2011-03-14

Let Me See My Feedback: A Phenomenological Exploration of the Feedback-Receiving Process at a University Counseling Center

David Doty Dayton
Brigham Young University - Provo

www.manaraa.com

Let Me See My Feedback: A Phenomenological Exploration

of the Feedback-Receiving Process at a

University Counseling Center

David Dayton

John Okiishi
Sally Barlow
Lane Fischer
Aaron Jackson
Steve Smith

Department of Counseling Psychology & Special Education

Brigham Young University

April 2011

ABSTRACT

Let Me See My Feedback: A Phenomenological Exploration

of the Feedback-Receiving Process at a

University Counseling Center

David Dayton

Department of Counseling Psychology & Special Education

Doctor of Philosophy

This study is a phenomenological investigation of psychotherapists' experience receiving quantitative and evaluative feedback on job performance. Participants were licensed psychologists working at a university counseling center. They were given feedback reports that compared their clients' psychotherapy outcomes with the outcomes of their colleagues' clients. Psychotherapy outcomes were measured using the Outcome-Questionnaire 45 (OQ-45), a self-report outcome instrument designed for tracking client progress through repeated measurement. Feedback reports included data about where psychotherapists' outcomes ranked (in quartiles) in comparison to the rest of the counseling center. Interviews were conducted with participants to gain a deeper understanding of their experience receiving quantitative and evaluative feedback.

Interviews were conducted, transcribed, and analyzed in accordance with the phenomenological method as explained by Wertz (2005) and the descriptive psychological phenomenological method as explained by Giorgi (2003). Content of interviews was grouped into four emergent themes: Ego Responses, Interpretation, Credibility, and Application.

Responses indicated that participants felt both threatened and reassured by their feedback. Those who reported feeling reassured were more inclined to see this as a validation of their approach to psychotherapy while those who felt threatened were more inclined to see the feedback as an assessment of identity. Many indicated that they struggled to understand terminology on the feedback reports as well as the statistical methodology used to analyze the data. Those who struggled to interpret the feedback reports were more likely to distrust or dismiss the results. While very few participants were dismissive of the notion that the feedback reports were valid measures of therapist efficacy, many were ambivalent about this question. Participants did not indicate making concrete behavioral changes as a result of receiving the feedback, although a few reported that the feedback induced introspection about "what is good psychotherapy," as well as dialogues with colleagues.

Keywords: feedback, performance assessment, qualitative, psychotherapy outcomes

Acknowledgements

Thanks to Stacy for providing the feedback that I know she did not have the energy to give. Also thanks for not rolling her eyes when she saw my "writing process."

Thanks to my committee members, Dr. Sally Barlow, Dr. Lane Fischer, Dr. Aaron Jackson, Dr. John Okiishi, and Dr. Steve Smith, for their patience and support. Thanks to John Okiishi and Steve Smith for going to bat for me.

Thank to the clinicians at the BYU Counseling and Career Center for allowing me to interview them and for taking this project seriously.

Thanks to Mom, Dad, and Grandma Doty for their unwavering support.

Table of Contents

List of Figures

viii

List of Appendices

ix

Introduction

Doctors are men who prescribe medicine of which they
know little to cure diseases of which they know less in
human beings of which they know nothing.
                                                        -Voltaire

In the past century, the quality measurement and management movement has

infiltrated industries around the globe (Dow, Samson, & Ford, 1999). It is known by

many names: performance assessment, total quality management (TQM), quality

improvement (QI), continuous quality improvement (CQI), and quality assurance, to

name a few. While it started in the industrial and manufacturing sector, the movement

has more recently made its way into the health care industry. The advent of Medicare

and Medicaid in the 1960s pressured the medical industry into initiating a more

detailed examination and evaluation of medical practices (Millenson, 1997). Evidence-

based medicine emerged and the subsequent improvement of services delivered as well

as dropping mortality rates testified to the effectiveness of the quality measurement and

management movement in the medical field (Millenson).

Providers of mental health services are now facing similar expectations to document

and deliver more effective treatment (Brown, Dreis, & Nace, 1999). Not only is there

pressure from health maintenance organizations (HMOs) to streamline the delivery of

services (as they view psychotherapy as a "commodity"), but, given the abundance and

accessibility of information, consumers are more perspicacious, litigious, and less

tolerant of inadequate services than ever. Globally, there is an increasing distrust and decreasing reliance on professionalism amongst consumers (Radin, 2006). Professionals in general, and mental health practitioners in particular, have been overly protective of their autonomy and unresponsive to consumer needs (Bickman, 1999). Further, psychology has an ethical and moral responsibility to self-monitor and self-regulate (Bickman, 1997; Moras, 2005). Such accountability not only serves the best interests of the public, but also reestablishes fading professional credibility. As Burlingame, Lambert, Reisinger, Neff, and Mosier (1995) noted over 10 years ago, "the age of treatment accountability in mental health care has arrived."

The performance assessment movement is not without problems, especially in the mental health care industry. While most clinicians agree there is some value in performance assessment, there is little consensus on to how engage in the evaluation and feedback process in a meaningful way. Many complain that performance assessment based on quantitative data foments insecurity and stifles creativity and intuition and there is legitimate concern about the unintended consequences of the uncritical application of feedback (Radin, 2006). Further, the feedback processes that have historically succeeded in the industrial and manufacturing complex are not necessarily transferable to the mental health care profession as they fail to address the complexity of the psychotherapeutic process. There are an abundance of theories that propose different ways of engaging in the feedback process but little research

(quantitative or qualitative) to support any of these theories (Grol, Bosch, Hulscher, Eccles & Wensing, 2007). The feedback and performance evaluation process is tenuous and should be approached cautiously.

The purpose of quality management is to evaluate current performance with the intention of reinforcing good behavior by professionals where outcomes are strong and influencing a change in behavior where there is a deficit in positive outcomes. While therapists today receive extensive training and supervision, they usually operate in the absence of quantitative information about clients' responses to treatment (Sapyta, Riemer, & Bickman, 2005). Yet, in spite of the lack of quantitative feedback, most clinicians believe they produce above-average outcomes. In a 2003 survey conducted by Dew & Riemer, 143 clinicians were asked to rate their performance on a scale from A+ to F. 66% of the clinicians rated themselves A- or better. Sapyta et al. argue that this inflation is likely due to the general human tendency to overvalue one's own work and that the lack of quantitative feedback severely limits the continued development of more effective care.

In 2005, a quantitative performance-assessment feedback system was proposed to be delivered to the clinicians at the Brigham Young University Counseling and Career Center (CCC). This proposed system was designed by researchers at the CCC (many of whom also practice psychotherapy at the CCC). Researchers at the CCC have enjoyed a highly collaborative relationship with clinicians at the Center over the years and have

access to an enormous database of Outcome Questionnaire 45 (OQ-45) scores. The OQ-45, a self-report outcome instrument designed for tracking client progress through repeated measurement, is administered prior to every individual psychotherapy session with every client at the CCC.

The proposed feedback system was based solely on extant OQ-45 data and was designed with rigorous statistical methodology ranking therapists on multiple domains. As part of the proposed feedback system, clinicians would receive an individual feedback report that graphically demonstrates how the outcomes of their clients stack up against the outcomes of their colleagues' clients. These feedback reports would be based on OQ-45 scores gathered from all of their clients. The proposed forms would also provide clinicians with information denoting in which quartile their outcomes fell in contrast to their colleagues.

As a research assistant at the CCC involved in the design of these feedback reports, I attended research and faculty meetings and saw how vociferously some clinicians opposed the proposed feedback system. Many claimed that OQ-45 outcome data was inadequate to assess clinical performance as it failed to account for the nuance and complexity of their clients. Even as I worked on the design of the forms, I remember thinking that it would be a blow to my self-efficacy as a clinician if I were to receive a feedback report that said my performance was "below average." While the debates about using the OQ- 45 as the basis for the feedback forms went on, I found myself

swayed back and forth by the different views of the argument (summed up as: "The OQ doesn't say everything!" vs. "The OQ doesn't say nothing!"). As clinicians were reassured that the feedback reports would be confidential, the proposal gained momentum and was eventually approved.

My ambivalence about the power of the OQ-45 was overshadowed by my interest and emerging preoccupation with the nature of discourse and factionalism within organizations. I was particularly struck by the seeming eagerness of some of the more vocal opponents of the OQ-45 to receive their OQ-45 feedback forms once the proposal was approved. I was also acutely aware that while the feedback forms were designed using rigorous statistical methodology, they were designed with little regard for the feedback *process*. As I began wondering what a more palatable feedback system would look like, I started exploring literature on feedback for psychotherapists—only to find that there wasn't much.

Simultaneously, in a course on research paradigms I gained exposure to qualitative research methodologies such as participatory action research, discourse analysis, narrative research, consensual qualitative research, and phenomenology to name a few. The idea to interview clinicians about their experience during the feedback process emerged.

Given the inevitability of performance assessment in psychology, it made sense to me that the design and implementation of performance assessment would be better off

if it was informed by a deeper understanding of the experience of receiving feedback. I figured that, given my involvement in creating the feedback forms at the CCC, I was in an ideal position to conduct such an investigation. I submitted a proposal to the clinicians at the CCC to conduct a qualitative investigation via interview of psychotherapists' experience receiving quantitative feedback on job performance. They approved the proposal and the study began.

<center>**Review of Literature**</center>

The purpose of the literature review is to familiarize the reader with literature on feedback processes in psychology as well as to create a context for the reader to interpret the data to be presented.

History of Performance Assessment

The quality measurement and management movement is a phenomenon of the 20th century. It can be loosely described as the process whereby individuals and systems are evaluated and receive feedback based on the collection and analysis of data with an emphasis on improving the quality of outcomes and performance (Russ-Eft & Preskill, 2001). The history of evaluation in the United States finds its roots in several factors: public dissatisfaction with education, the military-industrial need for efficiency, and the capitalist values of cutting costs and increasing profit (Russ-Eft & Preskill).

The first major indication of quality management in the mental health care industry was over 50 years ago when Hans Eysenck asserted that psychotherapy is innocuous relative to clients' presenting concerns (Eysenck, 1952). His rationale was that all evidence to the contrary was anecdotal and based on case studies. Outcome data collection and research was subsequently born in an effort to investigate the effects of psychotherapy. Initially, researchers were concerned with the question, "Is psychotherapy effective?" Studies were designed and conducted. The results

repeatedly confirmed that, in fact, distressed folks who go to psychotherapy are better off than distressed folks who do not go to psychotherapy.

By the 1970s, the effectiveness of psychotherapy was well established (Wampold, 2001). The question was no longer about *whether* psychotherapy was effective; the empirical evidence overwhelmingly debunked Eysenck's assertion about the impotence of psychotherapy by demonstrating that, for the most part, psychotherapy *is* helpful. The question had shifted to "Which type of psychotherapy is most effective?" This was followed by a barrage of studies that ostensibly "validated" certain theoretical orientations (that the researchers conducting the studies coincidentally happened to feel an allegiance towards).

As the performance assessment movement started to infiltrate the mental health industry, HMOs became more involved in subsidizing psychotherapy for their clients. In an effort to cut costs and increase profit (as is often the wont of large corporations in a capitalist economy), HMOs started to impose limits on the number of sessions they would subsidize for their constituents. Armed with research that supported the notion of "empirically validated treatments" (EVTs), they became reluctant to subsidize treatment that wasn't "proven" to be effective. This meant that unless a practitioner adhered to a treatment modality that had been empirically validated, she would not be included on the list of approved mental health care providers. Obviously, this was problematic for practitioners who disagreed with philosophical assumptions of certain

EVTs and the research that validated them—not to mention the clients who preferred different approaches to psychotherapy.

While EVTs continue to maintain a strong influence where HMOs and governmental funds are involved, outcome research has emerged in the past 15 years suggesting that theoretical orientation has almost no bearing on effective outcomes; that, in fact, effective outcomes in psychotherapy are largely the result of atheoretical components found in *all* therapeutic encounters referred to as *common factors* (Asay & Lambert, 1999, Maione & Chenail, 1999; Wampold, 2001). Common factors are the elements of psychotherapy that are incidental to all types of therapy: the therapeutic relationship, the therapist qualities, and the characteristics of the client, to name a few.

The common factors research has called into question the validity and methodology of the EVT research by suggesting that theoretical orientation has little to no bearing on outcomes—that, in fact, the positive outcomes found in EVT research were more a function of how the researchers defined "success" than psychotherapeutic techniques (Lambert & Hill, 1994). Even in the medical field, it is well-established that there are very few interventions that are beneficial to all patients (Rothwell, 2002). There has been increasing criticism directed at the generalizability of EVT studies conducted in laboratory settings because the results of these studies are often not replicable in real life settings (Bickman & Noser, 1999; Riemer, Rosof-Williams & Bickman, 2005). Research has also suggested that the best method of managing costs for HMOs is to

ensure that individuals receive effective care in the first place—although without any indication of how effective care is defined (Brown & Jones, 2005).

Given the inevitability of performance assessment in psychology, it appears to be incumbent on the industry to internally develop quality improvement processes that are customized to the complexities of psychotherapy—perhaps with an approach to quality improvement that is itself unique and reflective of the values of psychology. If this fails to happen internally, policy makers and regulators driven by market forces will do it (Miller, Duncan, Sorrell & Brown, 2005; Radin, 2006).

The ongoing debate about what constitutes effective therapy and how to measure it is a reflection of the arrival of the quality measurement and management movement in the mental health care industry. The aim of this section has been to highlight the inevitability of performance assessment practices in the mental health care industry. The following section will review the literature on popular performance assessment theories.

Performance Assessment Theory

The paucity of research on performance assessment in mental health services compels us to look for models of performance assessment in other industries such as medicine and, more specifically, psychiatry. This study presupposes that more ideological overlap exists between psychiatry and psychotherapy than say, automobile manufacturing and psychotherapy—and accordingly anticipates similar obstacles and

solutions in the performance assessment process. It is important, however, to be mindful of the ideological differences between psychiatry and psychology to preserve the distinction between the two—a distinction that is becoming increasingly blurry (Wampold, 2001).

Wampold explains that as HMOs become more involved in the subsidization of psychological services, the pressure to conceptualize psychotherapy as a medical treatment distorts the fundamental nature of psychotherapy. "Cast in more urgent tones," he explains, "the medicalization of psychotherapy might well destroy talk therapy as a beneficial treatment of psychological and social problems" (p. 2). We might even think of the field of psychology as going through an identity crisis that is akin to the individuation process. By exploring the boundaries that differentiate psychology from its "family" and simultaneously being mindful of similarities, psychology could work towards establishing its unique and distinct identity, especially as it stands in relation to medicine and psychiatry.

The summary of the literature on performance assessment theory is divided into three parts. The first part will focus on the research on the major obstacles to quality improvement in medicine and psychiatry. The second part will focus on the research on major obstacles to quality improvement in general. The final part will focus on recommendations for quality improvement in psychology.

Performance assessment research in medicine and psychiatry. The major obstacles

to performance assessment in the medical and psychiatric literature have been condensed into two themes: lack of leadership and clinician resistance.

   Lack of leadership. Hermann and Rollins (2003) have written that "foremost among obstacles to meeting information needs on quality of care is the absence of leadership" (p. 215). Such leadership, they suggest, will help professionals reach consensus on how to approach performance assessment, which measures to use, and how to interpret the meanings of those measures. The process of achieving consensus should involve all stakeholders in the mental health industry—administrators, clinicians, and clients (Hermann et al., 2004; McGrath & Tempier, 2003; Shortell, Bennett & Byck, 1998). Effective leaders can negotiate the conflicting needs and demands of the various stakeholders, create unity, and make difficult decisions when compromise is not feasible.

   In their research on the organizational dynamics, Kofman and Senge (1993) have observed that the perceived need for leadership reinforces a sense of powerlessness and, ironically, interferes with the emergence of leadership from within the organization. Typically we imagine effective leaders to be gifted, unique, and dynamic personalities that come along every once in a while. Such people are personable, intelligent, humble, and void of any major weaknesses; basically what amounts to what they call "myths of a heroes" (p. 17). An alternative to this individualistic and heroic notion of leadership so common in our culture is the concept of "servant leadership" (p.

18). Servant leadership emerges when people lead because they choose to serve other people and a higher purpose. An example of servant leadership frequently occurs on the battlefield where "the only leader whom soldiers will reliably follow when their lives are on the line is the leader who is both competent and who soldiers believe is committed to their well-being" (p. 18).

One of the specific challenges for leaders in a health care setting will be creating a culture where outcome and performance assessment can thrive. This will involve convincing nonclinician administrators that for some patients, improvement is at times imperceptible and that the quality of care cannot always be meaningfully measured (McGrath & Tempier, 2003). Likewise, leaders will need to figure out how to motivate, involve, and when necessary, tolerate uncooperative clinicians who too easily dismiss outcome data and performance assessment. All stakeholders would do well to remember that clinical competence is an extremely complex construct that requires "multiple, mixed, and higher order methods of assessment" (Howley, 2004, p. 300).

Clinician resistance. In 1835, French physician Pierre Louis published a book called Research on the Effects of Bloodletting. As the first medical doctor to use numerical data to support his assertions, Louis demonstrated that bloodletting, a very popular medical procedure at the time, was actually harmful. The medical community was outraged. It took the remainder of the 19th century for the practice of bloodletting to lose favor.

The implementation of new data and methods continues to be a slow and tedious process in medicine. There has always been a tendency for doctors to resist new information, especially when it contradicts strongly held theories. There are numerous examples of resistance to theoretically iatrogenic treatments shown to be effective (Rothwell, 2002). It is anticipated that if clinician resistance is strong in medicine and psychiatry, where more sophisticated and reliable outcome data is still considered "imperfect," such resistance will be even stronger in the mental health field where outcome data is more tenuous (Howley, 2004, p. 287).

While there is a certain irony to the resistance of feedback in the mental health profession, not all clinician resistance should be interpreted as insecurity and defensiveness or the desire for monopolistic control of the profession. What Bickman (2002) calls the "recalcitrance" of clinicians might also be understood as an understandable reaction to assessment processes that threaten and demean professional autonomy (p. 197). There is legitimate concern that performance assessment stifles creativity and intuition by funneling the decision-making process into a limited spectrum of predetermined options. The decision-making trees in medicine are built around the statistical values of probability and predictability—values that not all mental health care professionals champion. Additionally, the push towards efficiency is fundamentally at odds with the belief common in *all* professions that highly specialized and nuanced abilities cannot be standardized, rationalized, or commodified

(Freidson 2001). Reliance on statistically based benchmarks could "institutionalize" performance at the median (Mor, 2005, p. 341).

Ultimately however, performance assessment will fail if clinicians continue to resist the process. In order to create an environment where clinicians resist less, organizations need to adopt performance assessment practices that give equal importance to what researchers can learn from clinicians and vice versa (Kleingeld, Van Tuijl & Algera, 2004; Lavis, Robertson, Woodside, McLeod & Abelson, 2003; Rothwell, 2002).

Performance assessment research in general. In contrast to the previous section, this section summarizes the major obstacles to performance assessment in the quality improvement literature in general. The obstacles have been condensed into three general themes: confusion about the feedback, lack of infrastructure, and conflicting goals.

Confusion about the feedback. Even when there is a willingness on the part of professionals to receive feedback, the data is often ambiguous, confusing, or difficult to interpret. One meta-analytic study suggested that over one-third of feedback interventions had a negative effect on performance (Kluger & DeNisi, 1996). Those that don't understand feedback data are more likely to distrust it and become further entrenched in the belief that outcomes that reflect quality cannot be generated. (Huffman, Martin, Botcheva, Williams & Dyer-Friedman, 2004).

Even when feedback data *is* clear and understandable, it is still less helpful than

feedback delivered in the form of ideas or "actionable messages" rather than data (Bickman & Noser, 1999, Lavis et al., 2003). Lavis et al. notes that decision makers rarely use a regression coefficient when faced with making a decision. The challenge here for the mental health industry is to establish a link between outcome data and the specific processes, interventions, and/or factors that facilitated those outcomes, and deliver the feedback in way that is understandable and applicable (Lambert & Hill, 1994; Mor, 2005). One major challenge of creating and delivering feedback that is understandable and applicable is the lack of infrastructure.

Lack of infrastructure. Infrastructure refers to the features of organizational systems that are conducive to the development and implementation of feedback technology. Information systems technology develops exponentially and makes the analysis and delivery of exhaustive data in real time an actual possibility. For example, a firm in Japan has developed a "smart" toilet that, at your command, measures your weight, fat ratio, temperature, protein, and glucose and sends the data to your doctor (Boyle, 2001). The development and implementation of such technology progresses at a considerably slower rate in the mental health industries (Hermann & Rollins, 2003; McGrath & Tempier, 2003; Mor, 2005).

Part of the problem is that since outcome measures are deemed imperfect, there is increased demand for a broader array of data to compensate for that imperfection. One of David Boyle's "counting paradoxes" in his book *The Tyranny of Counting: Why*

Numbers Can't Make Us Happy (2001) is "The more we count, the less we understand."

There is a lack of agreement about which core measures should be used (Hermann et

al., 2004)). Even if there was agreement, there is no technology in place to synthesize

the data into useful messages for clinicians although medicine and psychiatry are

aggressively working towards this. Additionally, the ubiquitous demand for process

data is usually accompanied by an acknowledgement of how difficult it is to gather

(Lambert & Hill, 1994). Finally, there is concern about overwhelming the client with too

many questionnaires and measures (Hatfield & Ogles, 2004).

   An infrastructure that rapidly processes and delivers exhaustive data requires

enormous investments of time, money, and cooperation by all stakeholders. As

insurmountable and overwhelming as the process of developing a new infrastructure

seems, there is consensus in the literature that it is an important condition for

meaningful performance assessment to occur (Bickman, 1999; Bickman & Noser, 1999;

Bickman & Peterson, 1990; Donabedian, 1986; Lavis et al.; 2003; McGrath, 2004;

McGrath & Tempier; Mor, 2005).

   Conflicting goals. In her book, *Challenging the Performance Movement: Accountability,*

Complexity, and Democratic Values, Beryl Radin (2006) explains that most organizations

involve "a range of actors with different agendas and conflicting values operating

within a fragmented decision process." Organizational goals and the goals of

individuals within organizations can be specific, ambiguous, explicit, and/or implicit.

Radin explains that the multiplicity of conflicting goals compounds uncertainty and compels us to simultaneously attend to many factors when we make decisions. Performance assessment and measurement is a formalized process whereby we examine how well we have achieved our goals. The traditional approach to performance assessment is dominated by what Radin calls lateral thinking: a style of thinking that conceptualizes complex systems as fragmented entities made up of separate parts that operate independently from each other in a linear fashion.

Lateral thinking has led to success in conquering the physical world (e.g. irrigation, refrigeration, flying) through adhering to principles of the scientific method (Kofman & Senge, 1993). As the issues organizations face are increasingly systemic in nature, lateral thinking approaches to goal setting and problem solving that have historically been successful are too simplistic to deal with the interdependent complexity and nonlinear realities of dynamic and evolving organizations (Radin). With the intention of decreasing uncertainty, a lateral thinking approach to problem solving typically leads to overplanning, which in turn creates more uncertainty. A vicious cycle ensues where the uncertainty caused by overplanning creates more uncertainty. Radin believes this is evident in the tendency of many organizations to solve auditing problems with more auditing. Medical research already suggests "comprehensive plans with detailed targets for parts of the systems seldom improve patients' care in complex systems" (Grol, Bosch, Hulscher, Eccles & Wensing, 2007, p. 119).

Radin's recommendation that we conceptualize goals within a systems analysis framework is consistent with complexity theory—a theory that emphasizes viewing system behavior and change in terms of wholes rather than parts (Grol et al., 2007). As Kofman and Senge (1993) describe it, we view systems as "wholes within wholes" (p. 13). Plsek and Greenhalgh (2001) say,

> A complex adaptive system is a collection of individual agents with freedom to act in ways that are not always totally predictable, and whose actions are interconnected so that one agent's actions change the context for other agents. Examples include the immune system, a colony of termites, the financial market, and just about any collection of humans (for example, a family, a committee, or a primary healthcare team). (p.625)

Acknowledging the unpredictability and paradox inherent in systems increases both the clarity and difficulty of goals: clarity because goals become simpler and contextual and difficulty because we learn that we cannot realize conflicting goals simultaneously. A systems approach to a multiplicity of conflicting goals would involve harnessing the creative energy of an organization's constituents by fostering an environment of cohesiveness and autonomy and resisting the tendency to increase specifications and requirements (Plsek & Wilson, 2001).

Performance assessment recommendations. Recommendations will be condensed to the following sections: continuous quality improvement, epistemological pluralism,

and specific suggestions. These themes may inform the critique, redesign, and evaluation of the feedback system in the CCC.

Continuous quality improvement. The most written about and endorsed approach to performance assessment is continuous quality improvement (CQI), also known as total quality management (Bickman, 1999, Bickman & Noser, 1999; Dow, Samson & Ford, 1999; Grol et al., 2007; Hermann et al., 2004; McGrath & Tempier, 2003; Mor, 2005; Shortell, Bennett & Byck, 1998). CQI is not a program but "a philosophy that requires leadership and commitment" (Bickman & Noser, p. 250). Shortell, Bennett, and Byck (1998) operationally define CQI as "a philosophy of continual improvement of the processes associated with providing a good or service that meets or exceeds customer expectations" (p. 594). The adoption of a CQI philosophy is a major shift in organizational culture that is "probably the most difficult change that needs to occur" in mental health care organizations committed to performance assessment (Bickman & Noser, p. 253). Like going on a diet or going to the gym CQI is a lifestyle change with seemingly insurmountable resistance. Resources are already stretched so thin that people are reluctant to add yet another thing to do that could quickly devolve into a perfunctory and meaningless meanings and planning sessions.

The change in organizational culture is difficult because the infrastructure required for its success "is very demanding of individuals and organizations along multiple dimensions: cognitively, emotionally, physically, and some might say, spiritually"

(Shortell, Bennett & Byck, p. 605, 1998). It requires a shift towards systems thinking by members in the organization. Kofman and Senge (1993) warn that those who are not predisposed to systems thinking or disagree with its assumptions should not be excluded; often their initial non-responsiveness or confusion evolves into enthusiastic participation. Further, they warn, if those who disagree are excluded, the organization could devolve into a cult of devotees.

Difficulties in implementing CQI arise when "organizations look inward to the needs of their professionals, rather than outward to the needs of their customers…or because [clinicians] perceive that CQI is primarily a cost-control mechanism" (Shortell, Bennett & Byck, p. 608, 1998). In actuality, competition and financial incentives are antithetical to CQI philosophy, which should be viewed as a tool for positive change and not as a measurement of clinical performance (McGrath, 2004). Theories about organizational culture suggest that quality improvement ensues when organizations emphasize collaboration, customer focus, and continued learning (Ferlie & Shortell, 2001).

In his book *The Fifth Discipline: The Art and Practice of the Learning Organization*, Peter Senge (1990) says that service organizations need to become learning organizations. Garvin (1993) defines a learning organization as "an organization skilled at creating, acquiring and transferring knowledge, and at modifying its behavior to reflect new knowledge and insights" (p. 80). The culture of learning organizations is ideal for CQI

as it is characterized by "an experimental mind-set, curiosity about trying new things, a climate of openness, acceptance of debate and conflict, and an ongoing commitment to education, growth, and development at all levels" (Grol et al., 2007). Shortell et al. (1998) say that in medical organizations, CQI is more likely to be effective when physicians are "meaningfully involved in the governance of the organization" (p. 608).

CQI should not be viewed as a panacea to the challenges of performance assessment, as the notion of a panacea contradicts the assumptions of CQI (Bickman & Noser, 1999). However, the combination of scientific and humanistic values inherent in CQI leads to more holistic approaches to assessment that are actually quite compatible with professional values (source). This is particularly true in psychology where the approaches and philosophies are diverse.

Epistemological pluralism. A common concern in performance assessment is that outcome data does not consistently provide an accurate picture of what is happening with clientele (McGrath & Tempier, 2003). Critics are quick to highlight the limitations of data as well as the human tendency to either forget those limitations or reify the data—an artifact of our cultural obsession with calculating things that cannot be counted (Radin, 2006). This is particularly evident in the fields of psychology and education where progress, growth, and learning are vague concepts that are difficult to operationalize and attempts to do so are met with unabating criticism (Radin). Advocates of outcome data and those who remain skeptical about its possibilities are

polarized and entrenched in their respective positions.

Psychology has been dominated by positivistic and postpositivistic research paradigms (Ponterotto, 2005). Gary Walls (2006) details how the American Psychological Association's (APA) promotion of EVTs reflects the dominance of a "technocratic, positivistic epistemology" that leaves less room for psychoanalysis in the world of psychology (p. 654). The conflict looks something like this: those who embrace a positivistic (or postpositivistic) epistemology use positivistic arguments to justify their stance. The seemingly obvious and incontrovertible arguments are only obvious and incontrovertible in a context where positivism is championed as the "best" ways of knowing; a context their detractors aren't necessarily in. Maya Goldenberg explains that "the appeal to the authority of evidence that characterizes evidence-based practices does not *increase* objectivity but rather *obscures* the subjective elements that inescapably enter all forms of human inquiry" (pp. 2630-2631). These detractors (who might accuse positivists of suffering from "physics envy") may seem like broken records with their indefatigable skepticism and monotonous critiques of the limitations of positivistic and postpositivist methodologies (especially as they pertain to performance assessment). Walls' concern is that "the APA's commitment to the medical model and logical positivism undercuts a commitment to pluralism" (p. 660). Morrow (2007) expressed concern that as qualitative methodologies develop and gain credibility, "adherents to positivist and postpositivist paradigms will retrench, attempting to re-establish the

hegemony of their methods" (p. 225).

Epistemological pluralism is proposed as a philosophy that seeks to establish a foundation of mutual agreement where meaningful dialogue and collaboration can occur. Epistemological pluralism advocates an egalitarian approach to how we gather and interpret information. In the western world, empiricism and rationalism are almost considered to be synonymous with knowledge and have historically taken precedence over other ways of knowing such as intuition, tradition, and inspiration (Turkle & Papert, 1990). This is understandable since empiricism and rationalism have helped us achieve a sense of control and safety in the physical world by creating conditions with more predictability and less uncertainty (Kofman & Senge, 1993).

In psychology, where most research is data driven, there is disagreement about whether the ontological assumptions about the universe embedded in positivist and postpositivist methodologies are compatible with the nature of the human psyche and whether they can adequately address the complexity and nonlinearity of intra- and interpersonal dynamics, not to mention the question of meaning (Walls, 2006). The notion that there is potential for meaning in learning to tolerate uncertainty and coexist with ambiguity is also at odds with the drive to create certainty, predictability, and parsimony that is common to positivist approaches. Innovation and insight that come from paradox, intuition, creativity, and other less popular ways of knowing are more likely to be devalued and dismissed in a culture where ideas and arguments presented

without numbers sound wooly and not completely serious (Radin, p. 29).

It is important to emphasize that criticism of positivistic methodology is not commensurate with dismissal of positivistic methodology. In fact, the dismissal of data gathered through such means would be contrary to the notion of epistemological pluralism, not to mention unethical. While the term "epistemological pluralism" does not appear in the literature on performance assessment, it captures the essence of the ubiquitous suggestion that organizations employ a broad repertoire of measures and feedback processes as well as the nearly unanimous agreement that no single measure or process can provide comprehensive or exhaustive feedback. The actual measures and processes used will depend on the population being served, the infrastructure of the organization, and the outcome dimensions prioritized by the various stake holders" (Radin, 2006). The challenge for organizations will be in finding how to situate the data in the performance assessment process.

Epistemological pluralism is appropriate for mental health care organizations as it is reflective of psychology's commitment to multiculturalism (Yanchar & Slife, 1997). They both promote "equality and tolerance and a conceptual diversity that fosters human dignity, worth, and freedom" (p. 658). The success of epistemological pluralism as a philosophy relies on the flexibility and willingness to listen by all parties involved. The likelihood that some folks will continue to rigidly adhere to their already-held beliefs will vary from organization to organization and depend largely on the

organizational climate and the emergence of opinion leaders.

Many critics of quantitative methods still recognize their importance in performance assessment (Radin, 2006). In fact, it would be unethical to dismiss data that suggested a clinician was consistently facilitating negative outcomes, especially since clinical incompetence remains largely undetectable without quantitative analysis (Eddy, 1998). Quantitative measures do not need to be perfect in order to generate meaningful information (Howley, 2004). Wampold (2001) recommends that clinicians who consistently facilitate negative outcomes should receive additional training and supervision. The process of mining the meaningfulness of outcome data should begin with the question, "How imperfect is it?" (Howley, p. 287).

Another implication of epistemological pluralism is that clinicians could seek external feedback that is non-quantitative through activities such as peer supervision, co-therapy, individual psychotherapy, or other creative activities. The assumption that competence is obtained with licensure and maintained with experience and continuing education units is unsupported and unchallenged (Bickman, 1999). The infrastructure for some non-quantitative feedback processes is already in place in most clinics in the form of treatment teams.

Other suggestions. Remaining suggestions for improving performance assessment have been broken into three sections: customized feedback, the need for process research, and specific characteristics of the feedback message.

Customized feedback. Researchers have suggested that one-size-fits-all feedback interventions are less effective and have highlighted the need for feedback systems that are customized to the needs and goals of clinicians, organizations, and the population they serve (Bickman, 1999; Bickman, Riemer, Breda & Kelley, 2006; Davis et al., 1995; Grimshaw et al., 2001; Lavis et al., 2003; Oxman, A.D., Thompson, M.A., Davis, D.A., & Haynes, R.B., 1995; Radin, 2001). Leonard Bickman, director of the Center for Program Evaluation and Improvement at Peabody College, has written extensively about performance assessment in psychology and psychiatry. Bickman and colleagues (2006) have developed a system of continuous quality improvement for mental health service providers called Customized Feedback Intervention Training (CFIT).

The aim of CFIT is "to affect the culture of the organization and create a true learning environment" (p. 87). CFIT is based on research on individual and organizational change and focuses on the constructs of goal setting, cognitive dissonance, leadership, and organizational culture to name a few. The four major components of CFIT are (a) assessing the needs of an organization, (b) implementing a comprehensive treatment progress measurement system, (c) a feedback system, and (d) training.

The approach to performance assessment in CFIT was designed specifically for mental health organizations. It has the flexibility and customizability to adapt to the unique needs and resources of each organization. The components of CFIT are

consistent with the principles of both CQI and learning organizations.

The need for process data. In almost all literature on performance assessment reviewed for this study, regardless of whether it was specific to medicine, education, psychology, or organizations in general, there was a call for process data to supplement outcome data (Bickman, 1999; Bickman & Noser, 1999; Bickman & Peterson, 1990; Donabedian, 1986; Lambert & Hill, 1994; Lavis et al., 2003; McGrath & Tempier, 2003; Mor, 2005). While outcome data can help gain a better sense of positive and negative outcomes, it does not help know anything about the interventions and processes that facilitated those outcomes (Bickman, 2002). Process research in psychology seeks to identify the nature and quality of the therapeutic encounter between the client and therapist, often referred to as the "therapeutic alliance" (Lambert & Hill).

Not only should process data be gathered and reported, it should also be linked to outcome (Bickman & Noser, 1999). Horvath and Luborsky (1993) have commented on the need to identify specific therapist actions that facilitate alliance development or repair. Bickman and Noser go so far as to say that even if process feedback about therapeutic alliance is available, it is not helpful if it does not include recommendations on how to improve it. However, research on ruptures in the therapeutic alliance has suggested that repairing ruptured alliance is at the heart of therapeutic change (Safran & Muran, 2000). While a feedback message to a clinician that a rupture in the therapeutic alliance has occurred does not recommend specific actions, it can initiate

interventions that lead to positive interventions. Some research "indicates that clients frequently drop out of treatment before discussing problems in the [therapeutic] relationship, a therapist would be well advised to use the opportunity afforded by the [alliance feedback] to open discussions and work to restore the alliance" (Miller, Duncan, Sorrell & Brown, 2005, p. 201).

The push for process data is not new to psychology. Given the confidential nature of psychotherapy, the documentation of detailed processes is hardly pragmatic (Mor, 2005). Further, process researchers struggle to agree on which process variables to measure (Lambert & Hill, 1994). A breakthrough in process research would likely require a large and ambitious collaboration between practitioners and researchers. Non-empirical process information that explores the intangible and ineffable elements of therapeutic encounters can be obtained by inviting peers to observe encounters that, in the clinician's judgment, are especially problematic.

Specific characteristics of the feedback message. Another common recommendation in the literature is that feedback messages should be understandable and timely (Bickman, 1999; Bickman & Noser, 1999; Lavis et al., 2003). Feedback that is delayed for too long loses its impact. Lavis et al. detail the ideal conditions for knowledge transfer to occur. "Actionable messages" are preferred to research reports (p. 223). The caveat is that not all findings can be translated into actionable messages although they warn that such an excuse can be "overused" (p. 223). Lavis et al. also reports that "passive processes are

ineffective and that interactive engagement may be most effective, regardless of

audience" (p. 226). However, it is up to the stakeholders to decide which

measurements to use and fine-tune the strategies for implementing it.

Method

> *Is it possible, I wonder, to study a bird so closely, to observe and catalogue its peculiarities in such minute detail, that it becomes invisible? Is it possible that while fastidiously calibrating the span of its wings or the length of its tarsus, we somehow lose sight of its poetry? That in our pedestrian descriptions of a marbled or vermiculated plumage, we forfeit a glimpse of living canvases, cascades of carefully toned browns and golds that would shame Kandinsky, misty explosions of color to rival Monet? I believe that we do. I believe that in approaching our subject with the sensibilities of statisticians and dissectionists, we distance ourselves increasingly from the marvelous and spell-binding planet of imagination whose gravity drew us to our studies in the first place. This is not to say that we should cease to establish facts and to verify our information, but merely to suggest that unless those facts can be imbued with the flash of poetic insight then they remain dull gems; semi-precious stones scarcely worth collecting.* (Moore & Gibbons, p. 236)

The purpose of this study is to gain a more nuanced and in-depth understanding of the feedback-receiving experience—an understanding that might inform the design and implementation of subsequent feedback processes. More specifically, *this study is looking at the feedback-receiving experience of psychotherapists who have received evaluative and quantitative feedback on their job performance.* The study is qualitative in nature. In their oft-quoted paper on evolving guidelines for conducting qualitative research, Elliot, Fischer, and Rennie (1999) stated that "in qualitative research, the researcher attempts to develop understandings of the phenomena under study, based as much as possible on the perspective of those being studied" (p. 216). As this study intends to engage in an in-depth exploration of lived human experience via interview rather than behavior

observed under experimental conditions, a research methodology that is geared for uncovering deeper meaning and insight seems well-suited.

The phenomenological approach as described by Wertz (2005) and the descriptive phenomenological psychological approach as explained by Giorgi and Giorgi (2003) seem most congruent with this aim, as they are methods "shaped according to the intrinsic demands 'of the things themselves'—the psychological lives of human beings" (Wertz, p. 170). Wertz states that "in defining research problems and goals, the researcher reviews established knowledge and critically identifies its limits—some gap between knowledge and reality that requires qualitative knowledge, that is, an understanding of what occurs" (p. 170). The paucity of qualitative research on the feedback process for psychotherapists *is* the gap this study is attempting to fill.

This study intends to provide an in-depth description of the phenomenon in question—the feedback-receiving experience of practicing psychotherapists who have received evaluative and quantitative feedback on their job performance. The thoughts, feelings, opinions, attitudes, ideas, and beliefs expressed by participants during the interviews provided insight into how they experienced the feedback process and how they made sense of this experience. What follows is a description of the parameters of this study and an explanation of how the phenomenological method worked towards gaining this understanding.

Participants

    Selection. The participants in this study came from the pool of clinicians at the BYU CCC who received the evaluative and quantitative feedback reports (discussed later in more detail) and consented to be interviewed about their experience receiving these reports. This included 22 full-time clinicians currently employed at the CCC. Their age range varied from the 30s to the 60s. Some of the clinicians had less than 5 years of post-licensure experience while others were in the final 5 years of their career. Of the 22 clinicians, there were 14 men and 8 women, all of European American descent. All participants were licensed practitioners in the state of Utah and have received PhDs in either Counseling or Clinical Psychology, with one holding a PhD in Marriage and Family Therapy and one holding a PsyD. I had some form of personal contact with all clinicians prior to the interviews, whether as a supervisee, a student, group co-leader, or as a member of a clinical team.

    Although BYU is a private religious institution, the clinicians in the CCC enjoy the freedom to practice psychotherapy in the absence of any institutional mandates or pressure. These clinicians practice a wide range of therapies, including gestalt therapy, Rational Emotive Behavior Therapy, Dialectical Behavior Therapy, Acceptance and Commitment Therapy, client-centered therapy, psychodynamic approaches, integrative approaches, person-centered therapy, existential therapy, and so on.

The provision of individual, group, and couples counseling is available exclusively to students and partners of students (couples therapy only) enrolled in the university. The "fee" for therapy is covered by tuition costs and other subsidies for the university and is consequently a service with no co-pay fee or entanglement with insurance agencies. There are no limits on the number of sessions a client may attend. When this study was conducted, clients would undergo an intake interview (conducted by licensed psychotherapists and advanced trainees) and would then be referred to an in-house clinician to begin therapy. The referral decision was based on availability of clinicians as well as the severity of presenting concerns: more severe cases were screened for neophyte clinicians. There has been some discussion of switching to a random assignment protocol where the clinician conducting the intake interview continues to work with the client.

The data on the feedback forms that was delivered to clinicians was quantitative and evaluative. It was based on an aggregate of OQ-45 scores gathered from all clients seen in individual psychotherapy by each clinician at the CCC over the past several years. The presumption on the feedback reports was that there exists a relationship between OQ-45 scores and therapist efficacy (job performance). In other words, if there was a trend of deterioration of clients' psychological well-being (according to OQ-45 scores), some of this deterioration might be attributed to the quality of psychotherapy. Conversely, when there was an improvement of

psychological well-being according to OQ-45 scores, some of this improvement might be attributed to the quality of psychotherapy. The degree to which a psychotherapist influenced psychotherapy outcomes was referred to as therapist effects or therapist efficacy. The evaluative component of the feedback forms was that recipients were informed in which quartile their therapist effects ranking compares to the other clinicians in the Center (See Figure 1).

Therapist Info:			Selected Data:	You	Center
ID#:		X	Avg. # of Sessions	11.9	9.68
Date:		2/1/2006	HLM Slope	-1.20	-0.87
# of clients:		84	Slope x Avg. # of Sessions	-14.33	-8.05
			Pre-post Change	-14.19	-9.37
			Avg. Intake OQ	70.56	66.56
			HLM Intercept	69.00	64.25

Summary:

This report is based on clients seen at the Center over the preceding years who were seen for three or more sessions and by a therapist who saw a minimum of 30 clients (you saw 84). Using these criteria, outcomes across 71 therapists were studied. Your composite ranking was in the fourth quartile. On average your clients were seen for 11.9 sessions and their OQ -45 scores dropped 14.19 OQ-45 points during that period.

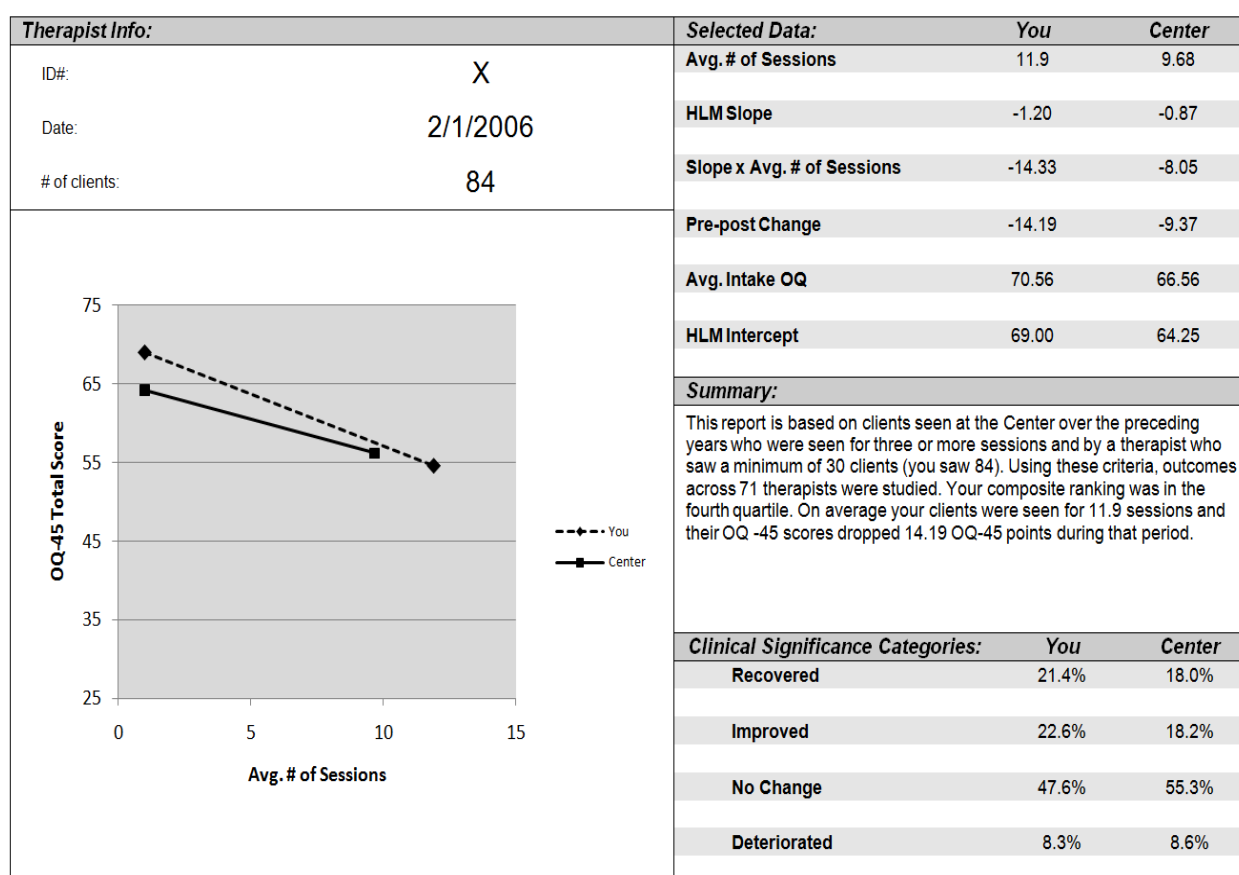| Clinical Significance Categories:	You	Center
Recovered	21.4%	18.0%
Improved	22.6%	18.2%
No Change	47.6%	55.3%
Deteriorated	8.3%	8.6%

Figure 1. Sample feedback form.

One condition of the delivery of feedback reports in the CCC was that clinicians who had not seen at least 30 clients would not receive feedback forms, as they lacked

enough OQ-45 scores to generate statistically significant feedback. The preparation and delivery of the forms was done following a strict blind procedure that preserved the anonymity of the clinicians. The delivery of these feedback forms to clinicians at the CCC was scheduled to occur independently of this study.

Ethical considerations. All clinicians signed an IRB-approved consent form prior to the interview (see Appendix A). The consent form gave permission to the primary investigator to conduct a study on the effects of the feedback system via interview. Descriptions of and quotes from all interviews included in this study are anonymous. Any potentially identifying information has been either altered or excluded from the study. As fewer women were interviewed, male pronouns have exclusively been used in all discussions of the interviews to preserve anonymity. One potential concern for clinicians is that while their responses are treated anonymously, they are not treated confidentially. It is possible that the publication of some statements made in interviews may contain value statements or disclosures that upset or agitate other colleagues and people in positions of power. Conversely, it is also possible that the publication of content from interviews gives a "voice" to folks who feel marginalized, unheard, or powerless to effect change in the current workplace milieu.

Data Collection Process

This section will detail how data was collected, including the process of "bracketing" and the method for generating interview questions.

Bracketing. Phenomenological research is rooted in research methodology originally developed by Edmund Husserl in the early 20th Century (Wertz, 2005). Husserl described processes researchers use to adopt an unbiased approach to their encounter with the research subject. Wertz describes a frame of mind to be adopted while conducting interviews for this study. He suggests that "phenomenological research requires an attitude of wonder that is highly empathic" (p. 172). Getting into this frame of mind involves becoming aware of one's own biases and assumptions, and then temporarily "bracketing" them; that is, setting them aside, in order to enter into what Husserl called the life-world (*lebenswelt*) of the participants. Wertz continues to explain that "this attitude is free of value judgments from an external frame of reference and instead focuses on the meaning of the situation purely as it is given in the participant's experience" (p.172). This is an attitude of "extreme care" that is critical in all stages of phenomenological investigation and analysis. As the primary investigator, I conducted all interviews.

Generating interview questions. The interviews were loosely structured. In studies of an exploratory nature, less structured interview formats are indicated (Suzuki, Ahluwalia, Arora & Mattis, (2007). One benefit of loosely structured interviews is that they allow for context-dependent flexibility in the order, form, or content of the questions. Some interview time was spent educating clinicians on how to interpret the data on feedback forms, as most clinicians were aware of my role in

designing the forms. All clinicians were asked a few of the same questions, such as

"Did you have any difficulty interpreting the data on the feedback forms?" and, "What

suggestions do you have for improving the feedback forms?"

The interviews consisted of open-ended questions such as "What was it like when

you first looked at the feedback form?" followed by prompts from the interviewer to

invite deeper exploration of emergent phenomena. When enthusiasm around a

particular idea emerged, questions inviting further exploration of that idea were asked.

All interviews were conducted one-on-one and face-to-face in the offices of the

interviewees, at their convenience.

Transcribing the interviews. Interviews were transcribed by an accredited

transcription service unaffiliated with BYU. The transcription service was instructed to

substitute an empty line in place of any names spoken in the interviews. Of the 22

clinicians who consented to be interviewed, 18 were interviewed (11 men, 7 women).

These interviews were conducted during a three week period within a month of the

delivery of the feedback reports. Repeated efforts were made to schedule interviews

with the four who were not interviewed. These four interviews were never conducted

because of scheduling conflicts and lack of response to my solicitations to schedule an

interview.

The durations of the 18 interviews ranged from five and a half minutes to 40

minutes. This does not include instances when, during the interview, the participants

asked that I stop the audio recorder so they could speak "off the record"—often for several minutes at a time. Numerous informal conversations with participants about the feedback process occurred spontaneously in hallways of the CCC, during a walk to the parking lot after work, or while we waited for a meeting to begin. No content from these undocumented encounters is included in the study.

   Addressing interviewer transparency.   As I have worked closely with the research team in the CCC, I anticipated the possibility that some clinicians would hold back or censor some of their thoughts and feelings, particularly those who were initially opposed to the proposal to generate feedback reports. My concern was that to whatever extent such participants experienced factionalism within the CCC, they might adopt a more defensive posture during the interviews, as they would possibly view me as a member of the "pro-OQ" camp. In anticipation of this possibility, I chose to adopt a transparent posture where I spoke openly about my thoughts and feelings on the OQ-45. When asked, I stated that while I don't believe the OQ-45 is a panacea to the problems of outcome measurement, I'm confident this feedback research is a small step in helping us find ways to exploit

the information it *is* giving us.

Data Analysis

   Figure 2 is a graphic illustration that provides an overview of the various steps in the process of analyzing and interpreting the interviews.
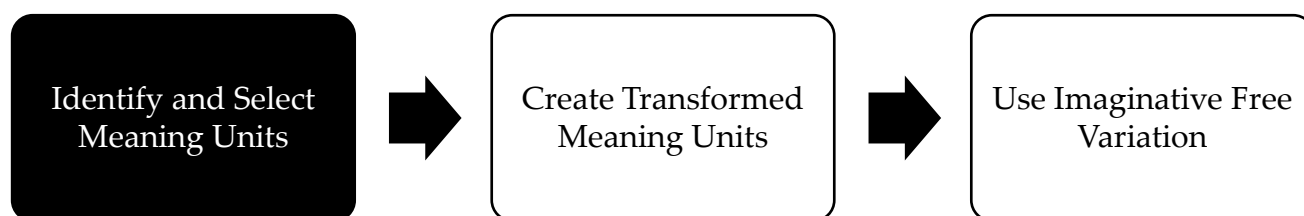
Identify and Select Meaning Units	→	Create Transformed Meaning Units	→	Use Imaginative Free Variation

Figure 2. Visual illustration of the analytical and interpretive process. The first step involves identifying and selecting meaning units.

Identifying meaning units. After interviews are transcribed into written form, Wertz suggests reading the interviews "without the research focus in mind in order to grasp the participant's expression and meaning in the broadest context" (Wertz, 2005, -p.172). This might also be thought of as "reading for the whole" to get a holistic sense of both content and context. In subsequent readings of transcripts, the researcher "differentiates parts of the description, identifying 'meaning units' that organize data for later analysis of parts" (p. 172). Redundant, unrelated, tangential, and irrelevant data were excluded—unless the themes of redundancy, unrelatedness, etc., were meaningful in the context of the research question. I conducted these readings over the course of several months. Additionally, I listened to the recordings of the interview to gain a familiarity with contextual elements of the interviews that can sometimes be lost in transcription, such as tone of voice, rate of speech, and syntax.

After reading and listening to the interviews, I parsed the content of the interviews into *meaning units.* This was done with the intention of organizing the data so that it

might be analyzed more conveniently. This operation of identifying meaning units is referred to as "situated description" or "individual phenomenal description." Giorgi states that there are no "objective" meaning units and that emergent meaning units are "correlated with the perspective of the researcher" (p.252).

As Giorgi (2002) suggests, the concept of validity is neither attainable nor sought after in phenomenological research. However, for this study, Dr. John Okiishi (the dissertation chair) was regularly consulted about the delineation of meaning units. Dr. Okiishi's role was that of a consultant. Although he was unable to listen to the audio recordings (as that would breach confidentiality), he and I periodically met to discuss potential meaning units in the transcripts and process my experience doing research. Subsequent steps in the interpretation and analysis of data also involved regular consultation with Dr. Okiishi.

Creating transformed meaning units. The next step involved creating *transformed meaning units* from the extant meaning units. Wertz explains this step:

> The researcher reflects on the relevance of each part of the described situation and of the psychological process involved, that is, what they freshly reveal for our knowledge about the phenomenon of interest. The phenomenological researcher does not remain content to grasp the obvious or explicit meanings but reads between the lines and deeply interrogates in order to gain access to implicit dimensions of the experience-situation complex. (p. 172)

This step (See Figure 3), according to Giorgi & Giorgi, "[expresses] the psychological meaning of the participant's everyday language more directly" (2003, p. 252). That is, the "meaning units that were originally in the language of the participant will be expressed with heightened psychological sensitivity with respect to the phenomenon under study" (p. 253). This is similar to the process in psychotherapy where, after listening empathically, the psychotherapist paraphrases and restates both the stated and unstated content in more direct and parsimonious language.

| Identify and Select Meaning Units | → | Create Transformed Meaning Units | → | Use Imaginative Free Variation |
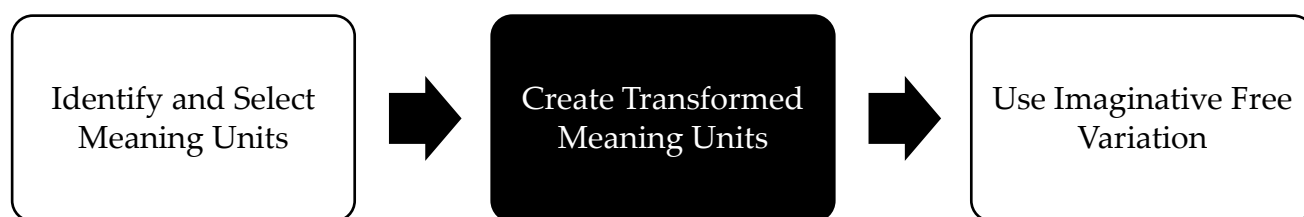
Figure 3. Visual illustration of the analytical and interpretive process. The second step focuses on transforming the meaning units into more psychologically sensitive and workable *transformed meaning units*.

The process of creating transformed meaning units is characterized by what is known as *thick description*, or in this case, *thick interpretation*. According to Ponterotto and Grieger (2007), thick description is the lynchpin of qualitative writing. They state that thick description "involves understanding and absorbing the context of the situation or behavior and, on that basis, ascribing present and future intentionality to the behavior" (p. 415). Giorgi & Giorgi encourage researchers to avoid psychological

jargon or theory-laden language when describing transformed meaning units (2003).
An exhaustive list of meaning units and transformed meaning units is provided in the
appendix.

   Using imaginative free variation. The next step (see Figure 4) in this process
involved identifying which transformed meaning units most accurately captured the
essence of the phenomenon under investigation (Giorgi & Giorgi, 2003). Wertz suggests
a distinctively phenomenological method called *imaginative free variation* to identify the
essence of phenomenon under investigation. The purpose of this step is to identify
which of the transformed meaning units are essential to or definitive of the feedback-
receiving experience and discard those transformed meaning units that are not essential
to this experience.

Identify and Select Meaning Units	→	Create Transformed Meaning Units	→	Use Imaginative Free Variation

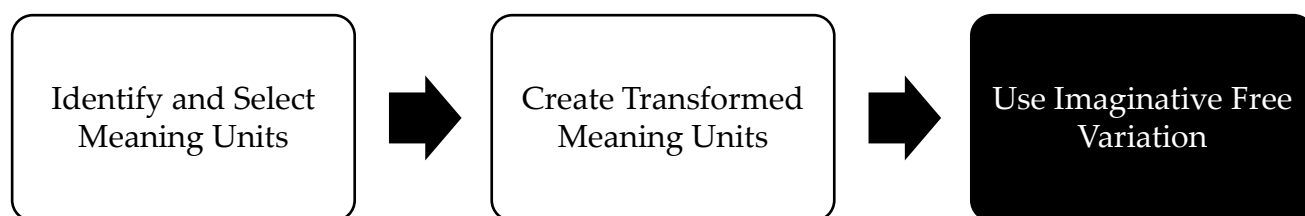*Figure 4*. Visual illustration of the analytical and interpretive process. The third step
uses *imaginative free variation* to identify which of the *transformed meaning units* capture
the essence of the phenomenon under investigation.

   To illustrate imaginative free variation, Giorgi & Giorgi (2003) invite the reader to
imagine a black, octagonal, and ceramic cup. If our task is to identify characteristics of a

cup that are essential to the essence of "cupness," we can imagine cups of various colors, shapes, and materials. We can imagine a round, red, plastic cup or a square, clear, glass cup. In these imagined instances, the shape, color, and material are not essential to the concept of "cupness." When such characteristics are deemed inessential to the "cupness," they "collapse." That is, they are not considered to be definitive of "cupness." However, we can agree that nonporous material and a concave shape are characteristics that *are* essential to concept of "cupness." We cannot imagine a flat or convex object or an object made of porous material maintaining its sense of "cupness."

For this study, it was assumed that not all transformed meaning units identified would be definitive of the essence of the phenomenon under investigation—specifically, the psychotherapists' experience of receiving evaluative and quantitative feedback on their job performance. Imaginative free variation was the process whereby transformed meaning units that are essential to the phenomenon under investigation were identified. It was anticipated that some of the transformed meaning units in this study would collapse, as they were not essential to the phenomenon under investigation. When a transformed meaning unit was under consideration, one question I asked was, "Is the phenomenological experience of receiving evaluative and quantitative feedback on job performance for psychotherapists characterized by [the transformed meaning unit in question]?" If the answer was "no," the transformed meaning unit under question collapses. Accordingly, transformed meaning units that collapsed were

excluded from the findings section of this study. The transformed meaning units that were found to be essential to the feedback-receiving process provide the structure for the findings section of the study.

A common criticism of qualitative research is that it is not generalizable. The phenomenological method, as Wertz has described it, does not presume to make universal findings or be universally generalizable. However, the process of imaginative free variation is crucial to the phenomenological process as it makes possible what Giorgi calls "context-bound generality" (Giorgi 1982). This concept suggests that "truths" that emerge from this study are applicable with confidence only to contexts and situations on the same level of abstraction. Specifically, the findings of this study are assumed to be generalizable to psychotherapists who receive evaluative and quantitative feedback on job performance based on clients' self-report measures.

Presentation of Data

As dozens of meaning units were extracted from the interviews, this study does not present an exhaustive walkthrough of the genesis and transformation of *every* meaning unit that emerges. However, a walkthrough of how some meaning units were identified, transformed, and subjected to imaginative free variation is described in detail to give the reader a sense of how the research process was executed.

It is important to note that identifying which transformed meaning units remain— or do not collapse—after subjected to imaginative free variation was *not* the final step in

this study. Rather, *these remaining transformed meaning units indicate which parts of the original source material are most deserving of examination*. In other words, the remaining transformed meaning units redirect focus to the parts of the interviews that most articulately and effectively capture the essence of the phenomenon under investigation. The foundation of the findings in this study was built upon those remaining transformed meaning units.

Findings

Nature of the Interviews Described

It may be instructive to note some general characteristics of the nature of the interviews before proceeding with the analysis of the content of the interviews. All interviewees were gracious, polite, and accommodating, even when they reported feeling upset or agitated by the feedback reports. In some interviews I sensed an initial defensiveness on the part of the participant, characterized by short and terse answers to my initial questions. Upon further reflection, I believe this presumed defensiveness was more often an artifact of my anxiety and *expectation* of defensiveness than actual defensiveness. In fact, not only were interviewees *not* defensive, they were eager to talk. In retrospect, I believe they realized there was something unique about this study before I did—that, due to the anonymous nature of these interviews, they could express opinions and feelings about their co-workers, the culture of the CCC, and their work experience without fear of being ostracized, challenged, or recriminated.

Transformed Meaning Units Identified

After the interviews were transcribed, I read them repeatedly over the course of several months to get a sense of the meanings. Additionally, I occasionally listened to them to re-familiarize myself with how the intonations of the participants affected the meaning of what they were saying. The purpose of repeated readings and listenings was to get a "sense of the whole" and develop a familiarity with both the content and

the context of each interview. I highlighted segments from each interview that struck me as possible meaning units and processed them with Dr. Okiishi.

In total, 143 meaning units were identified. While this study does not include a rationale for the delineation of every meaning unit, the following sections will detail two examples that illustrate the three steps in the process of how meaning units are identified, transformed, and subjected to imaginative free variation. All 143 meaning units were subjected to this process. The first example includes interview content that does not collapse when subjected to free imaginative variation. The second example is included, for the reader's convenience, to demonstrate what it looks like when content from an interview ultimately collapses. Every meaning unit in this study was subjected to these three steps and an exhaustive list of all meaning units, their corresponding transformed meaning units, and whether or not they collapsed, is included in Appendix B.

Example #1. The following interview excerpt is included to illustrate how two meaning units were identified and selected. In this excerpt, the participant talks about his beliefs on Hierarchical Linear Modeling (HLM), a popular statistical method that was used by CCC researchers to generate data presented on the feedback forms.

> *Interviewer*: Is there anything confusing about the way the data is presented or how to interpret it?

> *Participant*: HLM slope means nothing to me. I don't even know, number one, if I trust HLM. I think when we get too sophisticated of techniques, the data is messy and I have trouble trusting it, let alone, necessarily even knowing what it is. I've never done any HLM and I've never actually had a class on it and I've never even been in a class where it was even considered.

The participant's syntax, characterized by the use of unequivocal language ("HLM slope means nothing") and judgmental language ("the data is messy"), suggests that the statement is rich with emotional content, indicating that it may be denoted as a meaning unit. The point here is not to argue whether or not HLM slope should have meaning to the interviewee or whether or not the data is messy, but to tap into the phenomenological experience of receiving quantitative feedback for the participant. The operating assumption in this research paradigm is that emotionally-heightened content cues the researcher into the phenomenological experience of the participant.

From this excerpt, two meaning units have been delineated. The first one, "HLM means nothing to me," has a subtle but distinct meaning in contrast to the second one, "I don't know, number one, if I trust HLM." Whereas the first statement is about *understanding* HLM, the second one is about *trusting* HLM—two qualitatively different experiences. The remaining dialogue from this excerpt—"I have trouble trusting it, let alone, necessarily even knowing what it is"—serves to qualify this distinction between meaning units.

The process of creating transformed meaning units from these two meaning units involves restating the content with more direct and psychologically sensitive language that captures the between-the-lines meaning of the original statements. Whereas meaning units are presented in the original language of the interviewee, transformed meaning units are stated in the language of the researcher. This is what was referred to earlier as *thick interpretation*.

On the surface, the statement, "HLM slope means nothing to me," sounds like it may be an expression of indifference. However, when taken in the context of the interview, including the statement,"I've never done any HLM and I've never actually had a class on it and I've never even been in a class where it was even considered," it becomes evident that this is a statement is about understanding, not indifference. Additionally, on the audio recording, the interviewee clearly places emphasis on the word "nothing." If it were a statement of indifference, the interviewee would stress the word "means." (Try saying it aloud. "HLM *means* nothing to me" vs. "HLM means *nothing* to me"). Accordingly, the transformed meaning unit that emerges is, "I was confused by the data on the feedback form," as this statement most succinctly captures the between-the-lines meaning of the participant (see Figure 5).

| "HLM means nothing to me." | → | "I was confused by the data on the feedback form." | → | Does it "collapse" when subjected to imaginative free variation? |
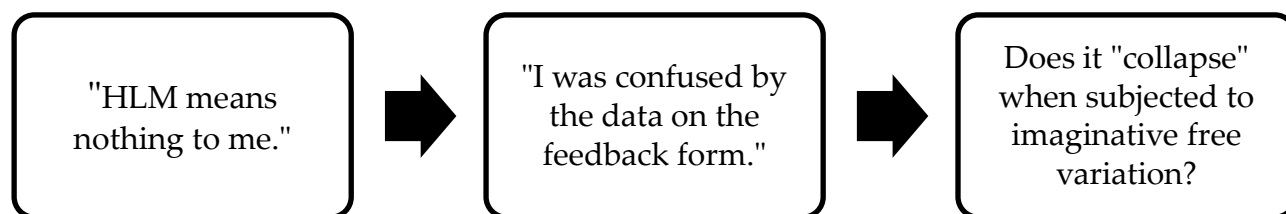
Figure 6: Visual illustration of Descriptive Psychological Phenomenological method in action.

The process of transforming the other meaning unit—"I don't even know, number one, if I trust HLM"—is more straightforward. As this statement is clearly about trusting the credibility of the statistical method used to generate the data on the feedback forms—HLM—it is transformed to, "I do not trust the statistical methods used to generate the data on the feedback form."

The final step involves subjecting these transformed meaning units to imaginative free variation to ascertain whether or not they are essential to the phenomenon under question—the psychotherapists' experience of receiving evaluative and quantitative feedback on job performance. To re-summarize how this process works, bring to mind the example of trying to understand the essence of "cupness" from chapter 3. If the transformed meaning unit (akin to the characteristics of the cup such as size, shape, color, porosity, and concavity) "collapses" when subjected to imaginative free variation, it is *not* considered to be essential to or definitive of the phenomenon under

investigation (the "cupness"). If it does not collapse, it *is* considered to be essential (in the case of the cup, porosity and concavity did not collapse and are therefore essential to the notion of "cupness"). The findings of this study are made up of those transformed meaning units that do not collapse when subjected to imaginative free variation.

In order to use imaginative free variation in this study, I imagine the phenomenon under investigation—psychotherapists' experience receiving evaluative and quantitative feedback on job performance—in hypothetical contexts where certain variables of the phenomenon are altered. These variables include the types of outcome measures used, the number of different measures used, the presentation of the feedback, the methods of analyzing the data employed, the type and nature of the organization (private practice, community mental health center, hospital, etc.), the level of experience of the clinicians, and their degree of professional training. I can be as creative as I want with changing these variables, so long as every hypothetical scenario involves the delivery of evaluative and quantitative feedback on psychotherapists' job performance. This means there are countless imaginable hypothetical scenarios. For example, I can imagine a scenario in which evaluative and quantitative feedback based on three different measures is delivered to clinicians at a hospital on a US military base—*in Chinese characters*. The absurdity of delivering feedback in Chinese characters to English speakers in this particular scenario helps answer an important question: Does

the transformed meaning unit, "I was confused by the data on the feedback forms" capture the essence of the feedback-receiving experience in this hypothetical context? The answer is "yes." If the feedback is delivered in Chinese characters, it is likely that, with the exception of participants who can read Chinese characters, even the most statistically adroit clinicians will be confused by it. In many of these hypothetical scenarios, it is plausible that the individual receiving feedback could struggle to interpret the feedback.

However, it also is possible to imagine hypotheticals in which the participants do not struggle to interpret the feedback—particularly scenarios that rely on simple statistical methodologies or scenarios where all of the participants in the study have an advanced comprehension of statistical methodology. What becomes evident in this process of imaginative free variation is that, regardless of the degree to which the participant struggles to understand the feedback, the *process* of understanding the feedback is invariably a part of the participant's feedback-receiving experience. In other words, the extent to which the participant is able to understand the feedback is characteristic of *every* participant's feedback-receiving experience. Therefore, the statement, "I was confused by the data on the feedback forms" does not collapse, as it captures one aspect of the process of interpreting feedback.

The other transformed meaning unit—"I do not trust the statistical methods used to generate the data on the feedback form"—also does *not* collapse when subjected to

imaginative free variation. In every imagined scenario, it becomes apparent that the participants' experience will invariably be characterized by the *question* of the credibility of feedback. Again, the point here is not to ascertain the actual credibility of the feedback, but to investigate the extent to which questions about the credibility of the feedback are a part of the participant's experience.

Example #2. This second example is included to give the reader a sense of the process of identifying and transforming a meaning unit that collapses during imaginative free variation. While the content of this interview is engaging and the interviewee's comments were characterized by an endearing vulnerability, the content ultimately does not provide insight into the phenomenon under investigation. During this interview, the participant spontaneously opened up about his thoughts and feelings about his experience of the culture of the CCC.

> *Participant:* You know for the first time I'm kind of thinking new thoughts here. That this really does, that your study really has the potential to really raise issues about our culture of trust here. Which, I don't open up to a lot of people here and for me that's interlaced with a lot of [LDS] church issues. I'm not as churchy as a lot of people here. I feel very different and I'm pretty reluctant to really share that at a deep level. And so if we really started to talk about who we are as human beings and how much we trust one another, there'd be kind of a church component in that that would complicate it. For me it would. I don't know

whether other people would say the same. Has this issue come up for other people?

Interviewer: The church issue?

Participant: Uh huh. My worthiness as a church member? My depth of belief. If I started to talk about who I was and really develop trust here, pretty soon it would bring up issues about the church. I have a lot of questions about the doctrine, the practices of the church. And if I were to develop a lot of trust here those issues would come up for me. And the kind of culture we have now makes it really hard for me to share that. Does that make any sense to you?

Interviewer: Because you might be ostracized?

Participant: Yeah. I'd be seen as questioning the church too much. Too much of a doubting Thomas. Maybe that's not right. Maybe I should test that more.

This excerpt is an expression of feeling both isolated from colleagues and caution about being more open and vulnerable around them. Implicit is a desire to interpersonally connect with colleagues on a deeper level. The essence of this entire excerpt is most succinctly captured by the interviewee's statement, "I feel very different and I'm pretty reluctant to really share that at a deep level." This sentence has been delineated as the sole meaning unit from this entire excerpt. While other segments from this excerpt were considered to be possible meaning units, they were ultimately not delineated as such. When taken in the context of the entire interview, it became evident

that these other statements served more to qualify and deepen our understanding of the primary meaning unit than to introduce a separate and distinct meaning unit.

This meaning unit—"I feel very different and I'm pretty reluctant to really share that at a deep level"—is transformed to, "I feel interpersonally isolated from my colleagues and I'm afraid to open up to them." The transformation of this meaning unit went through several iterations. The speaker's original word "different" and the word "alienated" were considered but, given the context of the statement, "interpersonally isolated" seemed to most accurately express the intentionality of the participant. "Different" is ruled out as it is too vague an emotional expression (is "different" really an emotion?). "Alienated" is not chosen as it is too easily associated with being a pariah, which is incongruent with the between-the-lines meaning of what was expressed. "Interpersonally isolated" seems to capture both the sense of loneliness and the craving for more collegial connection implicit in the participant's expression.

When subjected to imaginative free variation, this meaning unit collapses. That is, it is *not* deemed essential to the phenomenon under investigation. While it is conceivable that the sentiment of feeling isolated and reluctant to open up to colleagues could be experienced by a participant in all hypothetical feedback-receiving scenarios, this phenomenological experience occurs independently of receiving the feedback forms. The assumption here is that even if these feedback forms had never been delivered to clinicians at the CCC, this individual would have still experienced

interpersonal isolation from colleagues. This means that the essence of the feedback-receiving experience is *not* reliant on musings about the quality of one's interpersonal relationships with colleagues. It is worth mentioning that in several of the interviews, participants voiced their feelings, positive and negative, about the working culture of the CCC and the varying degrees of trust they felt towards their colleagues. It seemed that in many of these instances, these were expressions of sentiments that pre-existed the delivery of the feedback.

Categories Created

After submitting each of the 143 transformed meaning units to the process of free imaginative variation, 124 of them remained. These remaining 124 units most effectively capture the essence of the phenomenon under investigation—the psychologists' phenomenological experience with receiving evaluative and quantitative feedback on job performance. It becomes apparent that most of these remaining transformed meaning units naturally coalesce around broad themes. For example, transformed meaning units from several interviews contain a variation of the statement, "I want this feedback to be supplemented by additional information on my individual clients." The emergent theme in these statements is an expression of a desire for more detailed feedback.

The purpose of categorizing the remaining transformed meaning units into thematic groups is to organize and simplify the presentation of the findings. Over the

course of several months, repeated start-from-scratch attempts to categorize the transformed meaning units into themes were made. In one attempt, as many as 19 thematic categories were created. Eventually, the transformed meaning units were pared down to 4 broad thematic categories: Ego Responses, Interpretation, Credibility, and Application. A fifth category, Organizational Culture, was also created to serve as a repository for most of the transformed meaning units that collapsed during imaginative free variation, such as the one mentioned earlier, "I feel interpersonally isolated from my colleagues and I'm afraid to open up to them." My hypothesis is that statements about the organizational culture were prevalent because the anonymous nature of the interviews provided a venue for interviewees to safely vent pre-existing frustrations.

Not every remaining transformed meaning unit fits neatly into one of these broad themes. Many transformed meaning units have meanings that overlap into two and occasionally three of the broader themes. Others are only loosely connected to the themes. These units that overlap multiple themes or stray from themes add nuance and complexity to participants' experience and will be discussed in more detail. These four themes, made up of the transformed meaning units that withstood imaginative free variation, provide the structure for the remainder of the chapter.

Ego responses. One interviewee stated, "My initial response was one of, 'Why am I not good enough?' And, 'Better find another profession.' And that sort of thing." Later

in the same interview he said, "When the initial kind of ego-wound sort of started to heal I began thinking about what kind of clients I see and what I could do to improve my curve." This attitudinal shift from wounded to motivated was common in interviews.

Some interviewees described the emotions they experienced when they received the feedback reports or when they anticipated receiving them. These descriptions came in response to questions such as, "What was it like to receive this feedback?" A large range of emotions were described, ranging from feeling hurt by the feedback to feeling reassured by the feedback. Participants described thought processes that followed these initial emotional reactions; how they reconciled feedback that was incongruent with their perceptions of their own therapist efficacy, or how they were cautious about getting too excited about positive feedback results. Many reflected on, and in some instances questioned, their identity as a psychotherapist. One interviewee stated, "I had to rethink, ok now, 'What does this mean about me as a therapist and where I held myself in comparison with the rest of the clinic?'" For the most part, this feedback had a resonant impact on the participants; they took it seriously and in some cases, personally. The label "Ego Responses" has been designated as a term that broadly encompasses this wide range of emotional and identity-questioning responses to receiving evaluative and qualitative feedback on job performance.

Some participants reported having positive feelings after receiving the feedback. One interviewee, who said the feedback "bolstered [his] confidence," discussed the relationship between receiving this feedback and his approach to psychotherapy:

I, in the last couple of years, since the whole feedback thing started, have become more free [*sic*] and less constrained in my therapeutic style. I'm giving myself permission to share more of myself than I ever have. This might have had some confirmation for me to keep doing that, keep being more who I am and be less, kind of technique-oriented, and more, I would say, existential—where I am a human being that sort of encounters another human being. And to be freer with my feelings and my reactions. So this data moves me in that direction a little more.

It is worth emphasizing that this participant talks about how he felt more confident about his humanistic approach to psychotherapy (vis-à-vis feeling more confident about his self-worth). Such responses were common amongst participants who reported experiencing positive feelings after receiving the feedback reports. Another interviewee, while talking about how his more directive approach to psychotherapy is considered to be less popular by colleagues in the Center, stated, "It was kind of reassuring to know that actually my clients are doing fine, that it's not a bad thing to do what I'm doing." He qualified this feeling of reassurance by saying, "I felt like I didn't

feel personally validated exactly. However, I felt like what I do—it's good to know that it helps at a slightly above average rate."

This tendency to interpret the feedback as a reflection of one's approach to psychotherapy was absent in those who reported unpleasant experiences after receiving the feedback; these participants seemed more inclined to internalize the feedback as if it were an evaluation of personal worth. In one of the more vulnerable moments during the interviews, one participant said, "I found my results really upsetting and really struggled to make sense of them." He continued, "My initial response was one of, 'Why am I not good enough?' And, 'Better find a new profession.' That sort of thing." In a more masked expression of hurt, another participant said, "It's the threat, the ego threat. And I feel some of that, but I'm a little bit 'Whatever.' What are they going to do, fire me over it? Fine."

An expression of resentment about the feedback forms seemed evident when one participant stated, "For someone outside the system to say, 'Well you just spent 20 sessions doing nothing because they didn't get better,' that rattles my chains, because I think it did do something to help them." This presumable expression of resentment not only indicates how upset people can get by the notion of receiving evaluative and quantitative feedback, but it also approaches the crux of the ego issue: whether or not feedback is accurate, some people are going to get upset by it. The same individual continues, "[The OQ-45] is one piece of information, and it bothers me when it's some

grand total of everything that happens in therapy, because I think there are other things that happen in therapy besides symptom reduction." The irony here is that the creators of these feedback reports and the creators of the OQ-45 will be the first to state that the OQ-45 does not presume to be the, "grand total of everything that happens in therapy." So where is this hyperbolic expression coming from?

On a more positive note, those who reported feeling hurt or upset after receiving the feedback also demonstrated resilience in their ability to deal with upsetting feelings. The same interviewee who earlier reported feeling upset by the feedback continued, "When the initial kind of ego wound sort of started to heal, I began thinking about what kind of clients I see and what I could do to improve my [slope]." Some described how they dealt with feeling hurt or threatened by discrediting the feedback reports by highlighting flaws in the methodology (i.e. the lack of random assignment) or highlighting the limitations of the OQ-45. Others reported opening up to colleagues to process their results and seek emotional support. One participant described his experience processing his results with a colleague:

> And it was interesting that [another therapist] came to talk to me because he was in the [second] quartile and he was very upset by it. And then on finding out that I'm in the [second] quartile made him feel better.

Even those who received positive feedback expressed a desire to process the feedback with others. One of the participants who reported receiving positive feedback said,

"What I'd like to do is I'd like to find somebody in here who's willing to be pretty darn open with this data. And me and that person sit down and then talk about it, understand it. I'd like to do that with another person or small group." Although the feedback reports were produced and distributed in a highly secretive manner, it appears that many participants were eager to process their feedback with colleagues, perhaps to receive support, reassurance, or recognition.

One question asked during interviews was about participants' thoughts on a proposal to include more specific rankings on subsequent feedback forms (i.e. "Your results ranked 3rd out of 71 therapists") as opposed to the quartile rankings that were provided ("Your results were in the third quartile."). Many participants were unequivocal in their responses to this question. One participant spoke disparagingly of the notion, saying, "Ranking, to me, it reminds me of a beauty contest." His rationale was that ranking, "probably creates more political problems than it gives helpful information." Another stated that comparison with other therapists, "is where I feel less comfortable because it's more threatening, but I also don't know what it means."

The contrasting view—in favor of including specific rankings on the feedback forms—was exemplified by another participant's expression, "I want to know my exact ranking out of the total number of therapists. I want to know." In regards to those who do not want specific rankings, he stated his belief,

I think people who won't admit that are in a way lying to themselves. Like there's some people who avoid going to the dentist 'cause they know they're going to have a cavity, right? But if they really look at it hard, they don't want to avoid going to the dentist, right?"

Rather than being a reflection of the "right" or "wrong" view, these opinions about providing specific rankings give insight into the different ways people receive evaluative and quantitative feedback on job performance. One interpretation of this phenomenon is that those who oppose the inclusion of specific rankings are more inclined to experience a connection between job performance and self-worth whereas others are more inclined to experience it simply as a reflection of the work they've done.

In summary, participant responses indicated that, for the most part, there was an emotional component to the experience of receiving a feedback report. These emotional responses ranged from feeling upset to feeling confident. Those who experienced positive responses such as confidence or reassurance were more inclined to interpret the feedback as a representation of performance. Those who reported experiencing negative feelings such as getting upset were more inclined to interpret the feedback as an evaluation of identity or self-worth. They also demonstrated resilience in their ability to process their unpleasant thoughts and feelings. The relationship between self-esteem and job performance appears to be at the crux of how each individual responds to the feedback.

Interpretation. The variance in phenomenological experiences interpreting the feedback is captured in the contrast of the following quotes from two different interviews: "I struggle with what the data means," and, "I had no problem with [interpreting the data] at all." For the purposes of this study, interpretation refers to the participants' experience making sense of the data on the feedback forms. While this section is not about examining the interpretability of the forms themselves, it *is* about the quality of the participants' experience interpreting the feedback. Sentiments describing participants' experience interpreting the feedback forms were frequently shared in the interviews. Several interviewees commented on their experience interpreting the feedback and their experiences interpreting the feedback cover the gamut from, "I understood the data," to, "I did not understand that data."

Several of the interviewees who commented on their experience interpreting the feedback indicated that they struggled to interpret the feedback. Some struggled to understand the terminology on the forms, such as the categorization of clients into groups labeled "no change," "improved," "recovered," and "deteriorated." Others struggled to understand the statistical methodologies used to generate the data. A sampling of quotes from these interviews includes statements such as, "I don't know really how to interpret [HLM slope]," "I'm not sure what you mean by 'recovered,'" "I've always been a little confused by the HLM slope—the meaning of that," "Whatever

quartile I'm in, what does that exactly mean?," and, "So that information I wasn't really sure about."

Many of these expressions of confusion were about but not limited to difficulty understanding the statistical method used to generate the data—HLM. One interviewee stated, "I've had that explained to me about five times now. I keep forgetting it." It is also noteworthy that in several interviews I was asked to explain how HLM works. Several participants requested that a key be included with the feedback to help interpret the terms on the form as well as an explanation of how HLM works. At least one request was made for those who create the feedback reports to provide more face-to-face instruction on how to interpret the results.

Others expressed confusion about the quartile rankings on the forms. During the interviews it was pointed out by multiple interviewees that, in fact, there was an error with the way quartile rankings were presented on the forms. In statistics, those who are in the first quartile are said to be in the lowest quartile—the 25th percentile—and those who are in the fourth quartile are ranked in the top quartile, or the 75th percentile. The problem was that, on the feedback forms, those who were ranked in the fourth quartile were told they were ranked in the first quartile and vice-versa. It appears that clinicians intuited their way through this error either by themselves or through talking with colleagues.

A variation of the experience of interpreting the feedback forms that stood in contrast to expressions of confusion is exemplified by expressions such as "I had no problem with [interpreting the data] at all." Some participants indicated comfort with an imperfect understanding of the feedback with statements such as, "I don't understand [HLM] but I don't think I need a very sophisticated understanding to understand the data." This is congruent with the idea presented in chapter 2 that feedback does not need to be perfect to be meaningful.

There did not appear to be a relationship between participants' experience interpreting the feedback and reports of feeling either upset or reassured by the feedback. That is, not everybody who understood the feedback *liked* the feedback and not everybody who struggled to interpret the feedback was upset by it. The participant who stated that he, "had no problem [interpreting the data] at all," also talked about how this feedback has made him doubt his efficacy as a therapist. "So this was dissonance for me," he said, "that here was objective data that suggests I'm not as good a counselor." Conversely, the interviewee who stated that he did not, "know how to interpret [HLM slope]" also said, "I don't know if I can be up to articulating the reasons why I'm in the [fourth] quartile." In other words, this participant was not excited about his high ranking because he did not understand how that ranking was determined.

In summary, participants' descriptions of their experience interpreting the feedback reports indicated that many of them struggled to understand the data on the forms.

There was confusion about understanding the method used to generate the data on the forms, Hierarchical Linear Modeling, as well as confusion about some of the terminology on the forms. Not all participants struggled to interpret the data on the feedback reports and there did not appear to be a relationship between one's ability to interpret the form and one's emotional response to the feedback.

Credibility. The more dismissive extreme of participants' thoughts on credibility of the feedback was succinctly captured by one interviewee's statement, "I read [the feedback] and it wasn't helpful because I don't pay attention to the OQ very much." One the other hand, the more common response was captured by the interviewee who said, "Well I will say that of course there's error in the OQ, but I have more faith in the OQ than I have in anything else. It's a flawed instrument but it's better than everything else."

Credibility here refers to the extent to which participants believe the data on the feedback is valid and meaningful. The feedback is considered credible if the participant believes it has construct validity. Construct validity refers to the extent to which a measure correlates with the construct it purports to measure. In this case, the construct in question is therapist efficacy, or job performance. This is not a question of whether or not the feedback report is an accurate measure of therapist efficacy, but the extent to which the participants *believe* this to be the case. In other words, do they buy it? "If I could really see that data a little better," said one participant, "then I think I might buy

into it more." The degree to which participants "bought it" ranged from indifference (zero credibility) to "kind of bought it" (some credibility) to "bought it" (enough credibility).

The sentiment of those who were more indifferent to the feedback reports is succinctly captured in the following interview excerpt. During this interview, which was characterized by politeness and an almost apologetic tone, I sensed that the participant was eager to be done with the interview and move on to more pressing tasks. The following exchange occurred:

Interviewer: It doesn't sound like this [feedback] form was of that much interest to you at all.

Participant: You're right. Yeah. And I'm not sure why, because I know a lot of work has gone into it, and there's been a lot of data collection behind it. I mean, this is a big project, and I appreciate those that have done it. But it hasn't been a—what would you call it—an essential or critical part of the way that I work with the students I see.

Needless to say, this was one of the shorter interviews. While this interviewee was polite and expressed gratitude for the effort put into generating the feedback, it was evident that the feedback lacked credibility to him. Such responses of indifference to the feedback forms were rare, but they are important as they are indicative of what could occur in most organizations that delivered such feedback—that there may be

some folks who will view quantitative and evaluative feedback based on self-report measures as irrelevant and unimportant.

Other participants also conveyed an attitude of indifference to the feedback (often couched in criticisms or dismissals of the OQ-45). However, unlike the interviewee in the previous example, these participants also made statements at other points in their interviews that belied their presumable indifference; statements suggesting that, in fact, there was something credible about the feedback reports. The following exchange illustrates this idea:

Participant: I read it and it wasn't helpful because I don't pay attention to the OQ very much.

Interviewer: What wasn't helpful about it?

Participant: I'm just not sure why it would have been helpful?

Interviewer: Was there anything confusing or difficult to understand about the data?

Participant: I don't think so. I' mean, it's been a while, you know. It's been a month since I read it. I noticed that I had more deteriorators than the clinic average. I guess that was the part that I would have liked to have talked with you about. That probably would have been helpful for me to know, who those people were. So I could begin to think about what was going on with them.

This suggestion to include additional information on "deteriorators" — a term on the

feedback forms designated for clients whose trend in OQ-45 scores indicates a considerable increase of psychological distress—was the most ubiquitous suggestion in the interviews.   Regardless of how interviewees felt about the credibility of the feedback or the OQ-45, many wanted their feedback to be supplemented by additional information on their clients, especially the "deteriorators." The conflicting expressions in this particular excerpt—a dismissal of the OQ-45 coupled with an expression of a desire for more OQ-45-reliant data—are not easily reconciled. How does the participant simultaneously dismiss the OQ-45 and express interest in receiving more OQ-45-reliant data? What makes these conflicting expressions important is that they capture an inner tension that frequently surfaced in the interviews.

The tension in these conflicting expressions is definitive of a common variation of participants' experience in regards to the credibility of the feedback—ambivalence about the construct validity of the reports themselves. Many participants indicated ambivalence about the presumed relationship between the data used to generate the feedback reports (OQ-45 scores) and the notion of therapist efficacy. That is, they struggled to either embrace or dismiss this presumed relationship. One participant described the feedback as "a report card for therapists based on measures we like-slash-question." Notwithstanding the critiques of the limitations of the OQ-45, most participants made comments, suggestions, and requests that implied a belief that the OQ-45 itself has adequate construct validity—that it accurately enough measures

psychological distress. Even the most skeptical participants made statements indicating that they valued the OQ-45. Said one interviewee, "Well I will say that of course there's error in the OQ, but I have more faith in the OQ than I have in anything else. It's a flawed instrument but it's better than everything else." Even one of the more vociferous critics of the OQ-45 stated, referring to OQ-45 data on his form, "This at the bottom, yeah, this makes sense to me, and this is something that I would want to address."

Participants' discomfort with the notion that the feedback reports based on OQ-45 scores are accurate indicators of therapist efficacy is implicit in another one interviewee's statement, "I would like to have, maybe, more discussions among ourselves about the OQ as an outcome measure versus the OQ as a sort of indicator of how you should do treatment." Another interviewee hypothetically asked, "Is it a valid reflection of the work I'm doing versus the kind of clients I'm seeing?" suggesting that these feedback results are largely affected by the lack of random assignment of clients to clinicians. This tension points to an important and seemingly unanswered question that remains active for therapists: Does this feedback accurately measure therapist efficacy? When considering the indication of ambivalence in the interviews and the fact that, for many, the OQ-45 has some credibility, perhaps the more helpful question is, "To what extent is this feedback an accurate measure therapist efficacy?"

It is also instructive to observe that some of those who argued that the OQ-45 is not

a reliable indicator of therapist efficacy also reported feeling the most hurt and threatened by the feedback reports—implying that the feedback has some degree of credibility. This was in contrast to others who also indicated that the OQ-45 is not a reliable indicator of therapist efficacy but did not report feeling hurt or threatened by their results. If the feedback reports were actually dismissible and lacked credibility, how did some of those who perceive them as such become so upset? While some participants expressed concern that this type of feedback could someday threaten their job security (as administrators might use such data to make decisions about hiring and firing), such expressions were hypothetical and lacked the emotional immediacy and intensity of the more emotionally-charged expressions of feeling hurt or threatened. It seems numbers lend credibility and valence to a concept. According to the literature review, one possible explanation is that this tendency to feel threatened by data deemed to lack credibility is "an artifact of our obsessions with counting what cannot be counted." In other words, numbers with an evaluative component are powerful, irrespective of their actual credibility.

Imagine if all psychotherapists in an organization were asked to rank each other to generate a list of the top 25 therapists, based on a composite of these subjective rankings, much like what is done in college football polls. Many would likely be offended by such a suggestion, pointing out the flaws and problematic implications of such a method. And many of them would likely be eager to see the results. The point is

that credibility has both an emotional and an intellectual component and that the emotional component, at least in this instance, trumps the intellectual component. The take home point is that, in addition to refining the actual credibility of a measure, those providing feedback will also likely need to devote some energy towards ensuring that it is *perceived* as credible.

Application. Application refers to the extent to which participants applied or considered applying information acquired in the feedback process to their approach to psychotherapy. As there was an absence of actionable messages or specific suggestions on the feedback forms on how clinicians could apply the feedback, most seemed unsure about what to do with it. This commonly occurring sentiment is captured by one interviewee's statement that, "it was not so helpful either because it was positive or negative. I simply did not know how to make use of some of it."

Even participants who indicated that they believed the feedback was high on credibility appeared to believe it was also low on applicability. "It would be helpful to translate into terms I find clinically useful," said one interviewee. In the following quote, another participant explains that when his job performance is evaluated, he would appreciate more information about the relationship between what he's doing in psychotherapy and the data on the feedback reports:

I want to see my individuals and their notes. And—*like that's the bridge between this [feedback report] and making some kind of meaningful intervention* (emphasis

added). I mean, if I have a really crappy slope, like really bad slope, and I'm 70th out of 70, it's like, 'Okay, I need to change what I'm doing.' Or if I'm at the top, something about what I'm doing is working. But that's kind of what I want to do is find a way to make this more like, helpful, I guess. As it is right now, it's just kind of 'Ehh.'

Consistent with the view that the feedback is "kind of 'Ehh,'" participants did not cite concrete examples of interventions they made in therapy as a result of receiving this feedback. However, some described intentions to make general changes to their approach to psychotherapy. "I could probably terminate a few people a little more quickly," one interviewee stated, in response to data on the forms that compares the clinician's average number of sessions per client to the center average. Another participant reported that he has been more mindful of the therapeutic relationship since receiving this feedback (although there is no data on the feedback reports about the therapeutic alliance). He stated, "Actually, [the feedback] has got me thinking about one aspect. I used to be heavily worried about technique and if I was using the right techniques and interventions, and now I'm focusing more on the relationship aspect." It seems that, at least in this and a few other examples, the feedback forms fomented introspection about approaches to psychotherapy.

When asked whether or not he had done anything different in his practice as a result of the information on the feedback reports, another interviewee explained that

there was not a simple yes or no to the question:

> Now I know in a sense if you were to watch my work you wouldn't see any
> difference. But I have been very active in questioning and wondering and
> thinking and what does it mean and, 'how does it fit with this client?' And,
> 'what are the implications for this or for that?' I've been much more, what's the
> word—curious. I've been much more curious about how my colleagues have
> been doing work, and I have found that extremely useful. I've been kind of
> listening a little more carefully as cases are presented. So yeah, I've done a lot of
> [thinking] as a result of the feedback. So if you were to videotape my sessions
> you wouldn't be able to tell, just determine from one session. From even a long
> time ago to the very last session I did, you wouldn't see much change.

Evident in these quotes is an earnestness about making this data more useful. While
participants were not necessarily applying their feedback in tangible and behaviorally
observable ways, they were actively thinking about *how* to apply it. "When I look at
[my feedback form]," said one participant, "I want to know what went into that? Why?
Where can I improve? Where am I doing well?" Also implicit in such statements is a
belief that the feedback is credible.

That participants were brainstorming about how to use this data is evident in the
numerous suggestions and recommendations for improving the feedback reports made
during the interviews. Two participants talked about wanting a "bridge" between the

feedback and practice. Another participant stated, "Well, I think it would be a really good idea to give us some way to drill down into the data to look at all our unsuccessful clients." The suggestion to include additional information about "deteriorators" (i.e. individual graphs of clients' OQ-45 scores accompanied by case notes) was the most ubiquitous suggestion in the interviews. At least three interviewees expressed a preference to regularly receive such feedback at frequent intervals. Another popular suggestion on how to make the feedback more meaningful included processing feedback results with trusted colleagues. "In my mind," said one participant, "the most helpful thing that this [feedback] can do is open dialogue—dialogues with myself, dialogues with my colleagues about effective therapy. And if it opens the dialogue, then I think it's been a very helpful thing." The participant who said this feedback has had no impact on the way he does therapy made a recommendation later in his interview:

> If [so-and-so] gets the best results, put me behind the screen. Let me watch [him] every week, an hour a week for a whole year. I'd love it. Even now I wish I had more supervision. Put in a tape and say, [so-and-so], you got a half hour. Let's watch my tape.

Several expressions of concern about the misapplication of data were shared. Some participants said they were concerned that organizations might someday use this type of feedback data to make decisions about hiring, firing, and rank advancement. Even

though administrators at the CCC repeatedly attempted to reassure participants that this data would not be used to make administrative decisions prior to the delivery of these feedback forms, this did not appear ameliorate the suspicions of some participants. "I'm not absolutely sure that it wouldn't show up in rank advancement files," said one participant, "'cause that's one that I have very, very strong feelings about. I don't want that showing up in rank advancement files." With this quote being the one exception, all expressions of concern about how this data might be used by people in positions of power were speculative and lacked the emotional immediacy or urgency of other expressions of concern (such as expressions of feeling hurt). Others said they were not worried about how data would be used "against" them and described such concerns as "premature" and "paranoia." These expressions of concern or lack of concern about administrative use of this data appeared to be a continuation of a dialogue that had been going on amongst participants since the proposal for feedback was made.

Two interviewees shared the concern that some clinicians would adjust their approach to psychotherapy to influence their clients to fill out their OQ-45 questionnaires in ways that would make the therapist look more effective on the feedback reports. That is, they were worried that psychotherapists would somehow "coach" their clients to get lower scores on the OQ-45. One of these elaborated on this concern:

Well, I don't know if I want to say my total bias about it. Here's my concern

about just using the OQ. I don't usually gear my counseling to trying to get OQ

scores down, okay? *Which I think some people actually do* (italics added). I mean,

their goal in therapy is to talk with the client about their OQ score and how they

can—why it's high. And I don't. I mean, I do sometimes if it's really high, but

on a general basis, I'm not going to gear my therapy around the OQ…I'm not

sure I ever want to be in the top quartile, frankly, because that would say to me,

I'm just—I'm more concerned about my OQ scores than I am about the client.

The defensive tone of this statement provides insight into how threatening evaluative

feedback can be perceived and the measures that some may consider taking to improve

their feedback. Similarly, another participant stated, "If I really wanted to make myself

look good, I could do a number of things that would be really unethical," again

referring to the possibility of abusing therapeutic power. In each of these cases it

seemed that the interviewees were concerned about others engaging in such behaviors

and they did not appear to be making such considerations themselves.

In sum, participants' experience applying the feedback into practice was nearly

absent. While many seemed enthusiastic and eager to make the feedback applicable,

they struggled to make concrete interventions based on the data they were given. In

some cases the feedback fomented introspection and sparked dialogue with colleagues.

Several suggestions on how to make the feedback more applicable were made, such as

the suggestion to deliver feedback at regular intervals and supplement the feedback with additional data on individual clients. These suggestions confirm much of what has been said in the literature about positive performance assessment practices. Participants also shared concerns about the misapplication of feedback, particularly the way people in positions of power would use such information to make decisions as well as the fear that some clinicians would manipulate their clients for their own gain.

Discussion

Implications

The purpose of this study was to conduct a phenomenological exploration of psychotherapists' experience receiving evaluative and qualitative feedback on their job performance. Participants' descriptions of their experience receiving such feedback were categorized into four groups deemed essential to the feedback-receiving experience: Ego Responses, Interpretation, Credibility, and Application. Their expressions indicated that their experiences receiving feedback were multi-faceted and involved processing the feedback at affective, cognitive, and interpersonal levels, but there was no indication that any significant behavioral changes were made as a direct result of the feedback. In this chapter these four themes will be related to broader issues about providing quantitative and evaluative feedback on job performance to psychotherapists.

An assumption of this study was that a deeper understanding of the phenomenon of the feedback-receiving experience for psychotherapists would inform the design of subsequent feedback processes. This portion of the study discusses implications for the design of similar feedback systems for psychotherapists as well as implications for future research. It is assumed that the ultimate success of a feedback system does not rely solely on the shoulders of those who are designing it; recipients of feedback also shoulder some responsibility for learning how to interpret and apply the feedback.

Ego responses. Whether or not they should be perceived as such, numbers are powerful. Most psychotherapists who received the feedback took it seriously, at least on an emotional level. This may be interpreted as a vestige of what Beryl Radin refers to as "our cultural obsession with calculating things that cannot be counted," or the human tendency to reify numbers and statistics. While participants were strongly affected by the numbers on an emotional level, they struggled to make sense of them on a cognitive level. In order to be meaningfully interpreted on a cognitive level, numbers need to be situated in a plausible context. That is, clinicians need be able to understand what the feedback is saying about performance. If there is too much ambiguity in the feedback process, they will create their own context for interpreting results. This could lead to clinicians overinflating the value of positive performance or underestimating the seriousness of a problem. Even those who appeared to approach this feedback with a more nuanced and sophisticated understanding of how the data on the feedback forms was generated were not impervious to the power of numbers—particularly the rankings.

The literature review suggested that feedback with an evaluative component has more potential to effect change than feedback without an evaluative component. In other words, feedback that provides an assessment of performance such as a ranking is, theoretically, more effective. It appears that, in this study, ranking individuals in quartiles had enough power to rattle the participants but not enough power to effect

change. The dilemma for those providing feedback is to present evaluative data that is neither too threatening on one extreme (specific rankings) nor too impotent on the other extreme (by excluding evaluative data). One possible consideration for feedback designers is to use percentile rankings, which provides more specificity than quartile rankings *and* would likely be experienced as less threatening than ordinal rankings. The use of percentiles could segue into an eventual adoption of more specific rankings.

Additionally, if a clinician's performance is going to be quantifiably evaluated, he must believe that he can make behavioral changes that will improve dissatisfactory results. This means two things: (a) feedback should be delivered in regular intervals so the clinician can monitor improvement and (b), feedback measures need to be sensitive to changes in therapeutic approach. That is, if a therapist receives negative feedback and makes a change to his approach to psychotherapy, that change must be detectable by subsequent feedback reports. Clinicians want to believe they can make adjustments that will improve their results. To illustrate, consider the analogy of a baseball player who believes that if he makes adequate adjustments to the mechanics of his swing, he can improve his batting average. If a baseball player was told that there was nothing he could do to improve his batting average, he would learn to disregard that statistic and lose motivation to learn and change. Likewise, if feedback results cannot be improved upon by the clinician, the feedback is likely to be experienced as discouraging and eventually dismissed.

Feedback should include statements about why it is being delivered and the extent to which it presumes to measure therapist efficacy. If feedback does not intend to be the end-all measurement of therapist efficacy, it should indicate to what extent it *does* intend to measure therapist efficacy. Continuing with the baseball analogy, batting average is not the only meaningful statistic for hitters; slugging percentage, on-base percentage, and the more complicated on-base plus slugging statistic are also important statistics. But batting average *is* very important and players and coaches would be foolish to disregard or dismiss it.

Steps should be taken to educate those receiving feedback on the importance of feedback processes. Therapists need to understand that incompetence and poor performance are largely undetectable in a post-licensure context without the aid of external measures. Clinicians should be reminded that there is an ethical and moral obligation to protect consumers from poor practices. If organizational leaders follow the suggestion in the literature review that "service organizations become learning organizations," the ensuing paradigmatic shift would theoretically help clinicians get on board with more potent feedback practices, as they would be continuously seeking feedback and reevaluating their practice. In such a culture, discussions of problematic batting averages are less likely to be experienced as demeaning to professional autonomy and more likely to be seen as positive educational and developmental opportunities.

Interpretation. Having prepared every feedback report following a rigorous "blind" procedure to ensure anonymity, I came to regard these feedback reports as almost sacred and certainly very private. So I was surprised by the candid and casual manner in which several participants handed me their feedback forms during the interviews and asked me to help them interpret the data—no matter how positive or negative the numbers were. Initially I thought maybe they were comfortable sharing because there was nothing threatening about me or my low position on the hierarchical totem in the CCC. Reflecting on this now, I see that they were simply eager to make sense of their feedback.

The question of interpretation will always be a part of the feedback process. It appears that those who struggle to interpret the feedback are more likely to distrust and/or dismiss it. Those providing feedback should take pains to make sure the feedback is as understandable as possible.   This could involve providing a key with definitions of important terminology as well as explanations of methodologies used to help recipients to interpret the feedback. Face-to-face instruction with recipients may work towards reducing confusion. If all feedback recipients indicate that they are struggling to interpret feedback, those responsible for feedback processes may want to reconsider the methods they employ or take additional measures to facilitate understanding. However, it's important to keep in mind that those providing feedback are not solely responsible for the interpretation of the feedback. In effective feedback

processes, recipients of feedback are actively involved and invested in taking needed steps to be able to understand the feedback.

Credibility. In this study, many participants appeared to be ambivalent about the credibility of the feedback—they vacillated between attempts to dismiss or discredit it and indications that the feedback had some credibility. Interviewees' thoughts on credibility greatly varied. It appears that *actual* credibility and *perceived* credibility are related but separate constructs—and that, due to the emotional valence of numbers, many recipients of feedback are more likely to be preoccupied with how credible they *perceive* the feedback to be. This is congruent with reports in the medical field (discussed in chapter 2) of medical doctors' reluctance to adopt practices demonstrated to be effective, such as the discontinuance of bloodletting or use of aspirin to reduce blood pressure. While those developing feedback systems should work towards constructing truly *bona fide* and credible measures with adequate construct validity, they should also be mindful of the need to help their measures and methods be perceived as such. One consideration is to identify "servant leaders" in an organization (discussed in chapter 2) and involve them in the development of feedback processes. Enthusiastic participation of such leaders in feedback processes may work towards getting a critical mass of people on board with organizational goals.

Developers of feedback systems should be aware that resistance to feedback processes by individuals in an organization should not necessarily be interpreted to

mean that such folks believe proposed feedback lacks credibility. This is more likely an artifact of their personalities, conflicting value systems, or other unrelated organizational dynamics. That they are speaking up in the first place may mean they are interested in feedback processes, even if they are speaking critically. Proponents of performance assessment systems argue that dissenting views be heard, lest the organization turn into a cult of devotees. Those that seem uninterested and uninvolved are more likely to be indifferent to feedback processes.

Application. Running with the baseball analogy, when a baseball player has a low batting average, that does not tell us what is wrong with his approach to hitting the ball, only *that* something isn't working. Watching film to study the mechanics of the swing and consulting with batting coaches works towards identifying *what* isn't working and making improvements. Additionally, digging deeper into extant data may help hone in on the locus of *what* is not working. For example, identifying that a hitter with a .220 total batting average is hitting a significantly lower .100 against left-handed pitchers helps coaches make informed managerial decisions and players identify where to focus their efforts to improve.

The crux of the problem with making feedback applicable into practice is the difficulty identifying the relationship between therapist behaviors and outcomes. Thus, this study ends with yet another call for process research that links practices to outcomes. As "supershrinks" are identified and observed, researchers will come closer

to identifying such relationships. Until then, it will be difficult to include specific actionable messages in feedback reports. Digging deeper into the data by doing things such as looking closer at "deteriorators" may be one way of harnessing the energy that feedback recipients bring to the table, especially those inclined to more introspection and curiosity. Organizationally sanctioned venues for processing feedback (such as team meetings) and informal processing in groups or dyads may be effective ways of tapping into the creativity and resources of constituents of an organization. If feedback processes are thought of as circular, iterative, and long-term processes, people are more likely to be tolerant of imperfections and enthusiastic about learning and development.

Strengths of the Study

One strength of this study is that it takes information that would otherwise be withheld or shared in intimate and private contexts and puts it into a public venue. Due to the anonymous nature of the interviews, participants were freer to give voice to impulses that might otherwise be muted, suppressed, or ignored. The assumption here is that more transparency about what people are actually thinking and feeling increases clarity. It makes real the notion that organizations are made up of individuals with conflicting values and goals that cannot be simultaneously met. Such realizations change conversations. In contrast to data gathered via survey, an interview format gives participants more of a voice by allowing them use their own language. Additionally, interviews are able to detect nuanced meaning that is easily lost in a

survey, not to mention the depth of insight that can come from conversational exploration.

The content of the interviews appears to be congruent with and confirm ideas about feedback processes presented in the literature review. This includes the notion that people are more likely to distrust feedback they don't understand, the importance of the timely delivery of feedback in regular intervals, the tendency of clinicians to resist feedback, the need for process research, and the idea that feedback is more helpful when it includes actionable messages or suggestions.

Interviews gathered site-specific recommendations on how to make the feedback more helpful and applicable, such as the common suggestion to couple the feedback with additional data on individual clients such as the "deteriorators." Also of site-specific interest is the content from the interviews that collapsed when subject to imaginative free variation. There appeared to be considerable energy and venting from participants around the topics of organizational culture and organizational dynamics (see collapsed units in Appendix B). These unanticipated responses, while not specific to the phenomenon under investigation, may give insight into the relationship between organizational culture and feedback processes.

Limitations of the Study

While the qualitative methods in this study don't presume to be high on generalizability, measures are still taken to ensure that data is gathered and interpreted

in a fair, responsible, and reliable manner. Nevertheless, an investigation reliant on the subjectivity of one primary investigator has inherent limitations. These limitations include, the way the interviews were conducted, the questions asked, the relationship between investigator and participant, and the way the content of the interviews was interpreted. Using the same methodologies, a different researcher would have different relationships with the participants and ask different questions.   Even if the transcript material was identical, other researchers would analyze and interpret them differently and might identify different and/or additional meanings.

It was occasionally tempting to interpret a meaning unit out of context. Sometimes I found myself trying to force a meaning unit to fit one of my preconceived ideas rather than allowing the meaning unit to direct the process. The process of "bracketing" my opinions and values had to be done repeatedly throughout the analytic and interpretive process. This was more challenging with interview content that struck me as excessively defensive or preachy. When this occurred, I had to re-read interviews, reacquaint myself with the context, and adopt a more compassionate attitude that would allow me to connect with the felt meaning of what was being expressed.

I believe the skills and confidence I've developed as a counselor since I conducted those interviews would help me ask sharper questions that invited further exploration. In some cases I was too intimidated by participants and reluctant to push for deeper exploration. I'm confident that if I could conduct the interviews again, I would be

quicker to identify and ask about seeming incongruencies of thought and expression, giving participants an opportunity to respond. Perhaps my biggest dissatisfaction with this study is that I was unable to consult with interviewees about the transformed meaning units and give them an opportunity to respond to my interpretations. Those interested in developing feedback processes in an organization might be able to more effectively include participants in the design process by using a research methodology such as participatory action research. With this method, participants would likely feel less threatened by its experimental mindset. This method would also provide a systematic and boundaried way for researchers to involve participants in research design, which could be mutually beneficial by funneling participants' creativity to the researchers and helping participants feel heard.

One last concern I have about this study is that the phenomenon under investigation may be too broad. Subsequent research may be more effective by narrowing its scope by attempting to flesh out topics such Ego Responses or Ambivalence.

Conclusion

During one of the interviews, when asked what was confusing about the feedback, an interviewee jokingly replied, "The only confusing part was, is how against [the feedback proposal] people were at the beginning and how everybody is on board at the end. That was confusing to me, more than the [feedback] itself." Indeed, I have noticed

and tried to make sense of this same observation. I recall considerable resistance from clinicians at the Center to this feedback when it was initially proposed—as well as a seemingly collective eagerness to receive the feedback reports after the proposal was approved. While there is a bit of irony to the notion of psychotherapists feeling threatened by qualitative and evaluative feedback on job performance, this presumable flip-floppiness is indicative of something going on at a deeper level for psychotherapists: we *want* feedback. We really want it. And, at the same time, we are afraid of it. We're afraid of receiving it because, what if the feedback suggests our performance is inadequate or mediocre or incongruent with our beliefs about our own efficacy? It seems that in order for evaluative and quantitative feedback processes to be effective and meaningful for psychotherapists, we need to open ourselves to these possibilities. Regardless of how "good" our performance is, opening ourselves to these possibilities seems instrumental in learning and moving forward.

References

Asay, T.P., & Lambert, M.J. (1999). The empirical case for the common factors in therapy: Quantitative findings. In M.A. Hubble, B.L. Duncan, & S.D. Miller (Eds.), *The Heart & Soul of Change*. Washington, DC: American Psychological Association.

Bickman, L. (1997). Resolving issues raised by the Fort Bragg evaluation. *American Psychologist, 52*, 562-565.

Bickman, L. (1999). Practice makes perfect and other myths about mental health services. *American Psychologist, 54*, 965-978.

Bickman, L. (2002). The death of treatment as usual: An excellent first step on a long road. *Clinical Psychology: Science and Practice, 9*, 195-199.

Bickman, L., & Noser, K. (1999). Meeting the challenges in the delivery of child and adolescent mental health services in the next millennium: The continuous quality improvement approach. *Applied & Preventive Psychology, 8*, 247-255

Bickman, L., & Peterson, K.A. (1990). Using program theory to describe and measure program quality. In L. Bickman (Ed.), *Advances in program theory* (pp.61-72). San Francisco, CA: Jossey-Bass.

Bickman, L., Riemer, M., Breda, C., & Kelley, S.D. (2006). CFIT: A system to provide a continuous quality improvement through organizational responsiveness, measurement, training, and feedback. *Report on Emotional & Behavioral Disorders in Youth*, 6, 86-87, 93-94.

Boyle, David (2001). The *tyranny of numbers: Why counting can't make us happy*. London, UK: Harper Collins.

Brown, G.S., & Jones, E.R. (2005). Implementation of a feedback system in a managed care environment: What are patients teaching us? *Journal of Clinical Psychology, 61*, 187-198.

Brown, J., Dreis, S., & Nace, D.K. (1999). What really makes a difference in psychotherapy outcome? Why does managed care want to know? In M.A. Hubble, B.L. Duncan, & S.D. Miller (Eds.), *The Heart & Soul of Change*. Washington, DC: American Psychological Association.

Burlingame, G.M., Lambert, M.J., Reisinger, C.W., Neff, W.M., & Mosier, J. (1995). Pragmatics of tracking mental health outcomes in a managed care setting. *The Journal of Mental Health Administration, 22*, 226-236.

Davis, D.A., Thompson, M.A., Oxman, A.D., & Haynes, B.R. (1995). Changing physician performance: A systematic review of the effect of continuing medical education strategies. *Journal of the American Medical Association, 274*, 700-705.

Donabedian, A. (1986). Criteria and standards for quality assessment and monitoring. *Quality Review Bulletin, 12*, 99-108.

Dew, S., & Riemer, M. (2003). Why inaccurate self-evaluation of performance justifies feedback interventions. Paper presented at the 16th Annual Research Conference---A System of Care for Children's Mental Health: Expanding the Research Base, Tampa, FL.

Dow, D., Samson, D., & Ford, S. (1999). Exploding the myth: Do all quality management practices contribute to superior quality performance? *Productions and Operations Management, 8(1)*, 1-27.

Eddy, D.M. (1998). Performance measurement: Problems and solutions. *Health Affairs, 17*, 7-25.

Elliott, R., Fischer, C.T., & Rennie, D.L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38, 215-229.

Eysenck, H.S. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology, 16*, 319-324.

Ferlie, E.B., & Shortell, S.M. (2001). Improving the quality of health care in the United Kingdom and the United States: A framework for change. *The Milbank Quarterly, 79*, 281-315.

Freidson, E. (2001). *Professionalism: The Third Logic on the Practice of Knowledge*. Chicago, IL: The University of Chicago Press.

Garvin, D. (1993, July-August). Building learning organizations. *Harvard Business Review*, 78-91

Giorgi, A. (1982). Issues relating to the meaning of psychology as a science. In G. Foistad (Ed.), *Contemporary philosophy: A new survey* (Vol. 2, pp. 317-342). The Hague, the Netherlands: Martinus Nijhoff.

Giorgi, A. (2002). The question of validity in phenomenological research. *Journal of Phenomenological Psychology*, 33, 1-18.

Giorgi, A.P., & Giorgi, B.M. (2003). The descriptive phenomenological psychological method. In P.M. Camic, J.E. Rhodes & L. Yardley (Eds.), *Qualitative research in psychology: Perspectives in methodology and design* (pp.243-273). Washington, DC: American Psychological Association.

Goldenberg, M.J. (2006). On evidence and evidence-based medicine: lessons from the philosophy of science. *Social Science and Medicine, 62*, 2621-2632.

Grimshaw, J.M., Shirran, L., Thomas, R., Mowatt, G., Fraser, C., Bero, L.,…O'Brien, M.A. (2001). Changing provider behavior: An overview of systematic reviews of interventions. *Medical Care, 39 (8, Suppl. 2)*, pp. II-2-II-45.

Grol, R.P.T.M., Bosch, M.C., Hulscher, M.E.J.L., Eccles, M.P., & Wensing, M. (2007). Planning and studying improvement in patient care: The use of theoretical perspectives. *The Milbank Quarterly, 85*, 93-138.

Hatfield, D.R., & Ogles, B.M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice, 35*, 485-4891.

Hermann, R.C., Palmer, H., Leff, S., Shwartz, M., Provost, S., Chan, J.,…Lagodmos, G. (2004). Achieving consensus across diverse stakeholders on quality measures for mental healthcare. *Medical Care, 42(12)*, 1246-1253.

Hermann, R.C. & Rollins, C.K. (2003). Quality measurement in health care: A need for leadership amid a new federalism. *Harvard Review of Psychiatry*, 11, 215-219.

Horvath, A.O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology, 61*, 561-573.

Howley, L.D. (2004). Performance assessment in medical education: Where we've been and where we're going. *Evaluation & The Health Professions*, 27, 285-303.

Huffman, L.C., Martin, J., Botcheva, L., Williams, S.E., & Dyer-Friedman, J. (2004). Practitioners' attitudes toward the use of treatment progress and outcomes data in child mental health services. *Evaluation & The Health Professions, 27(2)*, 2004.

Kleingeld, A., Van Tuijl, H., & Algera, J.A. (2004). Participation in the design of performance management systems: A quasi-experimental field-study. *Journal of Organizational Behavior, 25*, 831-851.

Kluger, A.M., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254-284.

Kofman, F., & Senge, P.M. (1993). Communities of commitment: The heart of learning organizations. *Organizational Dynamics, 22*, 5-23.

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

Lavis, J.N., Robertson, D., Woodside, J.M., McLeod, J.A., & the Knowledge Transfer Study Group (2003). How can research organizations more effectively transfer research knowledge to decision makers? *The Milbank Quarterly*, 81, 221-248.

Maione, P.V., & Chenail, R.J. (1999). Qualitative inquiry in psychotherapy: Research on the common factors. In M.A. Hubble, B.L. Duncan, & S.D. Miller (Eds.), *The Heart & Soul of Change*. Washington, DC: American Psychological Association.

McGrath, B.M. & Tempier, R.P. (2003). Implementing quality management in psychiatry: From theory to practice—Shifting focus from process to outcome. *Canadian Journal of Psychiatry, 48(7)*, 467-474.

Millenson, M.L. (1997). *Demanding medical excellence: Doctors and accountability in the information age.* Chicago, IL: The University of Chicago Press.

Miller, S.D., Duncan, B.L., Sorrell, R., & Brown, G.S. (2005). The partners for change outcome management system. *Journal of Clinical Psychology, 61*, 199-208.

Moore, A., & Gibbons, D. (1986). *Watchmen*. Brooklyn, NY: DC Comics.

Mor, V. (2005). Improving the quality of long-term care with better information. *The Milbank Quarterly, 83*, 333-363.

Moras, K. (2005). Maintaining psychology's scientific and professional credibility and ethical responsibility to self-regulate: A comment on "validations" of mental health assessment instruments. *Journal of Clinical Psychology, 61*, 315-322.

Morrow, S.L. (2007). Qualitative research in counseling psychology: Conceptual foundations. *The Counseling Psychologist, 35*, 209-235.

Oxman, A.D., Thompson, M.A., Davis, D.A., & Haynes, R.B. (1995). No magic bullets: A systematic review of 102 trials of interventions to improve professional practice. *Canadian Medical Association Journal, 153*, 1423-1431.

Plsek, P.E., & Greenhalgh, T. (2001). Complexity science: The challenge of complexity in health care. *British Medical Journal, 323*, 625-628.

Plsek, P.E., & Wilson, T. (2001). Complexity, leadership, and management in healthcare organisations. *British Medical Journal, 323*, 746-749.

Ponterotto, J.G. (2005). Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *Journal of Counseling Psychology, 52*, 126-136.

Ponterotto, J.G., & Grieger, I. (2007). Effectively communicating qualitative research. *The Counseling Psychologist, 35*, 404-430.

Radin, B.A. (2006). *Challenging the performance movement: accountability, complexity, and democratic values*. Georgetown, DC: Georgetown University Press.

Riemer, M., Rosof-Williams, J., & Bickman, L. (2005). Theories related to changing clinician practice. *Child and Adolescent Psychiatric Clinics of North America*, 14, 241-254.

Rothwell, P.M. (2002). Why do clinicians sometimes find it difficult to use the results of systematic reviews in routine clinical practice? *Evaluation & The Health Professions*, 25(2), 200-209.

Russ-Eft, D., & Preskill, H. (2005). In search of the Holy Grail: Research on investment evaluation in human resource development. *Advances in Developing Human Resources, 7*, 71-85.

Safran, J.D., & Muran, J.C. (2000). Resolving therapeutic alliance ruptures: Diversity and integration. *Journal of Clinical Psychology, 56*, 233-243.

Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: theory, research, and practice. *Journal of Clinical Psychology, 61(2)*, 145-153.

Senge, P.M. (1990). Catalyzing systems thinking within organizations. In F. Massarik (Ed.), *Advances in organizational development* (pp. 197-246). West Port, CT: Ablex.

Shortell, S.M., Bennett, C.L., & Byck, G.R. (1998). Assessing the impact of continuous quality improvement on clinical practice: What it will take to accelerate progress. *The Milbank Quarterly, 76*, 593-624.

Suzuki, L.A., Ahluwalia, M.K., Arora, A.K., & Mattis, J.S. (2007). The pond you fish in determines the fish you catch: Exploring strategies for qualitative data collection. *The Counseling Psychologist, 35*, 295-327.

Turkle, S., & Papert, S. (1990). Epistemological pluralism: Styles and voices within the computer culture. *Journal of Women in Culture and Society, 16*, 128-157.

Walls, G. (2006). Meaning or medicine: The future of psychoanalysis in professional schools. *Psychoanalytic Psychology, 23*, 654-660.

Wampold, B.E. (2001). *The great psychotherapy debate: Models, methods, and findings.* Mahwah, NJ: Lawrence Erlbaum Associates.

Wertz, F.J. (2005). Phenomenological research methods for counseling psychology. *Journal of Counseling Psychology, 52*, 167-177.

Yanchar, S.C., & Slife, B.D. (1997). Parallels between multiculturalism and disunity in psychology. *American Psychologist, 52*, 658-659.

Appendix A

Informed Consent

Consent to be a Research Subject in the Therapist Effects Study

Introduction. The purpose of this study is to develop a procedure which will provide feedback to therapists about the outcome of their clients as measured by the Outcome Questionnaire (OQ-45). One assumption of this study is that reduction in OQ-45 scores is indicative of "good" therapy. The study will be conducted by David Dayton (faculty adviser is John Okiishi, PhD).

Procedures. By giving consent, you are giving permission for (1) OQ scores which were generated by your clients to be compared to group norms of the Counseling and Career Center (CCC) in the form of a feedback sheet, (2) CCC therapists to have the choice of retrieving or not retrieving their profile information, (3) David Dayton to investigate the effects of this feedback system by interviewing therapists/administering a questionnaire after they have retrieved their feedback packets.

Risks. There are minimal risks from participating in this study which may include the possibility of participants experiencing disappointment or discouragement and be caused to reflect on their role as a therapist. In such cases, therapists are encouraged to consult with colleagues in the CCC or seek qualified help outside the CCC.

Benefits. Previous research suggests that some psychotherapeutic and medical procedures are improved when practitioners are given feedback about their performance, CCC therapists' performance may improve by receiving this feedback.

Confidentiality. Therapist effects profiles are confidential. This information is not required to be shown or used in any way in CCC clinical team meetings, supervision meetings, stewardship interviews, rank advancement procedures or continuing faculty status reviews. Therapist's names will be linked to feedback packets in a blind procedure so that the therapist's identity cannot be associated with the feedback information by anyone other than the therapists themselves. Packets not retrieved within 60 days will be destroyed.

Participation. Participation in this research study is voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to your job.

Questions about the Research. You can contact David Dayton or John Okiishi if any questions come up. David can be reached at (801) 361-9235 or emailed at dayton@byu.edu. Dr. Okiishi can be reached at (80l) 422-6844 or john_okiishi@byu.edu.

Questions about your Rights as Research Participants. You can contact the Chair of the IRB, Renea Beckstrand, if you have any questions about your rights as a research subject. She can be reached at (801) 422-3873 or emailed at renea_beckstrand@byu.edu.

I agree to participate.   or   *I do not agree to participate*

_____   _____

(signature) (date)   *(signature) (date)*

Appendix B

Table 1

From Meaning Unit to Transformed Meaning Unit

Interview	Meaning unit	Transformed meaning unit	Collapse?
1	I don't understand [HLM] but I don't think I need a very sophisticated understanding to understand the data.	I'm satisfied with my imperfect understanding of this feedback.	No
	Is there some theme that I could make some sense out of my deteriorators?	This feedback has me wondering about my deteriorators.	No
	And I wouldn't want to threaten anybody and I wouldn't want to boast about my scores.	I'm concerned that, if made public, my positive feedback would alienate me from colleagues with less positive feedback.	No
	Um, it would be interesting to compare the BYU Counseling Center with another Counseling Center and get some sort of external comparisons.	I want to see how my feedback stacks up against clinicians in other organizations.	No
	I'd like to see my specific ranking out of all 71.	I want more specific ranking data.	No
	I wouldn't want you to jeopardize your study or your status or anything but I would like to see much more specific feedback.	I want more quantitative feedback.	No
	So what I'd like to do is I'd like to find somebody in here who's willing to be pretty darn open with this data and me and that person sit down and then talk about it, understand it. I'd like to do that with another person or	I want my quantitative feedback to be supplemented by dialogue with colleagues whom I trust.	No

small group.		
I would like the Counseling Center to be able to talk more openly about our data.	I feel dissatisfied with the degree of openness with this feedback in my organization.	Yes
I, in the last couple of years, since the whole feedback thing started, have become more free and less constrained in my therapeutic style. So this data moves me that direction a little more	After receiving this feedback, I feel more confident in my approach to psychotherapy.	No
This [feedback] bolstered my confidence and made me more enthusiastic about participating in this therapist effects study.	I felt confident about my approach after receiving this feedback.	No
I see myself as pretty open, willing to kind of share who I am. But that might be hard for me to talk about if my scores were in the bottom quartile.	I question my willingness to share my feedback with others.	Yes
I'm not sure how much I could invest in reinventing myself as a therapist. I think earlier in my career I would have been willing to reinvest a lot more.	I don't believe feedback could sufficiently motivate me to make major changes to my approach to psychotherapy.	No
I guess there could be an appeal made to me that, you know, this isn't that threatening if you're in the lowest quartile.	I don't believe this feedback is necessarily threatening.	No
So if the Center could help me in some way provide some training experience for me, that would help me be a better therapist, but would also help me be a better human being, then I'd be interested.	I am eager to participate in developmental interventions led by my organization.	Yes
I feel very different and I'm	I feel interpersonally isolated	Yes

	pretty reluctant to really share that at a deep level.	from my colleagues and I'm afraid to open up to them.	
2	I read [the feedback] and it wasn't helpful because I don't pay attention to the OQ very much.	I don't value the OQ-45 very much.	No
	I noticed that I had more deteriorators than the clinic average. I guess that was the part that I would have liked to have talked with you about, that probably would have been helpful for me to know, who those people were.	I want this feedback to be supplemented by additional feedback on my individual clients.	No
	I don't believe in holding back. I don't see why a person would.	I struggle to empathize with those who don't feel safe being open about their feedback.	Yes
	I don't use OQ analyzer at all. I don't even know how to get on there.	I am indifferent to some feedback based on the OQ-45.	No
3	I had no problem with [interpreting the data] at all.	I understood the feedback.	No
	When I got that [other] data back a few years ago, it really was, really was hurtful. Um, hurtful in the sense that it made me question, you know, if I was helping people as much.	Historically, this kind of feedback has made me question my efficacy as a psychotherapist.	No
	So this was dissonance for me—that here was objective data that suggests I'm not as good a counselor.	I experienced dissonance when I received feedback that was incongruent with my beliefs about my identity as a psychotherapist.	No
	When I made peace with it, then it's like, no I don't care that people know I'm in the third quartile.	Once I resolved this dissonance about the feedback, I was comfortable sharing it with others.	No
	And it was interesting that,	I felt reassured when a	No

	[another therapist] came to talk to me because he was in the third quartile and he was very upset by it and then on finding out that I'm in the third quartile made him feel better.	colleague whom I trust and admire shared [his] similar feedback results with me.	
	And at the same time, you know, I ask myself, if my OQ curve was really steep, I'd really need to ask myself what does that mean?	I am curious how I would have responded to this feedback had the results been more positive.	No
	The thing that happens in [team meetings] is expert teaching student. Um, and that there is a high value placed on expertise as opposed to a value placed on "I don't know."	The culture of my organization is too hierarchical and uncomfortable with ambiguity.	Yes
	See, it's not, people here are reluctant to engage in process. What I want to see happen in those team meetings is more process. And being able to give feedback in another way.	I want to engage in more collaborative and qualitative feedback processes in my organization.	Yes
	My biggest concern was not my slope but my deterioration rate, which would obviously affect my slope. I want to know what that means. Honest to goodness I don't know what that means exactly.	I want help making sense of my deterioration rate.	No
4	HLM slope means nothing to me.	I was confused by the data on the feedback forms.	No
	I don't even know, number one, if I trust HLM.	I do not trust the statistical methods used to generate the data on the feedback form	No
	Or it would be helpful to translate into terms I find	This feedback would be more meaningful if I knew how to	No

clinically useful.	apply it to my practice.	
I've had that explained to me about 5 times now. I keep forgetting it.	I do not understand this feedback.	No
The thing that I also found pretty useful is, I definitely want to know about deteriorators.	I want this feedback to be supplemented by additional feedback on my deteriorators.	No
I'd like to see what do peoples' numbers look like who are considered the best in the Center.	I'm curious about how my feedback results compare to the "best."	No
Honestly, the issue I have is I think the best methods require the most effort.	I'm unsure that investing in this feedback process is worth the effort.	No
But as far as, has [the feedback] affected what I do, you know, not really.	This feedback has not affected the way I do psychotherapy.	No
If I could really see that data a little better, then I think I might buy into it more.	If I could make more sense of the feedback, it might affect my approach to psychotherapy.	No
Here's your, what do they call it, clinical tools, and you've got a red [signal]. I don't like to do any of that stuff. It's too formal, it takes too much time.	I'm reluctant to invest time and effort into feedback processes that lack credibility in my mind.	Yes
I'm kind of weird like this, I'm more for openness than not. I'd be a little nervous to put my scores up on the screen in front of everyone.	I'm cautious about sharing my feedback with my colleagues.	No
Well, [the feedback form] says whether you're doing a good job or not. I mean, it's a report card for therapists, based on measures that we like slash question.	I question the credibility of feedback based on the OQ-45.	No
It's the threat, the ego threat. And I feel some of that, but	I feel threatened by this feedback.	No

	I'm a little bit, whatever. What are they going to do, fire me over it? Fine.		
	If [another clinician] gets the best results, put me behind the screen. Let me watch [that clinician] every week, an hour a week for a whole year. I'd love it.	I want to observe the psychotherapist who has the best results.	No
	Even now I wish I had more supervision.	I crave qualitative feedback.	Yes
5	I think it would be fun to have [feedback] on a semi-frequent basis.	I want to receive this type of feedback more frequently.	No
	I mean, I have curiosities about several of my clients and how they may have impacted my slope.	I want this feedback to be supplemented by additional feedback on individual clients.	No
	Yeah, the only confusing part was, is how against [the feedback forms] people were at the beginning and how everybody is on board at the end. That was confusing to me, more than the study itself.	I was fascinated by the evolution of this feedback system within my organization.	Yes
	I'm okay with being only an average clinician.	I'm did not feel threatened by this feedback.	No
	More data. I don't know what it would look like on another measure either. That would be kind of fun.	I want this feedback to be supplemented by additional measures.	No
	I really had to rethink, ok now, what does this mean about me as a therapist and where I held myself in comparison with the rest of the clinic.	This feedback compelled me to reexamine my identity as a psychotherapist.	No
6	But the way it came, because it's not clear how to interpret some of it, it was not so	I struggled to interpret and apply some of this feedback.	No

	helpful either because it was positive or negative, I simply did not know how to make use of some of it.		
	I think, generally speaking, we'd be better off if we had a little more training in how to understand the concept of slope.	I want help interpreting the statistics on the feedback forms.	No
	I would recommend that we invite people to get into peer-coaching relationships with a colleague.	I value engaging in collaborative feedback processes.	Yes
	One last thought, the assumptions that lie behind the OQ stuff…need to be discussed and made up front so interpretations can be made. Because there are other ways of measuring therapy besides the OQ.	I feel threatened by evaluative feedback based on one measures.	No
7	I found my results really upsetting and really struggled to make sense of them.	I was upset by this feedback.	No
	My initial response was one of, "Why am I not good enough?" And, "Better find another profession" and that sort of thing.	I noticed my tendency to engage in self-defeating thoughts when I received this feedback.	No
	When the initial kind of ego-wound sort of started to heal I began thinking about what kind of clients I see and what I could do to improve my curve.	I attempted to rationalize the dissatisfying feedback results.	No
	I could probably terminate a few people a little more quickly.	This feedback has caused me to reflect on how often I meet with my clients.	No
	But what I began to feel worried about was that if I	I believe it's possible to manipulate clients to make this	No

	really wanted to make myself look good, I could do a number of things that would be really unethical.	feedback more positively reflect my performance.	
	But my concern is that there's almost this assumption that…most people should be getting better by five or nine sessions.	I resent the belief by some people in my organization that most clients should recover so quickly.	Yes
	And I'm really hopeful that our faculty and the research management team, working with the clinical management team, kind of be cautious about how do you tell who's doing good work and how do you tell whether the client needs more therapy.	I'm concerned that this feedback data might be abused by people in positions of power.	No
	And maybe [the comparison with other therapists] is where I feel less comfortable because it's more threatening, but I also don't know what it means.	I feel threatened when my performance is compared to the performance of my colleagues.	No
8	And so I don't know how to make sense out of that [data], actually.	I struggle to interpret this feedback.	No
	So I think it would be kind of interesting to look at those two groups [recoverers and deteriorators].	I want this feedback to be supplemented by additional data on individual clients.	No
	I don't usually gear my [treatment] plan to trying to get OQ scores down, okay? Which I think some people actually do.	I believe people in my organization manipulate their clients in order to receive more positive feedback on job performance.	No
	And I'm not inclined to use the OQ as a therapeutic focus. So for me it just doesn't	This feedback is based on an outcome measure that is not congruent with my beliefs	No

	exactly fit with my style and my beliefs about therapy.	about therapy.	
	And I actually think there are people—I mean I hear people talk about—"You should be looking at the OQ and that should be the first thing you address in your therapy." I mean I hear faculty say that. I personally don't tend to agree with that.	I disagree with the belief held by some in my organization that the OQ-45 is very important to how you do psychotherapy.	No
	But I would like us to have, maybe, more discussions among ourselves about the OQ as an outcome measure versus OQ as a sort of indicator of how you should do treatment.	While I believe the OQ-45 is an adequate measure of outcomes, I question its utility as a measure of therapist efficacy.	No
	I mean I probably shouldn't say this, because I don't know who's in the top quartile, but I've heard so-and-so's in the top quartile, and happen to know that that person gets a lot of transfers—you know what I mean.	I believe the types of clients a clinician sees has significant bearing on the feedback forms.	No
	I'm not sure I ever want to be in the top quartile, frankly, because that would say to me, I'm just—I'm more concerned about my OQ scores than I am about the client.	I believe those who received positive feedback are more interested in looking good than helping their clients.	No
	Well, to me, getting this data is good. I mean I—you know, initially I was like, "Ahh, this is way too threatening." But when I got it I was like, "Whoa, this is okay."	This feedback was not as threatening as I anticipated.	No
	I don't know what they are,	I value some of this feedback.	No

	'cause I'm not that good at statistics, but this stuff at the bottom, to me, is really valuable.		
	But in terms of actual points, I wonder how really valid or significant that is.	I question the validity of this feedback.	No
	I think I'd like to include descriptions of what HLM means.	I don't understand HLM.	No
9	I think I understood it all (the feedback form). The only thing I didn't originally understand was the quartile ratings, and then you explained that to me.	I understood most of this feedback.	No
	When I look at [my feedback form] I want to know what went into that. Why? Where can I improve? Where am I doing well?	I want to be able to apply this feedback to my practice.	No
	So the more information I get, the more helpful the information is to me.	I believe additional data will help me interpret this feedback.	No
	So in this setting, to me, ranking probably creates more political problems than it gives helpful information.	I believe specific rankings would cause problems in the organization.	No
	But it's not helpful to me to compare myself vis-a vis my colleagues, unless there's a considerable difference.	I believe comparison is helpful only when there are meaningful differences between those being compared.	No
	Ranking, to me, it reminds me of a beauty contest.	I feel threatened by the possibility of specific rankings.	No
	And I felt really good. And that felt good to be able to say I'm doing a good decent job.	After I received this feedback, I felt reassured about my performance.	No
	So, for example, when I looked up how many deteriorators I	I believe I can make more sense from this data if provided with	No

had, I thought, "Why?" So just having a number doesn't tell me anything, it's just I have a lot of deteriorators. Who were they? Why? Is there a particular type of client I don't do well with? Is there something I can do to improve that? What does that mean when you say "deteriorators"?	supplemental information about individual clients, especially, "deteriorators."

When I looked at "No change," I thought, "Does that mean we've spent ten sessions for nothing?" Because the name of that would imply that –no change. So I just wasted ten hours of somebody else's life and my life.	I resent the implication of the terminology on the feedback form.	No
So just saying "no change" and then to rank that as, saying—and then for someone from outside the system to say, "Well you just spent 20 sessions doing nothing because they didn't get better. That rattles my chains, because I think it did do something to help them.	I resent the implication of this feedback that I have not been an effective clinician.	No
However [the OQ-45] is one piece of information, and it bothers me when it's some grand total of everything that happens in therapy, because I think there are other things that happen in therapy besides symptom reduction.	I'm upset by the implication of this feedback that symptom reduction is the only thing that matters in psychotherapy.	No
It's just, to me, a principle I was taught about in psychology and assessments,	I want quantitative feedback based on multiple psychological measures.	No

	I worry more about foreclosing on the truth by making premature interpretations of the truth based on limited information.	I am concerned the misapplication of OQ-45 data.	No
	And I think there is a mentality and a fear that where the OQ is going to be used as a hiring and firing tool. I think people fear that and I think that's premature.	I am not concerned that this feedback will be abused by people in positions of power.	No
	When you bring in alternative interpretations of data, you get shut down.	I feel voiceless in this organization when I disagree with popular opinion.	Yes
10	I'd like to get them every year, because I would love to know if what I'm learning is generally helping more or not.	I want to receive this type of feedback more frequently.	No
	However, everyone here is a competent therapist, and I think the mindset of comparing ourselves to each other is a little risky.	I am concerned about the notion of providing specific rankings on these feedback forms.	No
	I would love to know who to listen better in their clinical advice.	I want to be able to trust someone else's clinical expertise.	Yes
	And it was kind of reassuring to know that actually my clients are doing fine, that it's not a bad thing to do what I'm doing. I didn't feel personally validated exactly, however, I felt like what I do, it's good to know that it helps at a slightly above average rate.	I felt reassured about my performance by this feedback.	No
	But I would be interested in how mine do compared to how others do, because I don't know if that's average.	I want feedback that helps me know how effective I am at my job.	No

	I think I do better with OCD clients than most folks. However, I would love to know if I'm just making that up or not.	I want feedback that reassures my belief that I do well with certain clients.	No
11	Well, I want to know. I want to know my exact ranking out of the total number of therapists. I want to know.	I want to know my exact ranking on the feedback.	No
	That's like saying—you go to the doctor and he takes your weight and he doesn't tell you. And I want to know exactly where my weight falls in the set of normal {muddled}.	I feel entitled to receive all feedback data about my performance.	No
	Oh I 'd like to know every piece of information you have about me.	I want whatever data you can give me.	No
	And I think people who don't (want more data), I think people who won't admit that are in a way lying to themselves. Like there's some people who avoid going to the dentist 'cause they know they're going to have a cavity, right? But if they really look at it hard, they don't want to avoid going to the dentist, right?	I believe people in my organization who are opposed to receiving more data on their feedback forms are in denial.	No
	And this is nonsense about not giving me the score because it's bad for me. Well, it makes sense not to give someone a score if they don't understand it, but we understand it.	I believe people in my organization are competent enough to interpret this data.	No
	Well I will say that of course there's error in the OQ, but I have more faith in the OQ	I trust the OQ-45 as an outcome measure.	No

	than I have in anything else….it's a flawed instrument but it's better than everything else.		
	Well, I think it would be a really good idea to give us some way to drill down into the data to look at all our unsuccessful clients.	I want to exploit this data as much as possible, especially in regards to my clients who were "deteriorators."	No
12	I don't understand [the data] enough to know how I would [analyze it] differently.	I struggled to understand the feedback.	No
	There's more information here than I was expecting.	I did not feel prepared to receive these feedback forms.	No
	I don't think the form is relevant to the acceptability and the openness of the feedback.	The presentation of feedback was unrelated to the content of the feedback.	No
	I think there's some hesitance on my part about how it would be used still. I'm not absolutely sure that it wouldn't show up in rank advancement files, for example, 'cause that's one that I have a very, very strong feelings about. I don't want that showing up in rank advancement files.	I am concerned about how this feedback would be used by people in positions of power within my organization.	No
	You know, what is it that I do that's worthy of emulation, or what is it that I do that needs to be corrected. Those I see as extremely useful conversations and would hope that we would find a way to have them.	I want to have more dialogue with colleagues about how we do psychotherapy.	Yes
	The whole philosophy and concept behind that type of	I'm concerned about including rankings on the feedback	No

	thinking (ranking according to OQ data) is what I'm worried about.	forms.	
	Yeah, the feedback that I would like would be peer review, co-therapy, that sort of thing. To have somebody watching you work, somebody that's a peer. I don't want a dean of college-human performances. Is that what they're calling it now?	I crave more qualitative feedback from my peers.	Yes
	And I think this is useful stuff. I don't in any way want—I was one of the ones asking for it. I want it. I don't want it to be the end-all.	I believe this feedback is an important part of the whole (of feedback).	No
	(in response to differences since receiving feedback). Now I know in a sense if you were to watch my work you wouldn't see any difference. But I have been very active in questioning and wondering and thinking and what does it mean and how does it fit with t his client and what are the implications for this or for that. I've been much more, what's the word, curious.	This feedback has fomented introspection about my approach to psychotherapy.	No
	The piece that I'm looking for, are the ID numbers of the clients that were included in the study.	I want this feedback to be supplemented by additional information on my individual clients.	No
13	Actually, [the feedback] has got me thinking about one aspect. I used to be heavily worried about technique and if I was using the right	As a result of this feedback, I find my focus in therapy has shifted from techniques to the quality of the therapeutic relationship.	No

	techniques and interventions, and now I'm focusing more on the relationship aspect.		
	(in response to question about whether or not he'd like to see specific ranking) Yeah. I would.	I want to know my specific ranking in the feedback.	No
	Well, if it wouldn't be too labor intensive, it'd be interesting to see client that were particularly deteriorated, maybe all my deteriorators.	I want my feedback to be supplemented by data on my clients who were "deteriorators."	No
	That's probably actually in my interview so far, been like the most common kind of thing—is I want to see my individuals and their notes—like that's the bridge between this and making some kind of meaningful intervention.	Having looked at this feedback, I now want to reflect on the relationship between the qualitative and quantitative data I have for my individual clients.	No
	But that's kind of what I want to do is find a way to make this more like helpful, I guess. As it is right now it's kind of, "Ehh."	I struggled to apply this feedback into practice.	No
	I mean, I think [concern about rank advancement] is just flat out paranoia. I have no reason to believe that there's any directors or anybody around here that [has that] mentality—or even the threat of it.	I trust that people in positions of power in my organization will not abuse this information.	No
14	I don't know really how to interpret [HLM slope].	I don't know how to interpret some of this feedback.	No
	…is that an artifact of the way I see my clients for [fewer] sessions than somebody else, therefore, got to pat myself on	This feedback has caused me to reflect on and question my work as a psychotherapist.	No

	the back because, "Gee, I must be a good therapist with what I'm doing in session," or is this outcome because I just treat my clients less?	
	I don't know if I can be up to articulating the reasons why I'm in the [top] quartile.	I don't know understand the relationship between my positive feedback and what I do in psychotherapy.	No
	Now one of the things that maybe could help is if I were to get individual cases. So if I could look at individual trajectories, individual graphs and charts, and I were to look at all the deteriorators I had, and if I could look for some kind of pattern—gee, maybe I would know something about these people that would help me understand why they deteriorated.	I believe supplemental information on my "deteriorators" would help me identify thematic problems in my performance.	No
	I've told myself and told my supervisors now for three or four years that I consider myself a B+ therapist. I now have confidence that I'm at least a B+ therapist (laughs). That's not derogatory—I don't say that and feel like I'm denigrating myself. I feel like that's not bad. It also does not make me just want to be content and think, "Well, that's good enough."	This feedback is congruent with my beliefs about how effective I am as a psychotherapist.	Yes
	I don't worry about political or administrative uses of this feedback. I don't have that concern.	I am not concerned that people in positions of power will abuse this feedback.	No

	I don't want to discount [the OQ-45], in fact, I find it very helpful. I just think it's a piece of the puzzle and not the whole.	I consider the OQ-45 to be an important piece of a larger puzzle.	No
	I just don't think most of us are that sophisticated to take very much data and really make real good sense out of it.	I believe it takes considerable effort on the part of clinicians to make this feedback meaningful as it has been presented.	No
	In my mind the most helpful thing that this [feedback] can do is open dialogue—dialogues with myself, dialogues with my colleagues about effective therapy. And if it opens the dialogue, then I think it's been a very helpful thing.	This feedback has been helpful by facilitating dialogues about therapy with colleagues.	No
15	I think [specific rankings] would be interesting. I don't know if that would be any more helpful than just a quartile thing, but it would be interesting to know.	I'm ambivalent about including specific rankings on feedback forms.	No
	Again, it would help me more to see, "Okay well which ones may not have improved as much?" You know, just more specific demographic types of things.	I want this feedback to be supplemented by additional information on my individual clients.	No
16	Well, like you say, as far as the slope was concerned, I didn't understand that—er, the HLM intercepts. So that information there I wasn't really sure about.	I struggled to interpret the feedback.	No
	I'm not sure what you mean by the "recovered."	I don't understand the terminology on the feedback	No

		form.	
	(in response to my statement "Now it doesn't sound like this form was of that much interest to you at all?") You're right, it wasn't. Yeah.	I'm indifferent to this feedback.	No
17	I've always been a little confused by the HLM slope—the meaning of that.	I struggle to understand this feedback.	No
	And I think so long as we don't—well I guess everybody's fear would be if we start using the OQ to base salaries on, to base continued employment, to base rank advancement on, that we would be making a mistake.	I'm concerned that people in positions of power could abuse this feedback data.	No
	I think even more valuable, in some ways, is just looking at the individual OQ scores with each client all the time, and looking at their slopes, if you will.	I value the OQ-45 enough to discuss it with my clients.	No
18	Or whatever quartile I'm in, what does that exactly mean? That I could shift?	I struggle to interpret this data.	No
	I guess for me it's a number, it's a composite number, so I guess I'm not totally sure what I should do now with that number. But so what do I do with it? Does it mean I try to be more empathic? Do I try to be more directive?	I am unsure if I want to make the effort to be able to interpret this data.	No
	So qualitative, to me, is a little more meaningful in some ways.	I value qualitative feedback more than quantitative feedback.	No
	I'm just trying to brainstorm on what would be helpful to	I crave more open dialogue about psychotherapy	No

me when I get this number. And maybe just even meeting the better people or something and saying, "Hey what do you do in your practice?" If you were to say, "How do I improve?" what things would you look at?	approaches with my colleagues.	
(in response to question about how feedback has affected his approach to psychotherapy, if at all) Well I think I try to pay closer attention to my relationship to the client. So yeah, I think I try to really be more attuned to them, and focus on it.	After receiving this feedback, I find myself paying more attention to the quality of the therapeutic relationship.	No
My basic philosophy is [that] the relationship is critical, that if that's not there, no matter how much training I get, or feedback, it's probably not going to help me that much.	I believe quantitative feedback is relatively impotent.	Yes