The Question of Expertise in Psychotherapy

Scott D. Miller, Mark A. Hubble, and Daryl Chow
International Center for Clinical Excellence

Correspondence: Scott Miller, info@scottmiller.com

Abstract
Although it is well established that, on average, psychotherapy is effective, outcomes have remained flat for more than five decades. Since the 1990s, the effort to identify “empirically supported treatment” approaches has done little to alter this fact. Even more sobering, studies either fail to show therapists improve with specialized training or their outcomes steadily decline with time and experience. The aim of this paper is to illuminate how findings from the literature on expertise and expert performance illuminate new paths for the field of psychotherapy. Results to date point to new possibilities for helping practitioners realize improvements in the quality and outcome of their work.

Keywords
psychotherapy, psychological treatment, professional development, deliberate practice, routine outcome measurement

“We are all apprentices in a craft where no one ever becomes a master.” ~ Ernest Hemingway

Introduction
Following World War II, public and professional interest in psychotherapy exploded (Herman, 1995). In 1949, prominent psychologists convened in Boulder, Colorado, to “define therapy and establish criteria for adequate training” (Pope, 2003, p. 82). For all the time, energy, effort, and expense, at the end of the meeting, conference secretary, George Lehner, felt compelled to conclude, psychotherapy is “an undefined technique which is applied to unspecified problems with a nonpredictable outcome. For this technique, we recommend rigorous training” (Lehner, 1952, p. 547).

Since that time, much has changed. The field has professionalized. In the United States, psychological problems have been codified and refined in five successive editions of The Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013). Rigorous standards for preparation and clinical practice are in place and enforced by graduate training programs and regulatory bodies. At the same time, hundreds of methods have been developed, tested, and disseminated. Finally, decades of research provide overwhelming empirical support for the efficacy of the work (Hubble, Duncan, & Miller, 1999; Duncan, Miller, Wampold, & Hubble, 2010). Across a wide number of approaches, populations, and presenting complaints, in terms of client reported outcomes, the average treated person is found to be better off than 80% of those who do not receive treatment (Wampold & Imel, 2015).

The Problem of Training in Psychotherapy
For all the apparent progress, one major problem, originally captured in Lehner’s sardonic summary, continues to plague psychotherapy. In particular, the belief that “rigorous training” currently required for entering the field makes a difference in the quality and outcome of care practitioners provide. In the United States, doctoral training programs in psychology take between four and six years to complete and leave students, on
average, $100,000 in debt (Winerman, 2016). And yet, study after study reveals degree
professions perform no better than students (Boswell, Castonguay, & Wasserman, 2010; Christensen & Jacobson, 1994; Lambert & Ogles, 2004; Miller, Hubble, & Chow, in press). Millions are also spent annually on continuing education, including workshops, books, journals, instructional videos, and the like. Although mandatory for maintaining a license to practice, no evidence exists of any effect on results (Neimeyer, Taylor, & Wear, 2009; Webb, DeRubies, & Barber, 2010).

Another requirement for entering the field is working under the supervision of a senior clinician. While varying somewhat from jurisdiction to jurisdiction, and discipline to discipline, approximately 3,000 hours of supervision is the norm (Caldwell, 2015). Nevertheless, after reviewing research spanning a century, Watkins (2011) writes: “We do not seem any more able to say now (as opposed to 30-years ago) that psychotherapy supervision contributes to patient outcome” (p. 235). Using a large, five-year naturalistic dataset consisting of 6521 clients, seen by 175 therapists, who were supervised by 23 supervisors, Rousmaniere, Swift, Wagner, Whipple and Berzins (2016) confirmed and extended Watkins’s conclusions Once more, supervision was found not to be a significant contributor to client outcome. Going further, the supervisors’ experience level, profession (social work vs. psychology), and qualifications did not predict differences between supervisors in client outcomes.

The findings related to specialized training in so-called “evidence-based” approaches round out this grim assessment. In 1993, a Task Force within the American Psychological Association (APA) was organized to identify and promote a psychological formulary—“treatments of known efficacy” (Chambless & Ollendick, 2001, p.686). Though celebrated as an advance that would finally put the field on par with medicine (Nathan, 1997), subsequent research provided little support. In hundreds of randomized controlled trials pitting one method against another, none proves superior (Wampold & Imel, 2015; Wampold et al., 2017). For example, cognitive behavior therapy is compared with other bona fide approaches, such as interpersonal therapy, emotion-focused therapy, psychodynamic therapy, etc. Bona fide psychotherapies are treatments that are designed to be therapeutic, delivered by a trained therapists based on psychological principles, considered to be a viable form of treatment that has been presented to the psychotherapy community (i.e., via dedicated treatment manuals or books [Wampold et al., 1997]). Yet, training clinicians to use these approaches makes no difference in client outcomes (Rousmaniere, Goodyear, Miller, & Wampold, 2017).

As so much of conventional wisdom regarding what matters most for a good result has been shown to be immaterial, irrelevant, and inconsequential, it should come as no surprise that the overall outcome of psychotherapy has not improved in more than 40 years (Miller, Hubble, Chow, & Seidel, 2013). In their comprehensive review of the literature, Wampold and Imel (2015) report, “From the various meta-analyses conducted . . . the aggregate effect size related to absolute efficacy is remarkably consistent” (p. 94).

What has occurred at the “macro” level is reflected at the “micro” level. The evidence shows individual therapists do not get better with time and experience (Wampold & Brown, 2005, Chow et al., 2015). Worse, instead of improving, effectiveness plateaus early, then steadily declines (Miller & Hubble, 2011). In the largest study of professional development to date, Goldberg and colleagues (2016b) documented a diminution in performance, not unlike a slow leak from an inflated balloon. Importantly, the deterioration was unrelated to several factors often advanced as moderating variables, including client severity, number of sessions, early termination, caseload size, or various therapist factors (e.g., age, gender, theoretical orientation).

Absent a science of what reliably makes for improvement in the efficacy of psychotherapy, the field is poised to repeat the past. Fortunately, clues for a new direction and, ultimately, future are found in the behavior of a select group of practitioners.
The “Supershrinks” Show the Way

“Success leaves clues…” ~ Brad Thor

In 1974, researcher David F. Ricks examined the long-term outcomes of a “highly disturbed” group of adolescents treated in a large, metropolitan child guidance center. Not unexpectedly, results varied considerably. Some fared better in life. Others were eventually diagnosed as chronic schizophrenics. What was surprising was the variable found to account for the difference. Controlling for presenting diagnosis, gender, intelligence, social class, age, ethnicity, year of treatment, and frequency of psychotic or schizoid parents, the single most determinative factor was who had delivered the treatment. So good was one clinician, the children under his care called him “The Supershrink.”

While Rick’s (1974) report was occasionally cited over the next three decades, Okiishi, Lambert, Nielsen, & Ogles (2003) were the first to confirm empirically the existence of exceptional therapists in a large sample of practitioners. Gender, level and type of training, and theoretical orientation did not explain the difference between the most and least effective. What did, the authors wrote at the time, “remains a mystery” (p. 372).

Whenever the question of expertise in psychotherapy comes up—that is, highly effective and reliable performance—instead of reviewing actual client outcomes, studies have relied exclusively on peer nomination (i.e., the word of colleagues) for deciding who merits the distinction “master therapist” (Jennings et al., 2008; Jennings et al., 2005; Levitt & Williams, 2010; Skovholt & Jennings, 2005). It turns out, understanding the variability in outcome among individual clinicians would come from outside the profession. In particular, it was found in the extensive scientific literature bearing on the development of expertise (Colvin, 2008; Ericsson, 2009; Ericsson, Charness, Feltovich, & Hoffman, 2006). These findings are less concerned with the particulars of a given performance domain than how mastery of any human endeavor is acquired. This same research has been instrumental in identifying a set of processes anyone can follow to improve their performance (Ericsson et al., 2006).

In a series of articles, Miller and colleagues (2007, 2011) were the first to apply these results to the study of highly effective psychotherapists. They described three activities “the best” consistently pursue in their work: (1) establishing a baseline level of effectiveness; (2) obtaining regular, ongoing performance feedback; and (3) spending time outside of daily work in focused, systematic efforts to improve. Known in the expertise literature as “deliberate practice” (DP), meta-analytic studies document a significant association with performance (viz., .38° to .61) across a wide range of human endeavors (Macnamara, Hambrick, & Oswald, 2014; Platz, Kopiez, Lehann, & Wolf, 2014; Miller, Chow, Hubble, Wampold, Del Re, Maeschalck, & Bargmann, in press). The authors also noted the critical role context plays in encouraging and sustaining engagement in the three activities. This “culture of excellence,” as they called it, consists of a complex, interlocking network of people, places, and resources aligned to support and assist the learner.

In 2015, Chow, Miller, Seidel, Kane, Thornton and Andrews confirmed the applicability of the findings from the expertise literature to therapist development. The investigation examined the relationship between outcome and a variety of practitioner variables, including demographics, work practices, participation in professional development activities, beliefs regarding learning and development, and personal appraisals of therapeutic effectiveness. As has been found over and again, gender, qualifications, professional discipline, years of experience, time spent conducting therapy, and clinician self-assessments were not related to effectiveness (Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009; Wampold & Brown, 2005; Walfish, McAllister, O’Donnell, & Lambert, 2012; Malouf, 2012). On the other hand, DP was a significant predictor. In the first eight years of professional work, the top quartile of practitioners devoted, on average, 2.8 times more hours to the activity than those in the bottom three.

To date, only one study has been published documenting the impact of purposefully
implementing the three steps of the culture of excellence, as previously mentioned (Goldberg, Babbins-Wagner, et al., 2016). Specifically, attending to baseline performance via routine outcome monitoring, combined with systematic feedback and DP, incrementally improved the outcomes of individual therapists and overall agency results. Notably, therapists’ effectiveness increased every year, over 7 years, highlighting the potentially large, cumulative effect of small changes accrued over time (Ericsson et al., 1993; Imel, Sheng, Baldwin, & Atkins, 2015). The study further documented the importance of establishing a social context—including policies, procedures, administrative approval, and funding—for the pursuit of excellence (Ferlie & Shortell, 2001; Miller & Hubble, 2011).

The “Wet Edge” of the Paint

“You’ve got to bumble forward into the unknown.” ~ Frank Gehry

The use of standardized measures (i.e., routine outcome measurement [ROM]), and the role DP plays in improving effectiveness, are subjects of increasing interest among psychotherapy researchers (Boswell, Krause, Miller, & Lambert, 2013; Goodman, McKay, & DePhillips, 2013; Lambert, 2017; Prescott, Maeschalck, & Miller, 2017; Tilden & Wampold, 2017). For the first time in the history of the profession, trainees and clinicians have access to valid and reliable measures of therapeutic effectiveness and are able to compare their outcomes to nationally established norms (Shuckard & Miller, 2017; Delgadillo et al., 2017). For example, as of 2013, two systems (OQAnalyst and PCOMS) were reviewed and listed on the National Registry of Evidence-based Programs and Practices (NREPP) in the United States as well as being broadly used internationally. While studies so far confirm that ROM, and time spent in concentrated efforts to improve, can lead to better results (Chow, 2017; Chow et al., 2015; Goldberg et al., 2016a; Miller, Hubble, & Chow, 2017), as of yet, no study has identified how best the time should be spent.

The problem, as Tracey, Wampold, Lichtenberg, and Goodyear (2014) observe, is the lack of consensus “about how psychotherapy produces benefits, and the difficulty of using the information that does exist to improve . . .” (p. 219; Miller et al., 2013; Prado-Abril, Sanchez-Reales, & Inchausti, 2017). As noted earlier, much more is known about what does not work (e.g., training in specialized therapies for specific disorders, participation in supervision, and attending post-graduate continuing education). Despite this, some are using the term “deliberate practice” to refer to any training that involves elements of repetition and feedback—even in those just cited that contribute little if anything to outcome (c.f., Beck Institute for Cognitive Behavior Therapy, 2016; Milne & Reiser, 2014). Instead, based on past definitions of deliberate practice (Ericsson et al., 1993; Ericsson & Charness, 1994; Ericsson & Lehmann, 1996), the conceptualization of DP introduced here contains four primary elements:

1. A focused and systematic effort to improve performance based on clear identification of learning objectives, pursued over an extended period;
2. Guidance from a coach/teacher/mentor;
3. Immediate and ongoing feedback, and
4. Ongoing, successive refinement. (see Figure 1).

Figure 1. Four primary components of deliberate practice framework
One new, promising approach for conceptualizing and organizing DP efforts is based on empirically substantiated factors that have a high degree of influence on the outcome of psychotherapy, regardless of the method or diagnosis (Wampold & Imel, 2015). Known in the literature as the “common or therapeutic factors” these include the following: (1) quality of the therapeutic relationship; (2) creation of hope and expectation of change; (3) provision of plausible rationale and healing rituals; (4) understanding and use of client strengths and resources; and (5) therapist self-regulation (Hubble, Duncan, & Miller, 1999; Duncan, Miller, Wampold, & Hubble, 2011; Norcross, 2011; Laska & Wampold, 2014).

In 2015, Chow and Miller created the Taxonomy of Deliberate Practice Activities (TDPA), an assessment tool based on the five factors. Briefly, the TDPA is designed to map a clinician’s overall performance, making it possible to determine the degree to which a therapist’s in-session behaviors reflect what is known to be curative (Miller, Hubble, & Chow, 2017). From this information, targets for remediation or professional development may be identified, prioritized, and organized into an individualized plan informed by the four components of DP. A supervisor’s version is also available. It provides a concurrent, external check as well as information that can be used for designing learning exercises and evaluating trainee progress.

An ongoing series of studies, titled Difficult Conversations in Therapy (DCT), shows how a deficit in performance identified by the TDPA can be successfully remedied (Miller, Hubble, Chow, & Seidel, 2015). It is not uncommon for clinicians to demonstrate weakness in their ability to maintain a positive working relationship—a key category on the tool and, as it is, one of the most significant predictors of therapeutic success—when they are the target of a client’s anger or resentment. Practitioners are shown a brief video depicting just such a challenging situation. After watching, they are asked to respond as though they were in the room with the client. Their responses are, in turn, scored on a standardized relationship scale by two independent raters. Over a series of five successive trials, specific instruction for improvement, suggested by each therapist’s scores, is provided.

Results from the DCT reported thus far show that therapist performance did not improve with self-reflection alone. Instead, improvement only occurred when therapists received specific feedback, and were given time to reflect and revise their responses. Small, steady, but significant improvements in clinicians’ ability to respond warmly, empathically, and collaboratively accrued with each trial. The newly acquired skills were also found to generalize to other challenging situations. A randomised clinical trial of the DCT study is currently ongoing with four different challenging scenarios (Chow, Lu, Tan, Kwek, & Miller, n.d.).

Conclusions
Findings from the literature on expertise, coming from outside the field are providing inspiration for new lines of thinking and research aimed at improving the effectiveness of psychotherapy. The current evidence is promising, showing that the principles of deliberate practice can help individuals achieve better results in psychotherapy (Chow et al., 2015; Goldberg et al., 2016a). Such findings are now serving as a stimulus for rethinking how to go about training in a field where the outcomes have remained flat for many decades and in which practitioners frequently experience a deterioration in performance over the course of their careers (Miller, Hubble, Chow & Seidel, 2013).

In mental health services, using ROM to establish baseline levels of performance (between and among clinicians and agencies), and the preliminary investigations of DP, are encouraging. First, ongoing, real-time feedback significantly improves client outcomes. Second, compared to their less effective peers, higher performing therapists devote more time and effort to DP.

Such positive results notwithstanding, many questions remain. The initial findings summarized here require replication and testing.
with prospective, randomized controlled trials (RCT) comparing conditions with and without deliberate practice. In 2013, Miller et al. noted that several large-scale RCTs were underway. In the ensuing time, none have been completed. Additionally, the TDPA (Chow & Miller, 2015), and other schemes for conceptualizing and organizing deliberate practice efforts, await further development and validation. Finally, the nature and design of learning conditions that can best sponsor the development of higher performing therapists need much more thought and exploration.

Given current levels of research and interest, reason for hope exists that answers will be forthcoming.

Footnote
1. In the original version of Macnamara, Hambrick, and Oswald (2014), the published value was .35. That value was corrected to .38 in May 2018.

Authors’ Declarations
The authors declare that there are no personal or financial conflicts of interest regarding the research in this article.

The authors declare that they conducted the research reported in this article in accordance with the Ethical Principles of the Journal of Expertise.

The authors declare that they are not able to make the dataset publicly available but are able to provide it upon request.

References
Colvin, G. (2008). Talent is overrated: What really separates world-class performers from everybody


The Diagnostic and Statistical Manual of Mental Disorders (APA, 2014).


Malouff, J. (2012). The need for empirically supported psychology training standards. Psychotherapy in


