graduate/competency

American Psychological Association. (2012). Benchmarks Rating Form all levels. https://www.apa.org/ed/graduate/benchmarks-evaluation-system

Anker, M. G., Duncan, B. L., & Sparks, J.A. (2009). Using client feedback to improve couple therapy outcomes: a randomized clinical trial in a naturalistic setting. *Journal of Consulting & Clinical Psychology*, 77, 693704. https://doi.org/10.1037/a0016062

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose–effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology*, 77(2), 203-211. https://doi.org/10.1037/a0015235

Barkham, M., Connell, J., Stiles, W., Miles, J. N., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose– effect relations and responsive regulation of treatment duration: The good

enough level. *Journal of Consulting and Clinical Psychology*, 74, 160 –167. https://doi.org/10.1037/0022-006X.74.1.160

Barkham, M., Rees, A., Stiles, W., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). Dose-effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology*, 64, 927-935. https://doi.org/10.1037/0022-006X.64.5.927

Baltimore, M. L., Hickson, J., George, J. D., & Crutchfield, L. B. (1996). Portfolio assessment: A model for counselor education. *Counselor Education and Supervision*, *36*(2), 113–121. https://doi.org/10.1002/j.1556-6978.1996.tb00377.x

Budge, S. L., Owen, J. J., Kopta, S. M., Minami, T., Hanson, M. R., & Hirsch, G. (2013). Differences among trainees in client outcomes associated with the phase model of change. *Psychotherapy*, *50*(2), 150–157. https://doi.org/10.1037/a0029565

Bland, A. M. (2018). Facilitating and assessing personal growth in helper development using Hart's (2014) Four Virtues. *The Humanistic Psychologist*, *46*(1), 6–29. https://doi.org/10.1037/hum0000078

Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, *25*(1), 6–19. https://doi.org/10.1080/10503307.2013.817696

Carr, M. M., Saules, K. K., Koch, E. I., & Waltz, T. J. (2017). Testing the dose-response curve in a training clinic setting: Use of client pretreatment factors to minimize bias in estimates. *Training and Education in Professional Psychology*, *11*(1), 26–32. https://doi-org.srv-proxy2.library.tamu.edu/10.1037/tep0000135

Crits-Christoph, P., Baranackie, K., Kurcias, J. S., Beck, A. T., Carroll, K., Perry, K., Luborsky, L., McLennan, A., Woody, G., Thompson, L., Gallagher, D., & Zitrin, C. (1991). Meta-

analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research*, *1*(2), 81–91. https://doi.org/10.1080/10503309112331335511

Cuijpers, P., Li, J., Hofmann, S. G, & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: A meta-analysis. *Clinical Psychology Review, 30*(6), 768-778. https://doi.org/10.1016/j.cpr.2010.06.001

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16

Daniel, L. T. (2006). Predicting response to counseling from patterns of symptoms at intake: Effects of treatment length and training level of counselor. *Dissertation Abstracts International: Section B: The Sciences and Engineering*. ProQuest Information & Learning. http://search.ebscohost.com.srv-proxy2.library.tamu.edu/login.aspx?direct=true&db=psyh&AN=2006-99016-045&site=ehost-live

Dodenhoff, J. T. (1981). Interpersonal attraction and direct-indirect supervisor influence as predictors of counselor trainee effectiveness. *Journal of Counseling Psychology*, *28*,(1) 47–52. https://doi-org.srv-proxy1.library.tamu.edu/10.1037/0022-0167.28.1.47

Driscoll, K. A., Cukrowicz, K. C, Reitzel, L. R., Hernandez, A., Petty, S. C, & Joiner, T. E., Jr. (2003). The effect of trainee experience in psychotherapy on client treatment outcomes. *Behavior Therapy*, 34, 165-177. https://doi.org/10.1016/S0005-7894(03)80011-4

Duncan, B. L., & Reese, R. J. (2016). Using PCOMS technology to improve outcomes and accelerate counselor development. In T. Rousmaniere & E. Renfro-Michel (Eds.), *Using*

technology to enhance clinical supervision. (pp. 135–156). *American Counseling Association*. https://doi-org.srv-proxy1.library.tamu.edu/10.1002/9781119268499

Eells, T D., Lombart, K. G., Kendjelic, E. M.,Tumer, L. C, & Lucas, C. P. (2005). The quality of psychotherapy case formulations: A comparison of expert, experienced, and novice cognitive-behavioral and psychodynamic therapists. *Journal of Consulting and Clinical Psychology*, 73, 579-589. http://dx.doi.org.srv-proxy2.library.tamu.edu/10.1037/0022-006X.73.4.579

Erekson, D. M., Lambert, M. J., & Eggett, D. L. (2015). The relationship between session frequency and psychotherapy outcome in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, *83*(6), 1097–1107. https://doi.org/10.1037/a0039774

Fauth, J., Gates, S., Vinca, M. A., Boles, S., & Hayes, J. A. (2007). Big ideas for psycho-therapy training. *Psychotherapy: Theory, Research, Practice, Training*, 44, 384–391.https://doi-org.srv-proxy1.library.tamu.edu/10.1037/0033-3204.44.4.384

Ford, J. D. (1979). Research on training counselors and clinicians. *Review of Educational Research,* 69, 87-130. https://doi-org.srv-proxy1.library.tamu.edu/10.2307/1169928

Freitas, G. J. (2002). The impact of psychotherapy supervision on client outcome: A critical examination of 2 decades of research. *Psychotherapy: Theory, Research, Practice, Training*, *39*(4), 354–367. https://doi.org/10.1037/0033-3204.39.4.354

Gordon, R. (2012). Where oh where are the clients? The use of client factors in counselling psychology. *Counselling Psychology Review*, *27*(4), 8–17. http://search.ebscohost.com.srvproxy1.library.tamu.edu/login.aspx?direct=true&db=psyh&AN=2013-04085-002&site=ehost-live

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D., Shimokawa, K. W., & Sutton, S. W. (2005) A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology: In Session*, 61, 155-163. https://doi.org/10.1002/jclp.20108

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, 9(3), 329-343. https://doi.org/10.1093/clipsy.9.3.329

Hart, T. (2014). *The four virtues.* Atria.

Hart, T., & Hart, M. (2014). Spiritual assessment matrix. http://thefourvirtues.com/ survey

Holloway, E. L. (1992). Supervision: A way of teaching and learning. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology* (2nd ed., pp. 177–214). Wiley.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist*, *41*(2), 159–164. https://doi-org.srv-proxy2.library.tamu.edu/10.1037/0003-066X.41.2.159

Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, 51(10), 1059–1064. https://doi.org/10.1037/0003-066X.51.10.1059

Iberg, J. R. (1991). Applying statistical control theory to bring together clinical supervision and psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(4), 575–586. https://doi.org/10.1037/0022-006X.59.4.575

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. http://dx.doi.org/10.1037/0022-006X.59.1.12

James, K., (2016). "Barriers to Treatment and the Connection to Maslow's Hierarchy of Needs". Counselor Education Capstone. 18. http://digitalcommons.brockport.edu/edc_capstone/18

Jensen, G. M., & Saylor, C. (1994). Portfolios and Professional Development in the Health Professions. *Evaluation & the Health Professions*, 17(3), 344–357. https://doi.org/10.1177/016327879401700307

Kadera, S. W., Lambert, M. J., & Andrews, A. A. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *Journal of Psychotherapy Practice & Research*, *5*(2), 132–151. http://search.ebscohost.com.srv-proxy1.library.tamu.edu/login.aspx?direct=true&db=psyh&AN=1996-03833-004&site=ehost-live

Kleine-Kracht, D. (2019). *Predicting treatment response and no-show appointments for low-income and rural populations at a community mental health clinic*. [Unpublished doctoral dissertation]. Texas A&M University.

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, 6 2 ,1009-1016. doi: 10.1037//0022-006x.62.5.1009.

Kozina, K., Grabovari, N., De Stefano, J., & Drapeau, M. (2010). Measuring changes in counselor self-efficacy: Further validation and implications for training and supervision. *The Clinical Supervisor*, *29*(2), 117–127. https://doi-org.srv-proxy1.library.tamu.edu/10.1080/07325223.2010.517483

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, 267276. https://doi.org/10.1080/10503307.2011.563249

Ladany, N., & Melincoff, D. S. (1999). The nature of counselor supervisor Nondisclosure. *Counselor Education and Supervision*, 38, 161–176. https://doi.org/10.1002/j.1556-6978.1999.tb00568.x

Lambert, M. J. (1986). Some implications of psychotherapy outcome research for eclectic psychotherapy. *Journal of Integrative & Eclectic Psychotherapy*, *5*(1), 16–45. http://search.ebscohost.com.srv-proxy2.library.tamu.edu/login.aspx?direct=true&db=psyh&AN=1987-25823-001&site=ehost-live

Lambert, M. (1992). Implications of outcome research for psychotherapy integration. In J. Norcross & J. Goldstein (Eds.), Handbook of psychotherapy integration (pp. 94–129). https://doi-org.srv-proxy1.library.tamu.edu/10.1093/med:psych/9780195165791.001.0001

Lambert, M. J. (2005). Early response in psychotherapy: Further evidence for the importance of common factors rather than "placebo effects." *Journal of Clinical Psychology*, 61, 855-869. https://doi.org/10.1002/jclp.20130

Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, *17*(1), 1–14. https://doi.org/10.1080/10503300601032506

Lambert, M. J., Hansen, N. B., Umpress, V., Lunnen, K., Okiishi, J., Burlingame, G. M., & Reisinger, C. W. (1996). *Administration and scoring manual for the Outcome Questionnaire (OQ-45.2)*. American Professional Credentialing Services.

Lambert, M. J., & Hawkins, E. J. (2001). Using information about patient progress in supervision: Are outcomes enhanced? *Australian Psychologist*, 36(2), 131–138.

https://doi.org/10.1080/00050060108259645

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*, *10*(3), 288–301. https://doi.org/10.1093/clipsy/bpg025

Lumadue, C. A., & Duffey, T. H. (1999). The role of graduate programs as gatekeepers: A model for evaluating student counselor competence. *Counselor Education and Supervision*, *39*(2), 101–109. https://doi.org/10.1002/j.1556-6978.1999.tb01221.x

Liu, X. (2009). Ordinal regression analysis: Fitting the proportional odds model using Stata, SAS and SPSS. *Journal of Modern Applied Statistical Methods*, *8*(2), 30.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50, 370-396.

Mayfield, W. A., Kardash, C. M., Kivlighan, D. M., Jr. (1999). Differences in experienced and novice counselors; knowledge structures about clients: Implications for case conceptualization. *Journal of Counseling Psychology*, 46, 504—514. http://dx.doi.org.srv-proxy2.library.tamu.edu/10.1037/0022-0167.46.4.504

McHugh M. L. (2013). The chi-square test of independence. *Biochemia medica*, 23(2), 143–149. https://doi.org/10.11613/bm.2013.018

Michael, K. D., Huelsman, T. J., & Crowley, S. L. (2005). Interventions for child and adolescent depression: Do professional counselors produce better results? *Journal of Child and Family Studies*, 14, 223-236. https://doi.org/10.1007/s10826-005-5050-8

Miller, S. D., & Duncan, B. L., & (2000). The Outcome and Session Rating Scales: Administration and scoring manuals. http://www.talkingcure.comMohr, D. C., Beutler, L. E., Engle, D., Shoham-Salamon, V., Bergan, J., Kaszniak, A. W.

& Yost, E. B. (1990). Identification of patients at risk for non-response and negative outcome in psychotherapy. *Journal of Consulting and Clinical Psychology*, 58, 622-628. https://doi.org/10.1037/0022-006X.58.5.622

Nyman, S. J., Nafziger, M. A., & Smith, T. B. (2010). Client outcomes across counselor training level within a multitiered supervision model. *Journal of Counseling & Development*, *88*(2), 204–209. https://doi.org/10.1002/j.1556-6678.2010.tb00010.x

Propst, A., Paris, J., & Rosberger, Z. (1994). Do therapist experience, diagnosis, and functional level predict outcome in short-term psychotherapy? *Canadian Journal of Psychiatry*, 39, 168-176. https://doi.org/10.1177/070674379403900309

Reese, R. J., Norsworthy, L. A., & Rowlands, S. R. (2009a). Does a continuous feedback system improve psychotherapy outcome*? Psychotherapy: Theory, Research, Practice, Training,* 46, 418-431. https://doi.org/10.1037/a0017901

Reese, R. J., Usher, E. L., Bowman, D. C., Norsworthy, L. A., Halstead, J. L., Rowlands, S. R., & Chisolm, R. R. (2009b). Using client feedback in psychotherapy training: An analysis of its influence on supervision and counselor self-efficacy. *Training and Education in Professional Psychology*, 3, 157-168. https://doi.org/10.1037/a0015673

Seligman, Martin. (1995). The effectiveness of psychotherapy. The Consumer Reports study. *American Psychologist*, 12, 965-974. doi:10.1037/0003-066X.50.12.965

Samstag, L. W., & Norlander, K. (2019). Characteristics of trainees' early sessions: A naturalistic process-outcome study tribute to Jeremy D Safran. *Psychoanalytic Psychology*, *36*(2), 148–158. https://doi.org/10.1037/pap0000239

Steenbarger, Brett N. (1994). Duration and outcome in psychotherapy: An integrative

review. *Professional Psychology: Research and Practice*, 25, 111-119. https://doi.org/10.1037/0735-7028.25.2.111

Stein, D. M., & Lambert, M. J. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced? *Journal of Consulting and Clinical Psychology*, 63, 182–196. https://doi.org/10.1037/0022-006X.63.2.182

Stoltenberg, C. D., McNeill, B. W., & Crethar, H. C. (1994). Changes in supervision as counselors and therapists gain experience: A review. *Professional Psychology: Research and Practice*, 25, 416–449. http://dx.doi.org/10.1037/0735-7028.25.4.416

Stoltenberg, C. D., McNeill, B. W., & Delworth, U. (1998). IDM supervision: An integrated developmental model for supervising counselors and therapists. Jossey-Bass.

Stoltenberg, C. D. (2005). Enhancing Professional Competence Through Developmental Approaches to Supervision. *American Psychologist*, *60*(8), 857–864. https://doi.org/10.1037/0003-066X.60.8.85

Spurgeon, S. L., Gibbons, M. M., & Cochran, J. L. (2012). Creating personal dispositions for a professional counseling program. *Counseling and Values*, *57*(1), 96–108. https://doi.org/10.1002/j.2161-007X.2012.00011.x

Uckelstam, C.-J., Philips, B., Holmqvist, R., & Falkenström, F. (2019). Prediction of treatment outcome in psychotherapy by patient initial symptom distress profiles. *Journal of Counseling Psychology*, *66*(6), 736–746. http://dx.**doi**.org/10.1037/cou0000345

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(2), 639–644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644

Wampold, B. E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist*, 62, 857-873. https://doi.org/10.1037/0003-066X.62.8.857

Wampold, B. E., & Budge, S. L. (2012). The 2011 Leona Tyler Award Address: The relationship—And its relationship to the common and specific factors of psychotherapy. *The Counseling Psychologist*, *40*(4), 601–623. https://doi.org/10.1177/0011000011432709

Wampold, B. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14 (3), 270-277. https://doi.org/10.1002/wps.20238.

Wolgast, B. M., Lambert, M. J., & Puschner, B. (2003). The Dose-Response Relationship at a College Counseling Center: Implications for Setting Session Limits. *Journal of College Student Psychotherapy*, *18*(2), 15–29. https://doi-org.srv-proxy2.library.tamu.edu/10.1300/J035v18n02_03

APPENDIX A

Table A1.

Descriptive Statistics of Number of Sessions for Bachelor's Counselors who Achieved Reliable Change

Counselor ID	*M*	*Mdn*	Mode	Range
22	4.0	4.0	-	0
34	4.9	4.0	3	8
16	7.7	10.0	10	7
77	5.0	3.0	3	10
94	7.3	5.0	2	32
31	5.0	4.0	2	10
74	3.8	3.0	3	7
24	6.9	6.0	2	11
36	4.5	5.0	5	4
42	4.0	3.5	3	5
52	4.5	4.0	4	7
57	5.0	6.0	6	9
33	6.6	7.0	7	14
30	3.8	3.0	3	5
39	4.3	3.5	2	7
78	3.3	3.0	3	1
87	7.2	4.0	3	14
53	3.1	2.0	2	6
64	7.8	6.0	-	15
8	7.0	8.0	-	9
51	5.5	3.5	2	13
41	2.9	2.0	2	4
80	4.0	3.0	2	8
70	3.0	3.0	-	2
68	4.7	5.0	7	5
9	2.4	2.0	2	1
10	4.2	2.5	2	10
38	3.0	2.0	2	5
72	5.7	5.5	2	11
89	6.8	6.5	-	10
63	3.4	3.0	2	5
69	4.2	4.0	4	4
13	7.3	2.5	2	20
35	2.8	2.5	2	2
85	7.3	7.5	-	8
47	9.0	9.0	-	8
90	3.0	3.0	-	2

Counselor ID	*M*	*Mdn*	Mode	Range
2	3.4	3.0	2	4
50	5.6	3.0	3	15
65	3.8	2.0	2	6
27	3.3	2.5	2	5
4	4.0	3.0	-	5
11	3.6	2.0	2	5
19	15.0	15.0	-	0
28	3.6	3.0	2	4
40	7.7	6.0	-	13
46	4.2	4.0	4	2
55	4.0	4.0	-	0
67	9.0	9.0	-	14
73	2.0	2.0	-	0
48	3.5	3.5	-	3
23	3.0	3.0	-	0
25	2.0	2.0	-	0
60	2.8	2.5	2	2
44	5.0	5.0	-	4
59	2.3	2.0	2	1
21	-	-	-	-
37	-	-	-	-
95	-	-	-	-

Note. Dashes (-) were used for counselors who had 0 clients who reached reliable improvement.

Table A2.

Descriptive Statistics of Number of Sessions for Master's Counselors who Achieved Reliable Change

Counselor ID	*M*	*Mdn*	Mode	Range
17	6.0	6.0	-	0
45	7.0	7.0	-	0
76	2.0	2.0	-	0
79	19.0	19.0	-	0
81	2.0	2.0	-	0
3	3.2	2.0	2	7
92	4.9	3.5	3	10
29	2.7	3.0	3	1
83	2.9	2.0	2	4
43	13.0	7.0	5	46
6	4.4	4.0	4	3
49	7.0	6.0	-	11
75	5.7	4.0	2	19
62	3.8	3.0	3	4
1	3.9	2.0	2	8
26	4.8	4.0	4	9
32	5.5	6.0	8	6
56	6.0	5.0	10	8
61	3.7	4.0	2	3
71	2.0	2.0	-	0
86	4.0	4.0	4	4
20	3.0	3.0	-	2
82	3.0	2.5	2	3
84	10.3	14.0	14	11
88	5.7	4.0	4	5
14	3.3	3.0	3	1
58	4.5	4.5	-	3
66	8.5	8.5	-	5
12	5.0	5.0	-	6
91	2.5	2.5	-	1
54	3.0	3.0	-	0
15	-	-	-	-
18	-	-	-	-

Note. Dashes (-) were used for counselors who had 0 clients who reached reliable improvement.