



2014-04-09

Are Therapists Using Outcome Measures and Does It Matter? A Naturalistic Usage Study

Jared S. Klundt

Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

 Part of the [Counseling Psychology Commons](#), and the [Special Education and Teaching Commons](#)

BYU ScholarsArchive Citation

Klundt, Jared S., "Are Therapists Using Outcome Measures and Does It Matter? A Naturalistic Usage Study" (2014). *All Theses and Dissertations*. 4028.

<https://scholarsarchive.byu.edu/etd/4028>

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Are Therapists Using Outcome Measures and Does It Matter?

A Naturalistic Usage Study

Jared S. Klundt

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

Doctor of Philosophy

Mark E. Beecher, Chair
John Okiishi
Derek Griner
Aaron Jackson
Marleen Williams

Department of Counseling Psychology and Special Education

Brigham Young University

April 2014

Copyright © 2014 Jared S. Klundt

All Rights Reserved

ABSTRACT

Are Therapists Using Outcome Measures and Does It Matter? A Naturalistic Usage Study

Jared S. Klundt

Department of Counseling Psychology and Special Education, BYU
Doctor of Philosophy

Outcome measurement has been demonstrated to be beneficial when used as a routine part of psychological practice (Lambert et al., 2002), improving both treatment time and outcome of patients, and helping therapists determine which of their patients are not responding to their current treatments (Hannan et al., 2005; Whipple et al., 2003). Despite these proven benefits, many therapists are reluctant to implement and use outcome measures in their routine practice (Hatfield & Ogles, 2007). In addition, the research demonstrating the benefits of these outcome measures has used randomized experimental design to examine the effects of outcome feedback to therapists. The current study focused on examining the benefits of naturalistic usage of outcome measurement feedback in a setting where such measures are a part of routine practice. Therapists' regular use of the feedback from the Outcome Questionnaire-45 was found to have statistical but not clinical significance in the prediction of a patient's symptom change. Additionally, the regularity with which a therapist views a patient's outcome measure feedback was not significantly correlated with that patient's treatment length, overall change in symptoms, or symptom level at termination. Implications regarding these findings are discussed and recommendations regarding the future study of outcome measures are provided.

Keywords: outcome measurement, naturalistic, feedback

ACKNOWLEDGEMENTS

This hoop was not jumped through by my efforts alone. I owe a debt of gratitude to Mark, who has been an incredibly supportive chair. Without his gentle prodding, I would not have been able to finish this work. Thanks to John, my former chair and expert guide to OQ research, for both the initial seeds of the idea as well as frank feedback about my work. Derek, Aaron, and Marleen have all been so responsive as I have scrambled to meet deadlines, and I am grateful for their willingness to put up with my procrastination. And I am indebted to Lane Fischer, whose statistical knowledge is only superseded by his patience in explaining it to students who ought to have listened better.

I am grateful for the high value placed on education by my family, and their support of my pursuit of even further schooling. I am particularly indebted to my father, whose love for school and learning forms some of my earliest memories. My friends have been inexhaustibly supportive as they have listened to me ramble about topics and problems over which they have no control and even less interest. The Beesleys, in particular, are responsible for keeping me grounded, working, and laughing. And, of course, my amazing fiancée, who provided that final motivation that I doubt I would have found without her. Thank you, Cate.

TABLE OF CONTENTS

Introduction.....	7
Method.....	9
Participants.....	9
Measure.....	10
Procedure.....	13
Results.....	14
Percentage of Patients' Sessions for Which Feedback Was Reviewed.....	15
Therapist Vigilance.....	15
Discussion.....	16
Conclusion.....	20
References.....	22
Appendix: Literature Review.....	30
Historical/Political Influence on Outcome Measurement Development.....	30
Philosophical Constraints Overcome by Outcome Measurement.....	32
Practical Constraints Alleviated by Outcome Measurement.....	34
Patient-Focused Research Paradigm as Precursor to Outcome Measurement.....	36
Benefits of Outcome Measures.....	37
Therapist Resistance to Outcome Measures.....	39
Present Study Goals.....	41
References.....	44

LIST OF TABLES

Table	Page
1. Correlations with the Percentage a Therapist Views Their Patients' OQ Results.....	28

LIST OF FIGURES

Figure	Page
1. Sample OQ-45 digital feedback.....	29

Introduction

The evidence based practice (EBP) movement arose in an environment of increased financial pressure, government regulation, and decreased confidence in healthcare providers' expertise (Burlingame & Beecher, 2008; DeLeon, VandenBos, & Bulatao, 1991; Reed & Eisman, 2006). Health maintenance organizations (HMOs) and public policy makers, faced with spiraling healthcare costs, adopted the public idea that bringing healthcare practice into conformity with research (through EBP) would minimize and justify healthcare costs and increase efficiency (DeLeon et al., 1991; Tanenbaum, 1999). Psychologists responded to this increased scrutiny by establishing guidelines for determining best practices through empirical research methods examining two primary criteria: *treatment efficacy*, or whether or not a treatment works better than placebo; and *effectiveness* or *clinical utility*, the usefulness of a given treatment for a specific context (American Psychological Association, 1995; 2002a; 2005). These standardized guidelines were built upon philosophical assumptions of empiricism and universalism, two values with problematic implications for social science research and practice (Slife & Whoolery, 2006). For example, these philosophical foundations limit the type of practices that can be examined through the EBP paradigm, and ignore some of the fundamental tenets of ethical psychological practice (Sue, Zane, Nagayama Hall, & Berger, 2009; Westen, Novotny, & Thompson-Brenner, 2004). In response to these limitations, Howard, Moras, Brill, Martinovich, and Lutz (1996) introduced a new model for examining psychological treatment: the patient focused research (PFR) paradigm, which asks, "Is this treatment working for this patient's condition at the current time?" This question is answered by the use of pan-theoretical tools used to track individual patient progress called *outcome measures*, which provide evidence on which clinicians can base their practice while avoiding some of the philosophical problems of

empirically based methodologies. The benefits of using outcome measures in routine practice have been highlighted by multiple studies. For instance, therapists using outcome measures are better able to catch patients at risk for treatment failure (Hannan et al., 2005), encourage benchmarking across national healthcare systems (Barkham, Mellor-Clark, Connell, & Cahill, 2006), and improve outcomes for both normal progression and at-risk patients (Lambert, Harmon, Slade, Whipple, & Hawkins, 2005; Slade, Lambert, Harmon, Smart, & Bailey, 2008; Whipple et al., 2003). This research has resulted in the adoption of outcome measurement worldwide (de Beurs et al., 2011; Hatfield & Ogles, 2004; Mellor-Clark, Barkham, Connell, & Evans, 1999; Pirkis et al., 2005). Despite the clear benefits and the significant number of outcome measurement systems available, a majority of practicing therapists are not using outcome measures (Hatfield & Ogles, 2007), citing logistical concerns (e.g., ‘Adds too much paperwork,’ ‘Takes too much time’).

One of the more widely used (Hatfield & Ogles, 2004) outcome measures, the Outcome Questionnaire-45 (OQ-45; Lambert et al., 2001; Lambert, Whipple, Smart, Vermeersch, & Nielsen, 2001) has demonstrated clinical utility in tracking and advising therapists regarding patient progress in therapy, and detecting and improving patients at risk for treatment failure (Hawkins, 2004; Lambert et al., 2002; Whipple et al., 2003). These results show that therapist use of these outcome measures has clinically and statistically significant benefits when compared with treatment in the absence of such feedback. However, this research has used strict experimental methodology to examine the treatment effect of outcome feedback. In an effort to better understand outcome measurement utility in non-experimental therapy, the present study examined how clinicians naturally employ outcome measurement in a treatment facility that removes logistical barriers to usage by including it as a standard component of routine practice.

This study focused on two primary questions regarding consistent use of outcome measures in a naturalistic setting: (a) Are patients better served by their therapists' routine use of the OQ feedback? and (b) Are therapists who consistently use the OQ more effective?

Method

Participants

Data for this study came from a 21-month period (from April 2011 through December 2012) of the archival dataset continually collected by a large university counseling center. This span was selected as it was the earliest period that the outcome measurement software was able to independently confirm whether or not therapists had accessed the client feedback. Though outcome measurement is a part of routine practice at this center, patients must give their permission for their anonymized data to be included in research as part of the intake process; only consenting patients' data were included. This dataset was further narrowed by excluding patients who had less than two administrations of the OQ-45 or who were seen in group or couples counseling ($n=2824$ clients, 16349 OQ administrations). The archival dataset is stripped of all identifying information; each patient and therapist was assigned a database-specific identifying number, unconnected to identifiable demographic data, and none of the research team had access to delineating information. Due to the limiting effect of this confidentiality, specific demographic information about the included patients is not available. However, during the time the data was collected, this counseling center serviced 3,266 university students, who were 59.3% female, aged 16-58 ($M= 22.7$, $SD = 4.1$, $Mode = 22$), with ethnic backgrounds reported as 89.3% Caucasian, 4.5% Hispanic, 3.1% Asian, 3.1% Other. One-hundred and fourteen therapists contributed data to this study, including pre-internship students, internship, and post-internship levels of training, from clinical psychology, counseling psychology, and marriage and

family therapy training backgrounds. The therapists were either salaried faculty of the university or students in training, and did not receive a direct fee for services provided.

Measure

The Outcome Questionnaire-45 (Lambert et al., 2004) is a 45-item, self-report measure of psychological functioning. Items are measured on a 5 point Likert scale: 0=never, 1=rarely, 2=sometimes, 3=frequently, 4=almost always. The range of scores possible on the OQ-45 is 0 to 180, with higher scores reflecting greater distress. The OQ-45 assesses three broad domains of patient functioning: (a) subjective discomfort, (b) interpersonal relationships, and (c) social role performance. The 45 items encompass many items that are characteristic of persons with a psychological disorder, but also includes 9 items that target positive functioning and are reverse scored. The OQ-45 was normed on local and national populations, clearly discriminates between patient and non-patient samples (Lambert et al., 2004; Umphress, Lambert, Smart, Barlow, & Clouse, 1997), and is currently available in 17 languages. Using formulas developed by Jacobson and Truax (1991), clinical and normative data for the OQ-45 were analyzed by Lambert, Morton, et al. (2004) to provide cutoff scores for the reliable change index (RCI; 14 points) and clinically significant change (dysfunctional/functional cutoff: 64/63). These formulas for cutoff scores and classifications for change are supported by multiple studies (Bauer, Lambert, & Nielsen, 2004; Lunnen & Ogles, 1998) and provide consensus with other self-report measures like the Beck Depression Inventory (Beck, Steer, Ball, & Ranieri, 1996b) and the Symptom Checklist-90 (Derogatis, 1996).

Feedback from the OQ-45 consists of an electronic progress graph indicating the patient's status, as well as a breakdown of the clinical subscales and the total score. The electronic feedback also alerts the therapist to crucial item responses to the OQ-45. For example, if a

patient endorsed “sometimes” to the question assessing suicidality, this item, which is considered to be one of several critical items, would appear in red on the electronic feedback.

In addition to the total score and critical item responses, the OQ-Analyst system is designed to give feedback identifying patients who are not improving as expected from psychotherapy. In conjunction with the RCI, Lambert et al. (1996) used information regarding the dose response relationship, initial level of severity, and typical patient response patterns to create decision rules for identifying patients who are not improving as expected from psychotherapy. These rules place the patient within an expected response category at the session of interest; patients who are progressing as expected are designated as “On Track” (OT, alert status white, green, or blue), while patients who are not progressing and are in fact deteriorating are designated as “Not On Track” (NOT, alert status red or yellow) (Slade et al., 2008). This designation and its impact on therapy has been the primary focus of prior research on this measure (Harmon et al., 2007; Hawkins, 2004; Lambert et al., 2002; Whipple et al., 2003), with consistent results showing that the OQ primarily excels at identifying these NOT cases and alerting therapists to modify their approach for those clients. These studies pre-dated the electronic administration and feedback of the measure, and utilized strict control conditions, assigning patients and therapists to distinct research categories.

This study focused on the real-world usage of the OQ-45 as presently available to clinicians, which is significantly different than the feedback and administration system employed in the original clinical trials and the principal follow up studies. In particular, the original trials of the measure did not have access to immediate electronic feedback; all feedback to therapists was calculated and distributed by hand. This meant that there was a delay between the administration of the measure and the feedback given to the therapist for that measure; feedback was given to

the therapist each time a subsequent OQ was administered (Lambert et al., 2001). This delay is no longer present for therapists who use the current system, as all feedback is available immediately after the patient finishes taking the measure. However, accessing this electronic feedback requires a therapist to take a few extra steps to log into the OQ-Analyst, search for the client in question, and select the appropriate results. This requires initiative on the part of the therapist to seek out the feedback; previous studies printed out and delivered the feedback to the clinicians who participated in the research.

In addition to the conditions surrounding outcome feedback access, the feedback provided to clinicians in this study is itself distinct from the feedback provided in other validations of the OQ system. These previous studies noted that clinically significant effects of feedback were only noticed for NOT cases (Shimokawa, Lambert, & Smart, 2010). The present study notes that the current system does not consistently distinguish these cases correctly, due to a limitation in the OQ-Analyst software package that does not accurately classify patients if they have been seen by a different therapist previously in the same clinic (previous course of therapy). For the clinic from which this study's data was collected, between 20-25% of patients have had multiple courses of therapy. Due to this software limitation that was not present in previous studies that predated the digital feedback system, therapists are inconsistently made aware of NOT patients currently. This misclassification may affect the utility of the measure for those crucial patients for whom it has demonstrated the greatest clinical utility. For this study, the timing of feedback, effort required to access that feedback, and the content of that feedback was significantly distinct from previous works examining the OQ, and was thus more aligned with current, real world usage patterns.

Procedure

The current study utilized a naturalistic design; rather than assigning patients to feedback (FB) or no feedback (NFB) conditions randomly at intake, the clinician's use of the feedback tool available to them varied based on the clinician's normal, treatment-as-usual (TAU) practices. By allowing usage of outcome measures to vary naturally, this study attempted to eliminate bias on the part of therapists who recognize that their outcomes are being tracked, as well as examine the real-world usage of outcome measures. As with previous work at this university, patient

...treatment was often suspended at the end of the school year because of the necessity for many of the students to return home during the summer months. Most decisions to terminate treatment were client initiated or jointly agreed on. This study was purposely designed to impose as little as possible on the manner in which therapists practiced therapy and to reflect the effect of using feedback...in routine practice. In this regard, experimental control was deemphasized to maximize ecological validity (Whipple et al., 2003, p. 62).

The archival database tracked the following variables: a patient and therapist identification number assigned by and unique to the database, the OQ score for each administration, and whether or not the calculated feedback for that administration was reviewed. From this data were calculated the number of sessions attended by the patient, the patient's initial OQ score, their final OQ score, the change in score from their first administration to their last, and the percentage of those administrations that had feedback reviewed by the treating therapist.

The major goal of the current study was an examination of the results of consistent outcome measures usage on patient symptom reduction and therapist efficacy. This was accomplished by examining how often treating therapists examined their patients' outcome

scores, and the effect that had on both individual patients and overall therapist caseloads. Specifically, the primary goal was to see if the reported benefits of the use of outcome measures are maintained in a naturalistic setting where there was no external research emphasis on the results. The present study examined therapist effects for both outcome usage and treatment results; in other words, do therapists who are more conscientious about using outcome measures have better results (even if they are not perfectly consistent in their usage)?

To examine whether or not usage of the current OQ system has an effect on patients and whether there are differences between therapists who use the system, the following hypotheses were tested: (a) The percentage of a patient's OQ scores that are examined by the treating therapist will not significantly predict that patient's OQ change after accounting for the severity of their OQ score at intake, and (b) There is no significant correlation between the percentage of sessions that a therapist examines their patients' OQ results and (c) the average length of therapy, (d) the average final OQ score, and (e) the average change in OQ score for that therapist's patients.

Results

Analyses were conducted using the IBM Statistical Package for the Social Sciences (SPSS) to evaluate the data. The dataset included 2824 patients and 16349 administrations of the OQ. The average number of sessions attended was 5.79, the average initial OQ score was 68.53, the average final OQ score was 62.35, the average change in OQ was -6.18 points, and the average patient had 33.2% of their outcome results examined by their treating therapist. While 37.2% of outcome feedback was viewed overall, 39.9% of patients had none of their outcome results examined. This finding is somewhat surprising considering that the OQ has been administered to every client admitted to the center at every session for over a decade, and that

since 2008, the feedback corresponding to each administration has been available instantly due to the digital administration of the measure.

Percentage of Patients' Sessions for Which Feedback Was Reviewed

The first question asked by this study was “Are patients better served by their therapists’ routine use of the OQ feedback?” A multiple regression analysis was performed to examine the OQ change (ΔOQ) predicted by the percentage of feedback reviewed (*%REVIEWED*) by the treating therapist in the presence of the patient’s initial OQ score (*IntakeOQ*). A significant regression was found ($F(2, 2821)=419.44, p<.000$), with an R^2 of .229. The frequency with which a therapist reviewed a patient’s OQ feedback did significantly predict OQ change even after accounting for that patient’s OQ score at intake. However, though statistically significant, *%REVIEWED* only accounted for 0.23% of the variance in ΔOQ , with *IntakeOQ* accounting for 22.27% of the variance in the patient’s OQ change. In other words, the fact that a patient’s therapist consistently viewed their OQ feedback accounted for less than 0.25% of that patient’s improvement.

Therapist Vigilance

In addition to the effects on patients, this study sought to examine whether therapists who consistently used the OQ were more effective in their practice overall. Efficacy was defined as in previous work by the average length of therapy for that therapist’s patients, the average change in symptoms between intake and termination, and the average amount of symptom change across their entire caseload. A Pearson Correlation was calculated to examine the relationship between a therapist’s use of the OQ feedback (*AVG%REVIEWED*) and that therapist’s average treatment length (*AVGTREAT*), average final OQ score (*AVGLASTOQ*), and average change in OQ score of therapist’s patients (*AVG Δ OQ*). No significant relationship was found between

AVG%REVIEWED and any of the dependent variables. In other words, therapists who more consistently viewed their patients' OQ feedback did not have better outcomes than therapists who rarely used the OQ feedback.

Discussion

This study focused on two primary questions regarding consistent use of outcome measures in a naturalistic setting: (a) Are patients better served by their therapists' routine use of OQ feedback? and (b) Are therapists who consistently use the OQ more effective? The goal was to examine therapist use of the OQ-45 in a setting where it has been implemented as part of routine practice for over a decade, without research conditions placing unnatural emphasis on its use. This setting is as near as possible to an 'ideal' setting for outcome measure usage, in that every single patient is asked to take the OQ-45 at every session, and every therapist has immediate access to the electronic feedback system in their office. Additionally, this center has full-time staff who administer the OQ to every patient, researchers who maintain access to the outcome scores, and a long-standing institutional culture that emphasizes and encourages the use of practice-based evidence.

Despite the benefits of using this university counseling center, using just this site as the sole source of data introduces some significant limitations to the results of this study. This counseling center restricts its services to current, full-time students and their families, and those students are overwhelmingly Caucasian (89.3%). This may significantly limit the applicability of these results to populations with larger multicultural representation, or to settings not dominated by university students in early adulthood.

With no external, novel emphasis encouraging therapists to apply outcome measures to their practice, the therapists in this particular setting viewed the outcome feedback available to

them only 37% of the time. Nearly two-thirds (61.4%) of the therapists viewed less than 50% of their patient's OQ feedback, with 13 (11.4%) of the tracked therapists never reviewing a single feedback report during the 18-month period. At the other end of the spectrum, the ten therapists who most consistently checked the OQ scores for their patients reviewed, on average, 85% of the available feedback. This wide gap may be due to experience with the measure. Of the 13 therapists who never reviewed the OQ scores of their clients, 12 of them had fewer than 20 sessions during that time. However, even examining the therapists who had at least 25 sessions in that timeframe, the bottom 10 therapists on average viewed only 6.4% of their patients' feedback reports, while the 10 most consistent therapists with 25 or more sessions reviewed 79% of the available feedback.

Considering this clinic's historical involvement in and cultural valuation of the use of this outcome measure, one would expect far higher rates of usage by the treating therapists. One possible explanation for such low overall usage (37%) may be the effort required for therapists to view feedback for their patients. In order to view OQ-45 feedback, therapists must log into a web-based portal, search for their client, and then access the current feedback results. While this process takes less than sixty seconds, it is an additional step that must be added to a clinician's pre-session routine. Previous work has shown that therapist utilization is a barrier in even less naturalistic research, with the most recent examination of outcome measures at this center showing less than two-thirds of directly-delivered feedback being reported as viewed (Bailey, 2009). The present study's goal of examining outcome usage without direct encouragement by the researchers meant that there was even less external accountability on therapists than previous studies, thus a lower overall usage of outcome measures.

In contrast to this lack of motivating accountability, another possible explanation for such low outcome usage may be due to a perceived lack of utility of outcome measures on routine practice. The present results and other previous work on the OQ-45 reinforce the concept that the OQ is primarily useful only for patients who are deteriorating. Thus, treating therapists may not feel a need to check OQ scores if they believe the patient is progressing adequately. However, previous work has suggested that “therapists tend to over-predict improvement and fail to recognize clients who worsen during therapy (Hannan et al., 2005, p. 161),” further reinforcing the need for treating therapists to review outcomes with *more* regularity, not less.

Regarding study question one, this study found that a therapist’s consistent use of the OQ feedback system had statistically, but not clinically significant predictive power for a patient’s change in OQ score. As only 0.23% of the variance in OQ score is accounted for by using the OQ feedback system, it would suggest that consistent usage of the OQ system does not have a large general effect on client outcome. This is in line with previous work examining this measure that has “repeatedly demonstrated...OQ feedback interventions appear to be effective in enhancing outcome for NOT patients, while having little impact on OT cases” (Shimokawa et al., 2010, p. 89; 2010). The present study reinforces the concept that outcome measures are primarily useful for detecting and modifying deteriorating patients, with little to no effects on cases demonstrating expected progress.

With that in mind, the current study’s finding of such minimal impact would seem to be under-representing the positive effects of outcome measurement for those NOT cases. Previous studies have suggested that NOT patients account for an average of 23% of observed cases (Harmon et al., 2007; Hawkins, 2004; Lambert et al., 2001; 2002; Shimokawa, 2010; Slade et al., 2008; Whipple et al., 2003). This highlights one of the major limitations of this study; namely,

that the OQ-45 feedback software does not accurately classify patients who have had previous courses of therapy at the same clinic, thus miscalculating the NOT or OT classification. At the clinic used for this study, about 22.4% of patients fit into this category of utilizing multiple courses of therapy.

Only one of the previous studies examining the use of this measure utilized the immediate electronic feedback system (Slade et al., 2008), and circumvented the NOT categorization problem by utilizing only new patients at their first course of therapy. To do so, they notified clinicians and clients that they were participating in a study that was focusing on utilization of the feedback system, thus introducing possible bias into the clinicians' use of the measure (as well as patients' expectations of the measure's use). The present study, in order to avoid this potential bias, used an archival, anonymous dataset that helped to rule out any bias on the part of the therapists using the measure. However, the combination of the current limitations of the OQ software and the use of an anonymous, archival dataset eliminated the ability to track which patients had utilized the center previously, introducing a possible NOT or OT misclassification for over a fifth of the patients. Thus, the present study was unable to correctly classify and examine the specific impact of naturalistic usage of outcome feedback on NOT cases, trading a repeated validation of the outcome measure for information on the naturalistic usage of the measure. Further examination of naturalistic feedback with correct OT/NOT classification could confirm or question previous findings of efficacy for those clients in a routine practice scenario. This is likely to be possible shortly, as consistent feedback to the corporation that administers the OQ-45 has resulted in planned enhancements to the electronic feedback system that would correctly classify patients with multiple courses of therapy.

The second question this study investigated was whether or not therapists who consistently

use outcome measures are more effective in their clinical practice. Previous work examining therapist efficacy highlights two areas; (a) the length of treatment, with more effective therapists helping patients improve in fewer sessions, and (b) decrease in reported negative symptomatology, in either the absolute amount of symptoms reduced, or the patient's level of symptomatology being reduced to below clinical levels, or both (Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009; Beutler, Machado, & Neufeldt, 2004; Garcia-Toro et al., 2013; Okiishi et al., 2006). This study found no correlation between therapist use of the OQ feedback and average treatment length, overall change in OQ scores, or average final OQ score for their patients.

The finding that use of OQ feedback is not significantly correlated with a difference in treatment length is consistent with four previous studies (Harmon et al., 2007; Lambert et al., 2001; 2002; Slade et al., 2008) but is not consistent with the findings in two others (Hawkins, 2004; Whipple et al., 2003). This discrepancy raises some question about the factors underlying change; however, it suggests that just using an outcome measure does not a good therapist make. Therapy is a complex interaction of variables, both within and outside of the therapist's control. The OQ in particular is not designed to be a measure of therapist efficacy, but rather is designed to alert therapists of patients who are not responding to treatment. An alternative explanation for this finding may be that, due to the low frequency with which the majority of therapists in this study utilized the feedback system; effects may only reach significance after a certain level of consistent usage.

Conclusion

The preceding sections presented some of the modern concerns facing the mental health profession and the case for practice-based evidence in that context. The philosophical and

theoretical rationale concerning outcome measurement was suggested to have particular efficacy for addressing those concerns. Past work has supported the practical implementation of such measurement in clinical trials and novel research settings. This study has added to that body of knowledge an examination of a setting where outcome measurement has become a routine part of clinical practice. The primary goal was to examine the actual utility of outcome feedback to patients individually and globally in a naturalistic setting.

Results indicated overall low utilization of outcome feedback by clinicians practicing at the center in question. Patients whose treating clinicians were among the minority who regularly utilized that feedback demonstrated statistically but not clinically significant change, and even those high-utilization therapists demonstrated no significant difference in the outcomes of their patients. Limitations of the current study were reported, including the problems in the current feedback system which nullifies the ability for the outcome measure to track deterioration—those patients for whom it has demonstrated the greatest clinical utility.

References

- American Psychological Association. (1995). *Template for developing guidelines: Interventions for mental disorders and psychosocial aspects of physical disorders*. Washington, DC: Author.
- American Psychological Association. (2002). Criteria for evaluating treatment guidelines. *The American Psychologist*, 57(12), 1052–1059.
- American Psychological Association. (2005, August 1). *APA Policy Statement on Evidence-Based Practice in Psychology*. *apa.org*. Retrieved September 11, 2012, from <http://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf>
- Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology*, 65(7), 755–768. doi:10.1002/jclp.20583
- Bailey, R. J. (2009, August). *Evaluating a new intervention for not-on-track clients: The use of repeated feedback to therapists and an abbreviated Clinical Support Tool measure*. (Unpublished doctoral dissertation). Brigham Young University, Provo, Utah.
- Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research*, 6(1), 3–15.
- Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment*, 82(1), 60–70.
- Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, 67(3), 588–597. doi:10.1207/s15327752jpa6703_13

- Beutler, E. L., Machado, P., & Neufeldt, A. S. (2004). Therapist variables. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change (4th Edition)* (4 ed.). New York, New York: Wiley.
- Burlingame, G. M., & Beecher, M. E. (2008). New directions and resources in group psychotherapy: Introduction to the issue. *Journal of Clinical Psychology, 64*(11), 1197–1205. doi:10.1002/jclp.20534
- de Beurs, E., Hollander-Gijsman, den, M. E., van Rood, Y. R., Van der Wee, N. J. A., Giltay, E. J., van Noorden, M. S., van der Lem, R., Van Fenema, E., & Zitman, F.G. (2011). Routine outcome monitoring in the Netherlands: Practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clinical Psychology & Psychotherapy, 18*(1), 1–12. doi:10.1002/cpp.696
- DeLeon, P. H., VandenBos, G. R., & Bulatao, E. Q. (1991). Managed mental health care: A history of the federal policy initiative. *Professional Psychology: Research and Practice, 22*(1), 15.
- Derogatis, L. R. (1996). SCL-90-R: Symptom Checklist-90-R: Administration, scoring, and procedures manual.
- Garcia-Toro, M., Rubio, J. M., Gili, M., Roca, M., Jin, C. J., Liu, S.-M., Bastianoni, C., Blanco, C. (2013). Persistence of chronic major depression: A national prospective study. *Journal of Affective Disorders, 151*(1), 306–312. doi:10.1016/j.jad.2013.06.013
- Hannan, C. C., Lambert, M. J., Harmon, C. C., Nielsen, S. L. S., Smart, D. W. D., Shimokawa, K. K., & Sutton, S. W. S. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology, 61*(2), 155–163. doi:10.1002/jclp.20108

- Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist–client feedback and clinical support tools. *Psychotherapy Research, 17*(4), 379–392.
- Hatfield, D. R., & Ogles, B. M. (2004). The Use of Outcome Measures by Psychologists in Clinical Practice. *Professional Psychology: Research and Practice, 35*(5), 485.
- Hatfield, D. R., & Ogles, B. M. (2007). Why Some Clinicians Use Outcome Measures and Others Do Not. *Administration and Policy in Mental Health and Mental Health Services Research, 34*(3), 283–291. doi:10.1007/s10488-006-0110-y
- Hawkins, E. J. (2004). The Therapeutic Effects of Providing Patient Progress Information to Therapists and Patients. *Psychotherapy Research, 14*(3), 308–327. doi:10.1093/ptr/kph027
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy. Efficacy, effectiveness, and patient progress. *The American Psychologist, 51*(10), 1059–1064.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19.
- Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy, 3*(4), 249–258.
- Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology, 61*(2), 165–174. doi:10.1002/jclp.20113

- Lambert, M. J., Morton, J. J., Hatfield, D. R., Harmon, C., Hamilton, E. G., Reid, R. C., et al. (2004). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire)*. Salt Lake City, UT: American Professional Credentialing Services.
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., & Nielsen, S. L. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research, 11*(1), 49–68.
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology & Psychotherapy, 9*(2), 91–103.
doi:10.1002/cpp.324
- Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychology, 66*(2), 400–410.
- Mellor-Clark, J., Barkham, M., Connell, J., & Evans, C. (1999). Practice-based evidence and standardized evaluation: Informing the design of the CORE system. *Journal of Psychotherapy & Counselling, 2*(3), 357–374. doi:10.1080/13642539908400818
- Okiishi, J. C., Lambert, M. J., Eggett, D., Nielsen, L., Dayton, D. D., & Vermeersch, D. A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients' psychotherapy outcome. *Journal of Clinical Psychology, 62*(9), 1157–1172.
- Pirkis, J., Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D., & Dickson, R. (2005). Routine measurement of outcomes in Australia's public sector mental health services. *Australia and New Zealand health policy, 2*(1), 8. doi:10.1186/1743-8462-2-8

- Reed, G. M., & Eisman, E. J. (2006). Uses and Misuses of Evidence: Managed Care, Treatment Guidelines, and Outcomes Measurement in Professional Practice. In C. D. Goodheart, A. E. Kazdin, & R. J. Sternberg (Eds.), *Evidence-Based Psychotherapy: Where Practice and Research Meet* (1st ed.). Washington, DC: American Psychological Association.
- Shimokawa, K. (2010, August). *A Patient-Focused Psychotherapy Quality Assurance System: Meta-Analytic and Multilevel Analytic Review*. (Unpublished doctoral dissertation). Brigham Young University, Provo, Utah.
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78(3), 298–311. doi:10.1037/a0019247
- Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy*, 15(5), 287–303. doi:10.1002/cpp.594
- Slife, B. D., & Whoolery, M. (2006). Are psychology's main methods biased against the worldview of many religious people. *Journal of Psychology and Theology*, 34(3), 217–231.
- Sue, S., Zane, N., Nagayama Hall, G. C., & Berger, L. K. (2009). The case for cultural competency in psychotherapeutic interventions. *Annual review of psychology*, 60, 525–548. doi:10.1146/annurev.psych.60.110707.163651
- Tanenbaum, S. J. (1999). Evidence and expertise: The challenge of the outcomes movement to medical professionalism. *Academic medicine : Journal of the Association of American Medical Colleges*, 74(7), 757–763.

- Umphress, V. J., Lambert, M. J., Smart, D. W., Barlow, S. H., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment, 15*(1), 40–55.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological bulletin, 130*(4), 631.
- Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology, 50*(1), 59.

Table 1

Correlations with the Percentage a Therapist Views Their Patients' OQ Results

	Avg % Reviewed	<i>p</i>
AvgTreat	-0.075	0.43
AvgLastOQ	0.019	0.843
AvgΔOQ	0.077	0.417

Avg%Reviewed - Average percentage of patient's OQ feedback viewed by treating therapist

AvgTreat - Average treatment length (# of sessions) of that therapist's patients

AvgLastOQ - Average final OQ score of that therapist's patients

AvgΔOQ - Average change in final OQ score of that therapist's patients

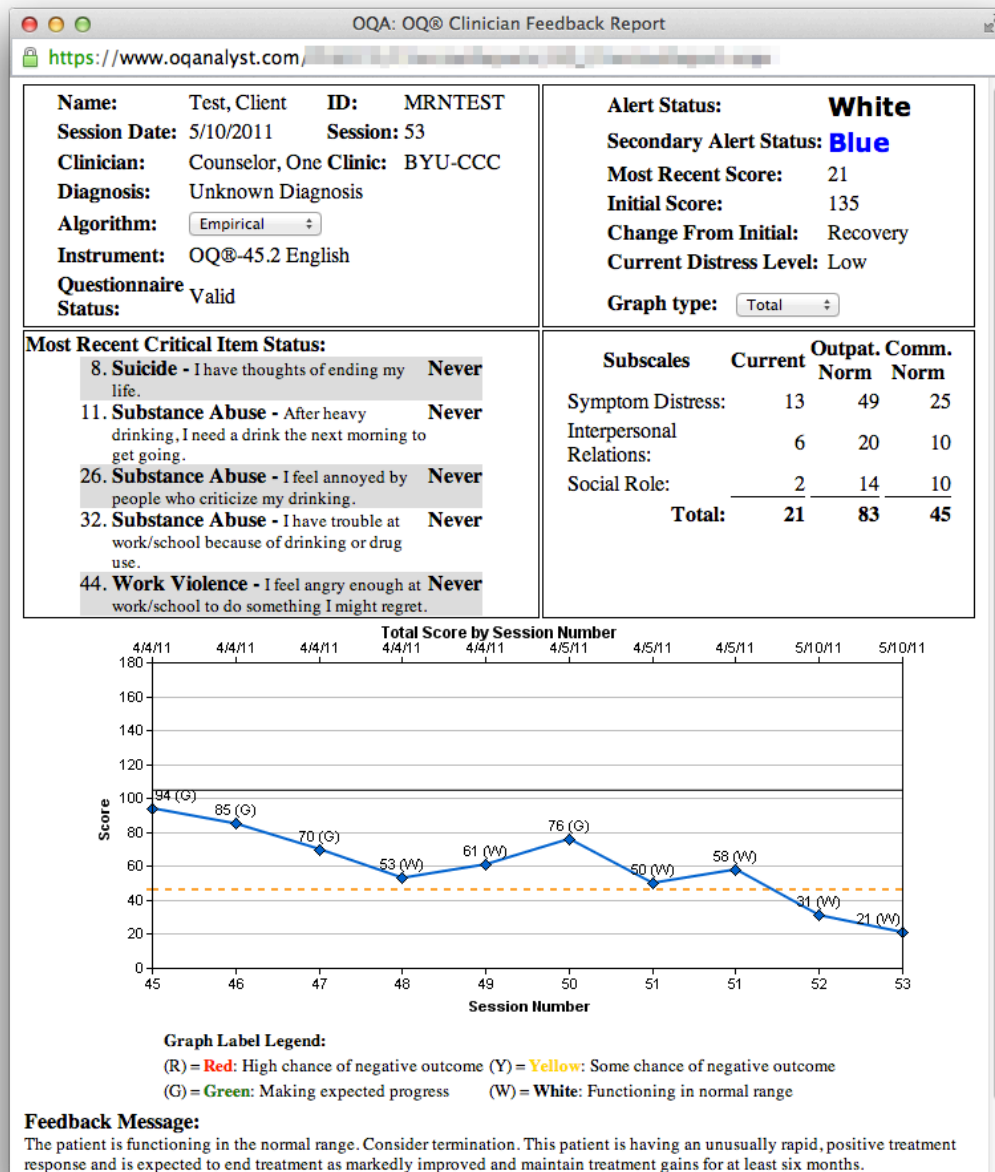


Figure 1. Sample OQ-45 digital feedback.

Appendix: Literature Review

The use of outcome measures in psychotherapy practice is a result of decades of philosophical, political, and economic influences. Evaluating the efficacy of those measures requires an understanding of the context in which their development occurred.

Historical/Political Influence on Outcome Measurement Development

The Health Maintenance Organization Act of 1973 officially encouraged the development of HMOs in the United States, providing qualifications to receive federal funding and circumvent some of the more restrictive state legislation (US Congress, 1974). HMOs were initially proposed to be a way to provide higher quality healthcare services while keeping medical costs down. However, as amendments were made to the federal code, HMOs' focus began to shift more heavily to the cost-reduction benefits, incentivizing the for-profit companies to provide only the treatments that were deemed absolutely necessary (DeLeon, VandenBos, & Bulatao, 1991). Government influence was increased with the creation of the Agency for Health Care Policy and Research (since renamed as the Agency for Healthcare Research and Quality (AHRQ)) in 1989; a branch of the U.S. Department of Health and Human Services whose work is to enable "services of the highest quality, with the best possible outcomes, at the lowest cost" (AHRQ, 2012, p. 1). These political influences further increased the economic pressure on HMOs, and created tension between themselves and their contracted healthcare providers, who felt their clinical experience devalued and their practice options increasingly restricted by these large organizations (Tanenbaum, 1999). Further increasing this tension, Wennberg (1984) found significant variation in rates for surgical and other medical practices, all across the globe; suggesting that healthcare professionals' judgment was inconsistent and needed to be more closely regulated. The combination of increased financial and policy incentive to keep costs

down and decreased confidence in those charged with supplying healthcare treatment made for an increasingly hostile and combative environment for healthcare providers (DeLeon et al., 1991; Hannan, Lambert, Harmon, Nielsen, Smart, Shimokawa & Sutton, 2005; Reed & Eisman, 2006; Tanenbaum, 1999; Whipple et al., 2003).

It is out of this politically and financially motivated climate that the evidence based practice (EBP) movement began. Wennberg's work, along with the political climate at the time, moved EBP into the spotlight as what Tannenbaum (Hatfield & Ogles, 2007; Tanenbaum, 1999) called a *public idea*. A public idea is a response to a public concern that, while not untrue, focuses public attention (and policy) on what is only a small slice of a large, complex problem. Some HMOs and other policy makers, faced with spiraling healthcare costs and further public scrutiny on their results, adopted the idea that these problems were caused by under-informed practice by health professionals. It was thought that bringing practice more in conformity with research through EBP would minimize these problems (Burlingame & Beecher, 2008; DeLeon et al., 1991; Reed & Eisman, 2006; Tanenbaum, 2003).

This scrutiny towards healthcare decisions was not directed solely at the medical field; the inclusion of psychologists on the lists of authorized HMO health professionals in the Health Maintenance Organization Amendments of 1986 (DeLeon et al., 1991; Tanenbaum, 1999; US Congress, 1982) allowed for those same pressures to apply specifically to mental health professionals. This evinced questions about the efficiency and efficacy of psychotherapeutic treatments. HMOs sought to trim costs by filtering out unnecessary or inefficient practices and incentivizing EBP by their contracted psychologists. In an effort to be sure such questions were asked and answered by those most qualified to do so, the American Psychological Association (APA) and other professional mental health organizations (AHRQ, 2012; American

Psychological Association, 1995; 2002a; 2005; Burlingame & Beecher, 2008), established task forces to set guidelines for examining and determining best practices (American Psychological Association, 1995; 2002a; 2006; Slife & Whoolery, 2006; Tanenbaum, 1999). These task forces outlined two primary criteria for scientifically evaluating treatments: *treatment efficacy*, or whether or not the treatment in question actually works; and *clinical utility*, or the usefulness of a given treatment for a specific context (American Psychological Association, 2002a; Sue et al., 2009; Westen et al., 2004). Treatment efficacy is primarily established as a function of the rigor of the scientific method employed by researchers examining a treatment plan, with randomized, controlled trials as the prescribed ideal methodology. Clinical utility depends on the generalizability of the researched practice to contexts outside of the rigorous experimental conditions.

These guidelines helped to standardize what was considered evidence-based practice and what was not, and helped to give clinicians and researchers a common language to discuss their results. However, this clarifying of the psychological application of the EBP paradigm also elucidates some of the both philosophical and practical problems inherent in the underlying assumptions of the framework. The ‘evidence’ of EBP is based on the philosophical principles of empiricism and reductionism/universality, which have problematic assumptions when applied to the social science of psychology.

Philosophical Constraints Overcome by Outcome Measurement

Empiricism argues that only that which can be observed can be knowable (Slife & Whoolery, 2006). This epistemology gives rise to the need to operationalize unobservable psychological constructs like depression, intelligence, or anxiety to be able to study them in a “carefully controlled empirical evaluation” as called for by the APA EBP criteria (American

Psychological Association, 2002a; Hannan et al., 2005). Operationalizing constructs removes the researcher a step from the actual phenomena; for example, researchers studying the concept of anxiety are forced to examine physical manifestations of anxiety, such as pulse, galvanic skin response, or breathing rate. Rather than the lived meaning and experience of anxiety, only its observable aspects can be studied by empirical science. This limitation restricts both the kind of knowledge that EBP can provide, as well as the applicability of that knowledge, to a detached and less-meaningful realm of understanding.

In addition to the restrictions of empiricism, EBP is also based on a reductionistic understanding of human experience. This philosophical assumption posits that “all change is ultimately reducible to, or governed by, unchangeable natural laws and principles” (Barkham et al., 2006; Slife & Whoolery, 2006), and that these natural laws will consistently apply to all cases of change. Here, too, the APA’s own guidelines demonstrate a belief that human behavior is reducible to broadly applicable principles made evident by their replicability:

The efficacy of interventions included in treatment guidelines can be supported by *multiple observations* by trained, knowledgeable, and experienced individuals...*consensus is more compelling than individual observation* but less compelling than carefully controlled empirical evaluation [emphasis added] (American Psychological Association, 2002a; Lambert et al., 2005; Slade et al., 2008; Whipple et al., 2003).

The idea that best practice in psychology is going to be universally consistent across individual cases minimizes the influence of personality and situational effects, which have significant modifying effects for most Axis I disorders (Blatt & Zuroff, 1992; de Beurs et al., 2011; Hatfield & Ogles, 2004; Mellor-Clark et al., 1999; Pirkis et al., 2005; Westen et al., 2004). For example,

Westen et al. (Hatfield & Ogles, 2007; Westen, Muderrisoglu, Fowler, Shedler, & Koren, 1997) found significant differences in depression between subjects diagnosed with Borderline Personality Disorder (BPD) Narcissistic PD, and Dysthymic patients. Others have shown clear distinctions in depressive symptoms and onset based on non-pathological personality differences in sociotropic motivation (social attachment/influence) versus autonomic motivation (personal achievement/environmental mastery) (Hammen, Ellicott, Gitlin, & Jamison, 1989; Hatfield & Ogles, 2004; Kwon & Whisman, 1998). Evidence-based practice's philosophical foundation of reductionism minimizes the importance of individual differences in mental health disorders.

Practical Constraints Alleviated by Outcome Measurement

In addition to the philosophical problems with the wholesale use of EBP, there are significant practical concerns. For example, research into the efficacy of treatments for ethnic minorities are severely underrepresented (Berger, 2010; Constantine, Miville, & Kindaichi, 2008; Lambert et al., 2001; 2001; Sue et al., 2009). A review examining 379 NIMH-funded clinical trials found that less than half of the studies included racial/ethnic information (Hawkins, 2004; Lambert et al., 2002; Mak, Law, Alvidrez, & Pérez-Stable, 2007; Whipple et al., 2003). This paucity of multiculturally competent information—a focus for psychology as a discipline (American Psychological Association, 2002b; Lambert et al., 2004)—reveals another of the weaknesses of EBP; namely, that keeping track of and investigating what is empirically supported for distinct populations with even a small subset of disorders is time intensive and Sisyphean in scope (Lambert et al., 2004; Umphress et al., 1997; Westen et al., 2004). Additionally, EBP's reliance on empiricism limits the kinds of treatments that can be evaluated—with a dependence on manualization in order to produce replicable, uniform treatment. Manualized treatments rely heavily on a relationship between therapist and patient

where the therapist sets the agenda and the topic of the session, replacing the collaborative relationship with one in which the therapist applies interventions to the patient in the same manner that a mechanic changes the oil in a car (Westen et al., 2004). This limits the kinds of therapeutic approaches that are available to be manualized (and therefore empirically studied), emphasizing the utility of short-form, directive treatments over long-term collaborative or exploratory theories. This format also discounts the significant proportion of variance explained by common factors like the therapeutic relationship (Messer & Wampold, 2002). Moreover, the wide variability demonstrated in the relevant literature indicates that therapists' adherence to the manually proscribed interventions is only inconsistently correlated with outcome (Barber et al., 2006; Bauer et al., 2004; Castonguay, Goldfried, Wisner, Raue, & Hayes, 1996; Loeb et al., 2005; Lunnen & Ogles, 1998).

These problems with empirical evidence based practice help explain why implementing EBP across the discipline has been a controversial and hotly debated topic (Beck, Steer, Ball, & Ranieri, 1996b; Burlingame & Beecher, 2008). Acknowledging the limited applicability of closely controlled, RCT research findings on the highly variable treatment environment in which most psychologists work, Howard et al. (Derogatis, 1996; Howard et al., 1996) proposed a different framework on which to examine best practice in psychology: the patient-focused research (PFR) paradigm. Rather than focusing on the broad efficacy (Is this treatment better than placebo?) or clinical utility (Is this treatment generalizable and feasible for broad application?) of a particular therapeutic practice, the core question that PFR attempts to answer is, "Is this patient's condition responding to the treatment that is being applied (Howard et al., 1996; 1996)?"

Patient-Focused Research Paradigm as Precursor to Outcome Measurement

Measuring patients' individual symptoms before, during, and after treatment is the primary focus of the PFR paradigm. These measurements are typically termed *outcome measures*, and provide evidence on which practice can be based while avoiding some of the more grievous philosophical restraints of the empirically supported EBP movement. Examining symptom progression regularly during treatment allows clinicians to “determine the appropriateness of the current treatment . . . [and] the need for further treatment . . . [and to] prompt a clinical consultation for patients who [are] not progressing at expected rates” for each individual patient's specific response to treatment (Howard et al., 1996; Slade et al., 2008). This avoids the reductionistic assumption that a given treatment will work for most patients and immediately informs the clinician if the treatment is appropriate and beneficial for this individual patient in their current context. Outcome measures also limit the distance imposed by empirical operationalization. By asking questions about the patient's current, lived experience, there is no need to redefine a construct to make it more easily measured; outcome measures simply ask a patient to describe their anger, sadness, or frustration based on their own experience of the feeling.

In addition to the philosophical benefits, the individually applicable nature of the PFR paradigm has practical benefits for HMOs, practitioners, and patients. HMOs are tasked by practicality and their stockholders to find the most efficient and effective treatment providers; often this is done by preparing (or adopting) a list of approved treatments for specific mental health disorders, and requiring contracted service providers to use only those catalogued treatments (Chambless et al., 1998; Harmon et al., 2007; Hawkins, 2004; Lambert et al., 2002; Whipple et al., 2003). However, as research expands and examines new modalities, such lists

require constant revision to remain current and relevant to practitioners. By having mental health professionals report outcome measures, costly and incomprehensive lists of treatments can be avoided, as “practitioners with good records of outcomes would in essence have proved their effectiveness with the actual patient population they are being paid to treat” (Reed & Eisman, 2006, p. 25). This helps alleviate the efficiency burden on HMOs, and allows more leeway for clinicians while still maintaining a high level of efficacy and accountability. This individual reporting also neatly eliminates the risks of patient/treatment mismatch (Norcross, 2001; Shimokawa et al., 2010), which has generally demonstrated negative effects on outcome (Atkinson, Worthington, Dana, & Good, 1991; Dunlop et al., 2012; Goates-Jones & Hill, 2008; Karno & Longabaugh, 2007; Whipple et al., 2003), and which requires costly case-by-case exception administration by HMOs and practitioners (Bailey, 2009; Hays et al., 2002; Zane et al., 2005).

Benefits of Outcome Measures

In addition to their practical and financial benefits, outcome measures have excellent therapeutic advantages. Outcome measures enable practicing clinicians to reliably detect treatment failure; a skill that both experienced and inexperienced therapists have been shown to lack (Hannan et al., 2005; Yalom & Lieberman, 1971). Hannan et al. (2005) asked 40 therapists seeing 550 patients over three weeks if they believed that their patient was deteriorating at the end of each session. Despite being informed of the historical failure rate for their treatment center (8%), the purpose of the study, and the criteria for what was deemed failure by the outcome measure, only three patients were predicted to deteriorate. Only one of those predicted three actually deteriorated, and the clinicians did not predict the 39 other patients who actually deteriorated. Additionally, feedback about predicted treatment failures has consistently

demonstrated a significant positive effect in helping therapists and patients to adjust the therapeutic interventions and improve therapy outcome (Harmon et al., 2007; Hawkins, 2004; Lambert et al., 2001; 2002; Shimokawa, 2010; Slade et al., 2008; Whipple et al., 2003). These five studies using the OQ-45 (Lambert et al., 2004; Slade et al., 2008) were heavily based in real-world practices: there was no specified diagnosis or lack of comorbidity; a wide variety of treatment orientations and therapist experience levels were included; patient ages and symptom severity levels ranged widely; the length of therapy was determined by the therapist and the patient, rather than by insurance mandate or research design; the outcome measure and algorithms detecting progress/failing cases were identical across all five studies; and each therapist saw both Feedback patients and No-Feedback patients, randomly assigned—minimizing therapist effects on outcome differences. In all five studies, therapists who received feedback about cases predicted to suffer treatment failure had significantly better outcomes than those who did not—reducing failure rates between 5-12% (Anderson et al., 2009; Beutler et al., 2004; Garcia-Toro et al., 2013; Lambert, 2007; Okiishi et al., 2006). Additionally, feedback resulted in more cost-effective service delivery, as patients who were progressing as expected towards a good outcome attended significantly fewer sessions (Harmon et al., 2007; Lambert et al., 2001; 2002; Slade et al., 2008). Formal treatment monitoring through the use of outcome measures has been shown to be a positive addition to therapists' abilities to help all of their patients.

These benefits have prompted clinicians worldwide to adopt systems of outcome measurement. In Australia, a national program of outcome measurement and data collection has been in place in every State/Territory since 2003, using the standardized clinician-report Health of the Nation Outcomes Scale (Hawkins, 2004; Pirkis et al., 2005; Whipple et al., 2003). The

Netherlands' Leiden University Medical Center (LUMC), in collaboration with local mental health clinics, began routine outcome monitoring (ROM) with a comprehensive battery of patient-report and therapist-report measures in 2002 (de Beurs et al., 2011; US Congress, 1974). Around the same time, a multidisciplinary group of mental health practitioners and researchers in the UK developed a standardized, patient-completed measure, the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) (Barkham et al., 2006; DeLeon et al., 1991; Mellor-Clark et al., 1999). A German program of outcome tracking, the Stuttgart-Heidelberg Model, was constructed and introduced in 2001 (AHRQ, 2012; Kordy, Hannöver, & Richard, 2001). And in the US, 1996 saw the release of the most current version of some of the most popular adult outcome measures used by American clinicians (Hatfield & Ogles, 2004; Tanenbaum, 1999); the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996a), the Symptom Checklist-90 (SCL-90; Derogatis, 1996), and the Outcome Questionnaire-45 (OQ-45; Lambert et al., 1996). These patient outcome measurement procedures offer researched and individually pertinent evidence on which patient focused research and practice can be based.

Therapist Resistance to Outcome Measures

Despite the international adoption of outcome measures, a majority of practicing therapists in the United States are not implementing routine outcome measurement into their regular practice. In a random survey of APA members, Hatfield and Ogles (2007) received 874 responses to a survey including the question “Do you use some form of outcome assessment in your practice?” The majority of respondents (62.9%) stated that they did not. The most commonly cited responses for not using outcome measures were practical in nature, i.e. “Adds too much paperwork” or “Do not have enough resources (money, personnel).” Despite the clear

agreement in the researched literature on the importance of outcome measures, nearly two-thirds of clinicians are not using those tools in their regular practice.

These naturalistic survey results demonstrate a particular distance from the research in the field, as implementing outcome measures into routine practice and research is consistently being called for in the literature. Lambert and Shimokawa (2011), after reviewing meta-analyses of two different outcome measures, strongly suggest that, “clinicians seriously consider making formal methods of collecting patient feedback a routine part of their daily practice (p. 1).” Bobbitt (2012) echoes the APA’s (2006) call for further research and examination into “providing clinicians with real-time patient feedback to benchmark progress in treatment and clinical support tools to adjust treatment as needed (p. 278).” And the APA further emphasized the importance of these tools by establishing a resource for member clinicians to be able to review the majority of outcome measurement instruments available for use (American Psychological Association, 2011).

While researchers and professional organizations call for private practitioners to implement outcome measures into their daily practice routines, practitioner responses demonstrate some of the distinctions between regular therapeutic practice and the research that ideally supports it. Research environments are often significantly different from naturalistic settings across a host of variables, including patient heterogeneity/severity, supervision and training of clinicians, typical treatment length (dosage), patient assignment, clinical caseload, and funding structure (Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009; Minami et al., 2008; Nathan, Stuart, & Dolan, 2000; Wampold & Brown, 2005). The more rigid a research paradigm, the more difficult it is to translate its results to real-world settings (Persons & Silberschatz, 1998). While these drawbacks to explicit research methodologies certainly do not eliminate the

clinical utility or knowledge gained from such investigation, they do present a perspective that may be somewhat removed from the reality of how a large majority of mental health professionals practice their trade.

Additionally, there is a significant gap between experimental therapy outcomes and naturalistic therapy outcomes (Barkham et al., 2008; de Jong, van Sluis, Nugter, Heiser, & Spinhoven, 2012). All of the yet reported research into outcome measurement uses explicit research methodologies to compare outcome measures groups to no-outcome measures groups. For example, Lambert et al. assigned patients to feedback conditions either by random assignment (Lambert et al., 2001) or based on semester enrollment, with patients "...assigned to the Feedback condition if they attended school during Winter Semester or Spring Term 2000, or to the No Feedback condition if they attended school Summer Term or Fall Semester 1999 (Lambert et al., 2002)." Whipple and colleagues (2003) randomly assigned patients to feedback or no feedback conditions, and allowed therapists receiving feedback to choose to receive additional feedback for patients at risk of deteriorating. Harmon et al. (2007) randomly separated participants into feedback conditions, and then compared those conditions with an archival control dataset. By the nature of the design used in these studies, participating therapists are all made explicitly aware that their work is being scrutinized. As therapists participating in research often have stronger allegiances to the studied phenomena (Luborsky et al., 1999; Shadish & Baldwin, 2005), this may significantly impact outcome measures research results.

Present Study Goals

The present study was an effort to respond to prior research reviews calling for further naturalistic examination (American Psychological Association, 2006; Bobbitt et al., 2012; Lambert & Shimokawa, 2011), to reiterate the strengths of routine outcome measurement as

effective evidence based practice, and to provide data from a naturalistic practice environment more removed from the biases of explicit randomized-control research. To do so, it examined routine outcome measurement usage by clinicians in a large college counseling center where outcome measurement is a standard part of routine practice. This is the next step from the previously mentioned studies using the OQ-45, as it includes two primary differences regarding both setting and methodology. First, because all patients at this clinic participated in outcome assessment, patients were not randomly assigned to outcome feedback (FB) or no feedback (NFB) conditions. This comparison was determined by the actual usage of the outcome measures feedback by the treating therapist. Second, therapists were not made explicitly aware that their outcome measurement usage was being tracked for the purposes of the present study. This provided further insight into how clinicians actually use outcome measures. By examining the software logs of the electronically administered outcome measure (OQ-45), the present study was able to examine naturalistic usage patterns of outcome measures, and the impact such usage had on patient outcome.

The use of software logs eliminated any reliance on self-report measures of utility or usage frequency as used in the past (Lambert et al., 2001), allowing for completely naturalistic examination of outcome use by clinicians. At the current site, every patient is asked to take the OQ-45 at every session; thus, treatment as usual (TAU) *includes* outcome measurement. This is beneficial in the context of the current study as it pertains to future applicability. Hatfield and Ogles (2004) cite a trend of increased usage of outcome measures among clinicians:

In contrast to earlier surveys, a surprising 37% of respondents reported gathering outcome data of one form or another. This is higher than the 29% of respondents in the CAPP (Phelps, Eisman, & Kohout, 1998) study and the 23% who reported using

standardized measures in Bickman et al.'s (2000) study of clinicians involved in practice with children or adolescents...this change might represent a shift in the field toward an increased use of outcome measures as part of clinical practice. If the current trend continues, an increasing number of practitioners will be using outcome measures in the future (p. 488-489).

As more and more psychologists implement routine outcome monitoring into their practice, we will start to see this as the norm, and understanding how to most effectively use this hopefully ubiquitous resource will become increasingly important.

References

- Agency for Healthcare Research and Quality. *Budget Estimates for Appropriations Committees, Fiscal Year 2013.*, Retrieved September 11, 2012, from <http://www.ahrq.gov/about/ataglace.htm>. Rockport, MD: Agency for Healthcare Research and Quality.
- American Psychological Association. (1995). Template for developing guidelines: Interventions for mental disorders and psychosocial aspects of physical disorders. *Washington, DC: Author.*
- American Psychological Association. (2002a). Criteria for evaluating treatment guidelines. *The American Psychologist, 57*(12), 1052–1059.
- American Psychological Association. (2002b). Ethical principles of psychologists and code of conduct. *The American Psychologist, 57*(12), 1060–1073.
- American Psychological Association. (2006). Evidence-based practice in psychology. *The American Psychologist, 61*(4), 271–285. doi:10.1037/0003-066X.61.4.271
- American Psychological Association. (2005, August 1). APA Policy Statement on Evidence-Based Practice in Psychology. *apa.org*. Retrieved September 11, 2012, from <http://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf>
- American Psychological Association. (2011, August 29). APA launches new outcome measurement database. *apapracticecentral.org*. Retrieved October 10, 2012, from <http://www.apapracticecentral.org/update/2011/08-29/measurement-database.aspx>
- Anderson, T., Ogles, B. M., Patterson, C. L., Lambert, M. J., & Vermeersch, D. A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65*(7), 755–768. doi:10.1002/jclp.20583

- Atkinson, D. R., Worthington, R. L., Dana, D. M., & Good, G. E. (1991). Etiology beliefs, preferences for counseling orientations, and counseling effectiveness. *Journal of Counseling Psychology, 38*(3), 258.
- Bailey, R. J. (2009, August). *Evaluating a new intervention for not-on-track clients: The use of repeated feedback to therapists and an abbreviated Clinical Support Tool measure*. (Unpublished doctoral dissertation). Brigham Young University, Provo, Utah.
- 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.
doi:10.1037/a0015235
- Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., & Gibbons, M. B. C. (2006). The role of therapist adherence, therapist competence, and alliance in predicting outcome of individual drug counseling: Results from the National Institute Drug Abuse Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16*(02), 229–240.
- Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research, 6*(1), 3–15.
- Barkham, M., Stiles, W. B., Connell, J., Twigg, E., Leach, C., Lucock, M., et al. (2008). Effects of psychological therapies in randomized trials and practice-based studies. *The British journal of clinical psychology / the British Psychological Society, 47*(Pt 4), 397–415.
doi:10.1348/014466508X311713
- Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment, 82*(1), 60–70.

- Beck, A. T., Steer, R. A., & Brown, G. K. (1996a). Manual for the Beck depression inventory-II. *San Antonio, TX: Psychological Corporation, 1*, 82.
- Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996b). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment, 67*(3), 588–597. doi:10.1207/s15327752jpa6703_13
- Berger, R. (2010). EBP Practitioners in Search of Evidence. *Journal of Social Work, 10*(2), 175–191.
- Beutler, E. L., Machado, P., & Neufeldt, A. S. (2004). Therapist variables. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change (4th Edition)* (4 ed.). New York, New York: Wiley.
- Bickman, L., Rosof-Williams, J., Salzer, M. S., Summerfelt, W. T., Noser, K., Wilson, S. J., & Karver, M. S. (2000). What information do clinicians value for monitoring adolescent client progress and outcomes? *Professional Psychology: Research and Practice, 31*(1), 70.
- Blatt, S. J., & Zuroff, D. C. (1992). Interpersonal relatedness and self-definition: Two prototypes for depression. *Clinical Psychology Review, 12*(5), 527–562.
- Bobbitt, B. L., Cate, R. A., Beardsley, S. D., Azocar, F., & McCulloch, J. (2012). Quality improvement and outcomes in the future of professional psychology: Opportunities and challenges. *Professional Psychology: Research and Practice, 43*(6), 551.
- Burlingame, G. M., & Beecher, M. E. (2008). New directions and resources in group psychotherapy: Introduction to the issue. *Journal of Clinical Psychology, 64*(11), 1197–1205. doi:10.1002/jclp.20534

- Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64*(3), 497–504.
- Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., et al. (1998). Update on empirically validated therapies, II. *Clinical Psychologist, 51*(1), 3–16.
- Constantine, M. G., Miville, M. L., & Kindaichi, M. M. (2008). Multicultural competence in counseling psychology practice and training. In S. D. Brown & R. W. Lent (Eds.), *Handbook of Counseling Psychology* (pp. 141–158). New York: Handbook of Counseling Psychology.
- de Beurs, E., Hollander-Gijsman, den, M. E., van Rood, Y. R., Van der Wee, N. J. A., Giltay, E. J., van Noorden, M. S., et al. (2011). Routine outcome monitoring in the Netherlands: Practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clinical Psychology & Psychotherapy, 18*(1), 1–12. doi:10.1002/cpp.696
- de Jong, K., van Sluis, P., Nugter, M. A., Heiser, W. J., & Spinhoven, P. (2012). Understanding the differential impact of outcome monitoring: Therapist variables that moderate feedback effects in a randomized clinical trial. *Psychotherapy Research, 22*(4), 464–474. doi:10.1080/10503307.2012.673023
- DeLeon, P. H., VandenBos, G. R., & Bulatao, E. Q. (1991). Managed mental health care: A history of the federal policy initiative. *Professional Psychology: Research and Practice, 22*(1), 15.
- Derogatis, L. R. (1996). *SCL-90-R: Symptom Checklist-90-R: Administration, scoring, and procedures manual*. NCS Pearson.

Dunlop, B. W., Kelley, M. E., Mletzko, T. C., Velasquez, C. M., Craighead, W. E., & Mayberg, H. S. (2012). Depression beliefs, treatment preference, and outcomes in a randomized trial for major depressive disorder. *Journal of psychiatric research, 46*(3), 375–381.

doi:10.1016/j.jpsychires.2011.11.003

Garcia-Toro, M., Rubio, J. M., Gili, M., Roca, M., Jin, C. J., Liu, S.-M., et al. (2013). Persistence of chronic major depression: A national prospective study. *Journal of Affective Disorders, 151*(1), 306–312. doi:10.1016/j.jad.2013.06.013

Goates-Jones, M., & Hill, C. E. (2008). Treatment preference, treatment-preference match, and psychotherapist credibility: Influence on session outcome and preference shift.

Psychotherapy: Theory, Research, Practice, Training, 45(1), 61.

Hammen, C., Ellicott, A., Gitlin, M., & Jamison, K. R. (1989). Sociotropy/autonomy and vulnerability to specific life events in patients with unipolar depression and bipolar disorders. *Journal of abnormal psychology, 98*(2), 154–160.

Hannan, C. C., Lambert, M. J., Harmon, C. C., Nielsen, S. L. S., Smart, D. W. D., Shimokawa, K. K., & Sutton, S. W. S. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology, 61*(2), 155–163. doi:10.1002/jclp.20108

Harmon, S. C., Lambert, M. J., Smart, D. M., Hawkins, E., Nielsen, S. L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist–client feedback and clinical support tools. *Psychotherapy Research, 17*(4), 379–392.

Hatfield, D. R., & Ogles, B. M. (2004). The Use of Outcome Measures by Psychologists in Clinical Practice. *Professional Psychology: Research and Practice, 35*(5), 485.

- Hatfield, D. R., & Ogles, B. M. (2007). Why Some Clinicians Use Outcome Measures and Others Do Not. *Administration and Policy in Mental Health and Mental Health Services Research, 34*(3), 283–291. doi:10.1007/s10488-006-0110-y
- Hawkins, E. J. (2004). The Therapeutic Effects of Providing Patient Progress Information to Therapists and Patients. *Psychotherapy Research, 14*(3), 308–327. doi:10.1093/ptr/kph027
- Hays, K. A., Rardin, D. K., Jarvis, P. A., Taylor, N. M., Moorman, A. S., & Armstead, C. D. (2002). An exploratory survey on empirically supported treatments: Implications for internship training. *Professional Psychology: Research and Practice, 33*(2), 207.
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy. Efficacy, effectiveness, and patient progress. *The American Psychologist, 51*(10), 1059–1064.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19.
- Karno, M. P., & Longabaugh, R. (2007). Does matching matter? Examining matches and mismatches between patient attributes and therapy techniques in alcoholism treatment. *Addiction, 102*(4), 587–596. doi:10.1111/j.1360-0443.2007.01754.x
- Kordy, H., Hannöver, W., & Richard, M. (2001). Computer-assisted feedback-driven quality management for psychotherapy: The Stuttgart-Heidelberg Model. *Journal of Consulting and Clinical Psychology, 69*(2), 173–183.
- Kwon, P., & Whisman, M. A. (1998). Sociotropy and autonomy as vulnerabilities to specific life events: Issues in life event categorization. *Cognitive Therapy and Research, 22*(4), 353–362.

- Lambert, M. J. (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. doi:10.1080/10503300601032506
- Lambert, M. J., & Shimokawa, K. (2011). Collecting client feedback. *Psychotherapy: Theory, Research, Practice, Training, 48*(1), 72–79. doi:10.1037/a0022238
- Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy, 3*(4), 249–258.
- Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology, 61*(2), 165–174. doi:10.1002/jclp.20113
- Lambert, M. J., Morton, J. J., Hatfield, D. R., Harmon, C., Hamilton, E. G., Reid, R. C., et al. (2004). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire)*. Salt Lake City, UT: American Professional Credentialing Services.
- Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., & Nielsen, S. L. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychotherapy Research, 11*(1), 49–68.
- Lambert, M. J., Whipple, J. L., Vermeersch, D. A., Smart, D. W., Hawkins, E. J., Nielsen, S. L., & Goates, M. (2002). Enhancing psychotherapy outcomes via providing feedback on client progress: A replication. *Clinical Psychology & Psychotherapy, 9*(2), 91–103. doi:10.1002/cpp.324

Loeb, K. L., Wilson, G. T., Labouvie, E., Pratt, E. M., Hayaki, J., Walsh, B. T., et al. (2005).

Therapeutic alliance and treatment adherence in two interventions for bulimia nervosa: A study of process and outcome. *Journal of Consulting and Clinical Psychology, 73*(6), 1097–1107. doi:10.1037/0022-006X.73.6.1097

Luborsky, L., Diguier, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., et al.

(1999). The researcher's own therapy allegiances: A 'wild card' in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice, 6*(1), 95–106.

Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable

change. *Journal of Consulting and Clinical Psychology, 66*(2), 400–410.

Mak, W. W. S., Law, R. W., Alvidrez, J., & Pérez-Stable, E. J. (2007). Gender and Ethnic

Diversity in NIMH-funded Clinical Trials: Review of a Decade of Published Research.

Administration and Policy in Mental Health and Mental Health Services Research, 34(6), 497–503. doi:10.1007/s10488-007-0133-z

Mellor-Clark, J., Barkham, M., Connell, J., & Evans, C. (1999). Practice-based evidence and

standardized evaluation: Informing the design of the CORE system. *Journal of*

Psychotherapy & Counselling, 2(3), 357–374. doi:10.1080/13642539908400818

Messer, S. B., & Wampold, B. E. (2002). Let's face facts: Common factors are more potent than

specific therapy ingredients. *Clinical Psychology: Science and Practice, 9*(1), 21–25.

Retrieved from <http://onlinelibrary.wiley.com/store/10.1093/clipsy.9.1.21/asset/>

[clipsy.9.1.21.pdf?v=1&t=hs1yf18&s=6e5864549034ca1b28bad60b37e023e1cd701559](http://onlinelibrary.wiley.com/store/10.1093/clipsy.9.1.21/asset/clipsy.9.1.21.pdf?v=1&t=hs1yf18&s=6e5864549034ca1b28bad60b37e023e1cd701559)

- Minami, T., Wampold, B. E., Serlin, R. C., Hamilton, E. G., Brown, G. S. J., & Kircher, J. C. (2008). Benchmarking the effectiveness of psychotherapy treatment for adult depression in a managed care environment: A preliminary study. *Journal of Consulting and Clinical Psychology, 76*(1), 116–124. doi:10.1037/0022-006X.76.1.116
- Nathan, P. E., Stuart, S. P., & Dolan, S. L. (2000). Research on psychotherapy efficacy and effectiveness: Between Scylla and Charybdis? *Psychological bulletin, 126*(6), 964–981.
- Norcross, J. C. (2001). *Empirically supported therapy relationships*.
- Okiishi, J. C., Lambert, M. J., Eggett, D., Nielsen, L., Dayton, D. D., & Vermeersch, D. A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients' psychotherapy outcome. *Journal of Clinical Psychology, 62*(9), 1157–1172.
- Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66*(1), 126.
- Phelps, R., Eisman, E. J., & Kohout, J. (1998). Psychological practice and managed care: Results of the CAPP practitioner survey. *Professional Psychology: Research and Practice, 29*(1), 31.
- Pirkis, J., Burgess, P., Coombs, T., Clarke, A., Jones-Ellis, D., & Dickson, R. (2005). Routine measurement of outcomes in Australia's public sector mental health services. *Australia and New Zealand health policy, 2*(1), 8. doi:10.1186/1743-8462-2-8
- Reed, G. M., & Eisman, E. J. (2006). Uses and Misuses of Evidence: Managed Care, Treatment Guidelines, and Outcomes Measurement in Professional Practice. In C. D. Goodheart, A. E. Kazdin, & R. J. Sternberg (Eds.), *Evidence-Based Psychotherapy: Where Practice and Research Meet* (1st ed.). Washington, DC: American Psychological Association.

- Shadish, W. R., & Baldwin, S. A. (2005). Effects of behavioral marital therapy: A meta-analysis of randomized controlled trials. *Journal of Consulting and Clinical Psychology, 73*(1), 6–14. doi:10.1037/0022-006X.73.1.6
- Shimokawa, K. (2010, August). *A Patient-Focused Psychotherapy Quality Assurance System: Meta-Analytic and Multilevel Analytic Review*. (Unpublished doctoral dissertation). Brigham Young University, Provo, Utah.
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology, 78*(3), 298–311. doi:10.1037/a0019247
- Slade, K., Lambert, M. J., Harmon, S. C., Smart, D. W., & Bailey, R. (2008). Improving psychotherapy outcome: The use of immediate electronic feedback and revised clinical support tools. *Clinical Psychology & Psychotherapy, 15*(5), 287–303. doi:10.1002/cpp.594
- Slife, B. D., & Woolery, M. (2006). Are psychology's main methods biased against the worldview of many religious people. *Journal of Psychology and Theology, 34*(3), 217–231.
- Sue, S., Zane, N., Nagayama Hall, G. C., & Berger, L. K. (2009). The case for cultural competency in psychotherapeutic interventions. *Annual review of psychology, 60*, 525–548. doi:10.1146/annurev.psych.60.110707.163651
- Tanenbaum, S. (2003). Evidence-based practice in mental health: Practical weaknesses meet political strengths. *Journal of Evaluation in Clinical Practice, 9*(2), 287–301.
- Tanenbaum, S. J. (1999). Evidence and expertise: The challenge of the outcomes movement to medical professionalism. *Academic medicine : Journal of the Association of American Medical Colleges, 74*(7), 757–763.

- Umphress, V. J., Lambert, M. J., Smart, D. W., Barlow, S. H., & Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment, 15*(1), 40–55.
- US Congress. (1974). *Health Maintenance Organization Act of 1973*.
- US Congress. (1982). *Tax Equity and Fiscal Responsibility Act of 1982*.
- Wampold, B. E., & Brown, G. S. J. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*(5), 914–923. doi:10.1037/0022-006X.73.5.914
- Wennberg, J. E. (1984). Dealing with medical practice variations: A proposal for action. *Health affairs (Project Hope), 3*(2), 6–32.
- Westen, D., Muderrisoglu, S., Fowler, C., Shedler, J., & Koren, D. (1997). Affect regulation and affective experience: Individual differences, group differences, and measurement using a Q-sort procedure. *Journal of Consulting and Clinical Psychology, 65*(3), 429–439.
- Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological bulletin, 130*(4), 631.
- Whipple, J. L., Lambert, M. J., Vermeersch, D. A., Smart, D. W., Nielsen, S. L., & Hawkins, E. J. (2003). Improving the effects of psychotherapy: The use of early identification of treatment and problem-solving strategies in routine practice. *Journal of Counseling Psychology, 50*(1), 59.
- Yalom, I. D., & Lieberman, M. A. (1971). A study of encounter group casualties. *Archives of general psychiatry, 25*(1), 16.

Zane, N., Sue, S., Chang, J., Huang, L., Huang, J., Lowe, S., et al. (2005). Beyond ethnic match: Effects of client–therapist cognitive match in problem perception, coping orientation, and therapy goals on treatment outcomes. *Journal of Community Psychology*, 33(5), 569–585.