Leeds Alliance in Supervision Scale (LASS)

Instructions:
Please place a mark on the lines to indicate how you feel about your supervision session

(Approach)

This supervision session was not focused   I--I   This supervision session was focused

(Relationship)

My supervisor and I did not understand each other in this session   I--I   My supervisor and I understood each other in this session

(Meeting my needs)

This supervision session was not helpful to me   I--I   This supervision session was helpful to me

Wainwright, N. A. (2010). The development of the Leeds Alliance in Supervision Scale (LASS): A brief sessional measure of the supervisory alliance. *Unpublished Doctoral Thesis.* University of Leeds.

Leeds Alliance in Supervision Scale (LASS)

Notes on use

The Leeds Alliance in Supervision Scale (LASS) was developed as a sessional measure of the supervisory alliance.

The LASS is based upon a number of research measures designed to tap the supervisory alliance, and a number of alliance theories that underpin these measures.

The LASS should be completed **at the end** of each supervision session, in the last 10 minutes. Completion of the LASS provides an opportunity for the supervisee to provide feedback on how they felt about the supervisory working alliance in that session. This feedback can then be used as a discussion point, allowing an open discussion about how the supervisee and supervisor feel about the supervisory alliance.

Completion of the LASS in each supervision session also allows for changes in the alliance to be monitored and discussed by both supervisee and supervisor.

The aim of the LASS is to promote open feedback and discussion about the supervisory alliance so that it can be fostered and used as an effective component of clinical supervision.

If the LASS is to be used to track change over several supervision sessions, care must be taken to ensure that the lines used in the Visual Analogue Scales are always 10cm long.

The Role of Deliberate Practice in the Development of

Highly Effective Psychotherapists[1]

Abstract

Little empirical research exists about highly effective psychotherapists, and none about the factors that mediate the acquisition and maintenance of superior performance (e.g., Ericsson, 1996; Ericsson, 2006; Ericsson, Krampe, & Tesch-Romer, 1993). In the full sample, a three-level multilevel modeling (Level 1: clients, Level 2: therapists; Level 3: organization types) of practitioner outcomes was used to examine the contribution of the therapist to treatment effectiveness. Consistent with prior research, in the full sample (n = 69 therapists; n = 4580 clients) found that therapist effects explained 5.1% of the variance in outcome, after adjusting for initial severity. Therapist gender, caseload, and age were not found to be significant predictors. In a subsample of therapists, the relationship between outcome and therapist demographic variables, professional development activities and work practices was analyzed (n= 17 therapists, n = 1632 clients). Therapist characteristics (e.g., years of experience, gender, age, profession, highest qualification, caseload, and degree of theoretical integration) did not significantly predict client-reported outcomes. Consistent with the literature on expertise and expert performance, the amount of time spent targeted at improving therapeutic skills was a significant predictor of client outcomes. Further, highly effective therapists indicated requiring more effort in reviewing therapy recordings alone than the rest of the cohort. Caveats and implications for clinical practice, continuing professional development, and training are discussed.

Keywords: Deliberate Practice, Expertise in Psychotherapy, Feedback, Professional Development, Psychotherapy Outcomes, Therapist Effects.

The Role of Deliberate Practice in the Development of

Highly Effective Psychotherapists

With the exception of a few recent proposals (Miller & Hubble, 2011; Miller, Hubble, Chow, & Seidel, 2013; Tracey, Wampold, Lichtenberg, & Goodyear, 2014), the field has yet to systematically examine the development of superior performance in psychotherapy. Studies document significant variability in outcome among therapists, showing that the clinician accounts for a greater proportion of effectiveness than treatment modality (0%-1%) (e.g., Wampold, 2001; Wampold & Brown, 2005), client-rated alliance (5-15%) (Crits-Christoph, Gibbons, Hamilton, Ring-Kurtz, & Gallop, 2011; Horvath, Del Re, Fluckiger, & Symonds, 2011), or utilization of an empirically supported treatment (0%-4%) (Wampold, 2005). While researchers have provided rich and detailed descriptions of therapist personality characteristics, professional development activities, and work practices believed related to outcome, no studies have empirically examined the link between such traits and activities with therapy outcomes (Levitt & Williams, 2010; Orlinsky & Ronnestad, 2005; Skovholt & Jennings, 2005).

The study of expertise in other fields provides a potential model for understanding the key mediating factors involved in the development of top-level performers in psychotherapy. Across a variety of domains, including music (Ericsson et al., 1993; Krampe & Ericsson, 1996), chess (Gobet & Charness, 2006), sports (Cote, Ericsson, & Law, 2005), business (Sonnentag & Kleine, 2000), medicine and surgery (Norman, Eva, Brooks, & Hamstra, 2006), researchers have found that engagement in extended, deliberate practice, facilitates incremental development resulting in superior performance.

Deliberate practice (DP) is a highly specialized process. According to Ericsson and Lehmann (1996), it is defined as, "Individualized training activities especially designed . . . to improve specific

aspects of an individual's performance through repetition and successive refinement" (pp. 278-9).

Available evidence makes clear: experience involving the routine and proficient execution of skills associated with a particular performance domain is not enough to lead to improvement. To be effective, DP must be focused on achieving specific targets just beyond a performers current abilities, guided by the conscious monitoring of outcomes, and carried out over extended periods of time. Studies show that engagement in DP not only facilitates the development of superior performance, but is also crucial for its maintenance. For example, not only is engaging in deliberate practice at the early skills acquisition phase important, but it continues to be essential for older expert pianists, in order to maintain their level of expert performance (Krampe & Ericsson, 1996). However, there is less known about DP in psychotherapy.

Accordingly, the present study sought to determine whether DP accounted for the development of superior performance by examining the link between the nature and amount of time spent in efforts to improve performance and individual clinician outcomes aggregated over time. First, therapist effectiveness based on actual client outcomes will be determined after adjusting for initial severity. Next, the relationship between outcome and time spent by therapists in DP (e.g., solitary practice aimed at improving therapeutic engagement) will be determined. Finally, the specific nature of deliberate-practice activities explored as well as their relationship to differences in outcome among more effective therapists.

METHODS

Participants

Therapists. Participants were psychotherapists, psychologists, social workers, marriage and family therapists, and counsellors, practicing independently within the Human Givens Institute

Practice Research Network (HGIPRN)[2] in the UK. In order to be eligible for inclusion, practitioners had to have caseloads of 10 or more clients. Similar to other naturalistic outcome studies (e.g., Clark et al., 2009), each of cases attended a minimum of two sessions).

Based on the inclusion criteria, the final sample consisted of 4580 clients working with 69 therapists from 45 organizations. To ensure independence at the organizational level (several practiced at more than one organization), organizations were partitioned into six conceptually distinct types. Nine out of the 69 therapists were working in more than one type of organization. For the purposes of analysis, each of the nine therapists was coded with a primary organization type based on where they had treated the largest number of clients. The mean caseload of each therapist (treating clients with at least 2 sessions) was 66.38 clients ($SD = 70.03$, $Mdn = 40.00$; $Min = 10$; $Max = 335$). This consisted of 28 (40.6%) male therapists and 38 (55.1%) female therapists (3 therapists unspecified). The majority (52.2%) were between the ages of 41 and 55. In terms of organization, the majority were worked in the voluntary sector (42%), followed by the private practice (39.1%), Primary Care (NHS) (8.7%), Secondary Care (NHS) (4.3%), Insurance Based (2.9%), and Occupational Health and Counseling setting (2.9%).

In the subsample, 17 of the 69 original therapists voluntarily completed an online questionnaire about their professional development and work practices (Chow, 2014). Mean therapist caseload in this group was 94.24 ($SD = 97.40$; Mdn = 46; Min = 10; Max = 335), average years of experience, 8.45 ($SD = 5.24$). Briefly, 52.9% of the therapists were male, the majority (64.7%) were between the ages of 41 and 55, 58.8% were working in a private setting, and 52.9% were practicing as a professional psychotherapist. The average years of experience were 8.45 ($SD = 5.24$).

---

[2] For further details about HGIPRN, see Andrews, Wislocki, Short, Chow, & Minami, 2013

Clients. Only data from clients who were at least 18 years old were included. In the full sample, the mean age for the 4580 clients was 40.04 (*Mdn*: 40.00; *SD*: 12.86); 2999 (65.5%) were females and 1580 (34.5%) were males (one client's gender was left unspecified), and the majority were self-identified as white (66.2%). In terms of ethnicity, the majority was White (90.1%), followed by Asian (1.6%), Black (African, Caribbean, Other Black, or Black British) (0.6%), others (0.6%), mixed ethnicity (0.1%), and 20.8% were not stated. Consistent with most treatment settings, the majority of clients presented with concerns relating to anxiety (*n* = 3670; 74.90%) and depression (*n* = 2690; 59.58%). The average number of sessions attended was 4.72 (*SD* = 3.83). A total of 2503 clients (54.7%) had a planned termination (i.e., mutual agreement between client and therapist) with their therapist, 947 clients (20.7%) indicated an unplanned termination (i.e., client stopping therapy abruptly), and 1130 clients (24.7%) did not indicate a planned or unplanned termination. A total of 3632 out of 4580 clients (79.3%) provided information about the use of medication, and 1659 (36.2%) of these were prescribed a psychotropic medication.

In the subsample, data from 1632 clients were included. The mean age for these clients was 40.19 (*Mdn*: 40.00; *SD*: 13.20), of which 985 (60.4%) were female and 646 (39.6%) were male, and the majority (83.5%) white. Gender was not specified for one of the clients. Anxiety (*n* = 1357; 57.41%) was the major presenting complaint, followed by depression (*n* = 1234; 49.69%). The average number of sessions attended was 3.89 (*SD* = 2.57). A total of 1087 clients (66.6%) had a planned termination with their therapist, 265 clients (16.2%) indicated an unplanned termination, and 207 clients (12.7%) did not indicate a planned or unplanned termination.

Measures

CORE. Clients completed the 10-item version of the Clinical Outcomes in Routine Evaluation (CORE) (Connell & Barkham, 2007; Evans, Mellor-Clark, Margison, & Barkham, 2000).

The CORE is a 10-item questionnaire that assesses symptoms (anxiety, depression, physical problems, trauma), functioning (general functioning, close relationships, social relationships), and risk (risk to self, risk to others). Items are scored using a 0-4 Likert-type response format (from *not at all* to *most or all of the time*) rated over the past week. The recommended cutoff between clinical and non-clinical populations is 10, and higher scores indicating more distress. (Connell & Barkham, 2007). Internal consistency and test-retest reliability for the measure are both high ($\alpha = .94$; 1-week test-retest reliability, Spearman's $\rho = .90$ [Evans et al., 2002]). Total scores for the longer version (34 item version) and this shorter form (10 items) are highly correlated ($r = .94$) and, when used to assess outcome, have been shown to result in similar overall effect sizes (Andrews, Wislocki, Short, Chow, & Minami, 2013). The current study's average effect size $d = 1.22$ is equivalent to past naturalistic outcome studies based on the same outcome measure, CORE-OM (e.g., Clark et al., 2009; $d = 1.09$ for two treatment sites combined). Consistent with its length, internal consistency for this sample is somewhat lower than the longer version ($\alpha = .83$), but still in the acceptable range ($\alpha = .82$).

RAPIDPractice. The Retrospective Analysis of Psychotherapists' Involvement in Deliberate Practice (RAPIDPractice) is a survey instrument specifically designed to assess the amount of time a clinician spends in activities aimed at improving therapeutic performance.   Items were developed from both a systematic review of the literature on deliberate practice and its research methodologies (e.g., Cote et al., 2005; Ericsson et al., 1993) , as well as consultation with K. Anders Ericsson, the leading researcher on expert performance (Ericsson, 1996; Ericsson, 2004; Ericsson, 2006; Ericsson et al., 1993). RAPIDpractice contains 32 items, 7 of which seek information regarding professional background including, gender identification, age, discipline, training, certification, years of clinical experience, theoretical orientation). The remaining items capture the amount of time therapists spend in practice outside of work aimed at improving therapeutic skills. In this section, respondents asked

to rate: (1) the frequency with which engage in 25 activities (the amount of time spent in the last typical work month); (2) the confidence they have in their frequency rating from 0 (not at all confident in my time estimate) to 10 (highly confident in my time estimate); (3) the relevance of the particular activity to their improving clinical skills from 0 (not at all relevant) to 10 (highly relevant); and (4) the cognitive effort required for engaging in the activity from 0 (no effort exerted at all) to 10 (highest possible effort exerted) (Chow & Miller, 2012).

Procedure

Over a four-year period, data was gathered from practitioners who routinely administered a standardized outcome measure as part of treatment services (Andrews et al., 2013). Therapists from the full sample were invited to participate in assessing the relationship between practitioner characteristics, work activities, and engagement in professional development and deliberate practice. In this instance, data was gathered via an online questionnaire (c.f., Chow, 2014).

Data Analyses

In the full sample, multilevel modeling (MLM) was implemented through SPSS's Linear Mixed Model procedure (SPSS Version 19). Model parameters were estimated with maximum likelihood. All non-categorical explanatory variables were grand mean centered to facilitate interpretation of the intercept (Raudenbush & Bryk, 2002).

Data from the subsample on deliberate practice data were analyzed with a series of generalized linear mixed models (GLMMs). The GLMM is an attempt to overcome the limitations of classic least squares regression procedures (McCulloch, Searle, & Neuhaus, 2008). The GLMM can handle several types of non-normal outcome variables and accommodate a mixture of fixed and random effects. After controlling for intra-therapist and intra-organization-type dependencies, relationships between severity-adjusted client outcome and each of the conceptually distinct groups of therapist

explanatory variables were examined. The fixed effects were the initial (pre-test) CORE score and the therapist variables described below. The final (post-test) CORE score provided the outcome variable. Following the GLMM analyses, descriptive statistics were computed using standard statistical procedures (ANOVA, one-sample t-tests, correlation) in order to further investigate the relationships between therapist explanatory variables and client outcomes.

In all, there were seven therapist variables; specifically, three therapist characteristics (demographics, caseload, and degree of theoretical integration), time spent on deliberate practice, and time spent in three deliberate practice subdomain activities (see Table 2). Prior to the primary analyses, the relationship between each of the seven predictors and the adjusted client outcome was examined with a separate GLMM. The aim of these analyses was to identify any predictors not related to the adjusted client outcome, and were dropped from all subsequent analyses in order to avoid suppressor effects (MacKinnon, Krull, & Lockwood, 2000). The remaining therapist variables and adjusted client outcomes were analyzed with one or more GLMMs. Bonferronni adjustments were made to test alpha levels within each of the analytical groups in accordance with the number of GLMMs that were tested within a given group.

RESULTS

Full Sample. In order to analyze the between-therapist variability in client outcome for the first study, an unconditional model (i.e., no predictors) was first introduced in the analysis. The results from the three-level multilevel model revealed that the intra-class correlation (ICC) of therapist effects for the CORE was 5.35% (3.193/[3.193+56.514] = 0.0535; Wald $Z = 3.84$, $p < .001$). Next, an adjustment of clients' initial severity of functioning (i.e., pre-test score added as a covariate) was conducted, which serves as a base model for comparison with subsequent models. This resulted in the therapist effects for the CORE accounting for 5.10% of the total variance in client outcomes

(2.4262/[2.4262+45.206] = 0.0510; Wald Z = 3.84, p < .001). A significant proportion of therapist variability was still unexplained, so the three-level model was retained for all subsequent analyses. There were no significant differences in adjusted client outcomes across the treatment organization types (F[3, 12] = .49, p = .695). Additionally, no other client or therapist variables reduced the proportion of variance in outcome attributed to therapists (results available as on-line appendix).

Comparing therapists' performance. Therapists were grouped into quartiles based on their outcomes in order to facilitate a closer examination of performance differences. Differences between the four groups of therapists on a variety of outcome variables are reported in Table 1. Subsequent analyses employed the full dimensional value of these therapist variables i.e., without grouping therapists into quartiles.

Subsample: Deliberate Practice. Four GLMM analyses were conducted for the effect of deliberate practice on outcomes: (1) time spent alone in deliberate practice alone; (2) time in solitary activities; (3) time in non-solitary activities; and (4) time spent in nontherapeutic activities[3].

The first variable, average number of hours per week spent alone in deliberate practice, was grand-mean-centered and entered in the regression model. This was a significant predictor of the adjusted client outcome $(b$ = -0.016, SE = 0.007, t[1549] = -2.09, p = .037, η^2 = 0.003). In other words, a reduction in client distress was predicted by therapists who spent more time alone outside of work in deliberate practice activities. None of the other three predictors was significantly related to the client outcome (ps > .05, see supplementary material[a]).

To further examine the impact of this first variable, time spent alone on deliberate practice, the therapists from the top quartile were compared with the rest of the cohort. Of the original sample of

---

[3] Separate GLMMs were conducted on five clinician demographic variables (years of experience, gender, age [three levels], profession [five levels], and highest qualification [seven levels]) in order to determine the relationship between the characteristics and outcome. None proved to be significantly related to client outcome. Subsequent analysis also showed that caseload and integrative theoretical orientation were likewise unrelated to outcome (see online appendix).

69 therapists, the 17 were ranked 1, 3, 5, 7, 8, 13, 14, 15, 22, 27, 29, 30, 31, 36, 43, 44, and 54 in terms of client outcomes. As seen in Figure 1, on average, the top quartile group of therapists invested about 4.55 times more time on "deliberate practice alone" in a typical work-week, compared with the rest of the therapists.

Along with the amount of time spent for each of the 20 domain-specific and 5 non-therapy-related activities, each respondent was asked to rate on a Likert scale: (1) the relevance of the item to improving effectiveness; and (2) the cognitive effort required to engage in the activity (with 0 being, *not at all relevant/no effort extended* to 10, *highly relevant/highest possible effort exerted*). The relevance and cognitive effort ratings were not included in the MLM. Instead, based on recommendations made by Keppel and Wickens (2004), the items was rank ordered based on their means. A series of one-sample t-tests were conducted comparing the grand mean for the relevance rating to each item mean. This strategy was repeated for ratings of cognitive effort. Results are summarized in Table 2.

Of the 25 different activities, the following received significantly higher than average relevance ratings as regards improving effectiveness: "reviewing difficult/challenging cases alone," "attending training workshops for specific models of therapy," "mentally running through and reflecting on the past sessions in your mind," and "mentally running through and reflecting on what to do in future sessions," Live supervision provided during sessions was rated the least relevant (all *ps* < .05).

With regard to cognitive effort, therapists only rated "clinical supervision as a supervisee (review of difficult/challenging cases and/or non-progressing cases)," and "attending training/workshops," as requiring significantly higher than average effort (all *ps* < .05). None of the relevance ratings were significantly correlated with therapist average outcomes. Among the cognitive effort ratings, only "reviewing of therapy recordings alone" was significantly correlated with therapist average outcomes

($rs = -.665$, $n = 10$, $p = .036$, two-tailed). With the exception of "live supervision during sessions," "reading/re-reading core counseling materials," "writing down plans for future sessions," and "self-care activities and tending to emotional needs" the perceived relevance of all items was significantly correlated with the cognitive effort ratings (r^s ranging from .56 to .92).

DISCUSSION

After adjusting for initial severity and accounting for organizational level data, the analysis determined that 5.1% of the variance in outcome was attributable to the therapist—a finding consistent with prior research in the area (Baldwin & Imel, 2013). Consistent with prior research, therapist demographic variables (age range, gender, highest qualification, professional discipline, years of experience) failed to predict client outcomes (Beutler et al., 2004). Other factors, including theoretical approach, degree of theoretical integration/eclecticism, and size of caseload were tested but likewise did not predict client outcomes. With regards to caseload, researchers have noted the clinical relevance of feeling psychologically burdened with too many clients (Norcross & Guy, 2007). It is possible, however, that the caseload measure in this study was confounded, reflecting the total number of cases treated instead of a count of concurrent clients.

Deliberate practice, or the amount of time therapists spent in time spent alone in deliberate practice, was significantly related to outcome. The finding is consistent with results from numerous studies in different professional domains, including sports, chess, business, computer programming, teaching, medicine and surgery (Ericsson, et al., 1993; Duckworth, Kirby, Tsukayama, Berstein, & Ericsson, 2011; Gobet & Charness, 2006; Keith & Ericsson, 2007; Starkes, Deakin, Allard, Hodes, & Hayes, 1996). By contrast, no significant relationships were found between the amount of time spent in any of the specific activities surveyed and outcomes. In other words, no specific activity engaged in by therapists reliably led to better outcomes. Although contrary to expectations, the finding is

consistent with Ericsson and colleagues' (1993) whose investigation of violinists also yielded no profile differences in terms of the ratings of relevance, enjoyment, and effort of related activities. Given that the overall amount of deliberate practice was related to effectiveness, it may be that the utility of engaging in a specific activity depends on the needs, knowledge, skills, and competencies of the specific clinician. To date, no study has yet examined a taxonomy of deliberate practice activities for therapists to improve effectiveness.

To illustrate the impact of deliberate practice on the acquisition of expertise, trajectories plotting the amount of time performers of varying ability spend in deliberate practice over time were created for therapists in this study. Using a methodology similar to those used in prior studies (Ericsson et al. 1993; Charness et al., 2005), the amount of time spent in deliberate practice was based on self-report in a typical work-week. This figure, in turn, was first multiplied by 52 (weeks per year), then by years of experience. As the average years of experience for the subsample was approximately eight years, only these years of professional experience are included (see Figure 2).

As figure 2 illustrates, the estimated accumulative time spent by the top quartile (most effective therapists) spent, on average, about 2.8 times more hours per week engaged in deliberate practice activities aimed at improving effectiveness than the rest of the other therapists. Although intriguing, there are some caveats. First and foremost, the sample is small. Second, the number of hours spent in deliberate practice were both self-reported and based on retrospective recall, not actual measurement. Third, therapists in the study were at different developmental stages in their professional careers. As a result, time estimates may be related more to skills maintenance than skills acquisition, since the figure was indicative of each therapist's current ratings of the time spent in working to improve their clinical skills.

It is possible to speculate on a differential effect of time spent engaging in deliberate practice during the early professional developmental phase of skills acquisition, compared with the current maintenance of the relevant psychotherapeutic competencies. Given that most clinicians are likely to have spent more time in training during the initial phases of their careers, it would be advisable for future studies to include beginning therapists (e.g., Budge et al., 2013). This would be less susceptible to bias recall of time spent in deliberate practice alone, and also test the hypothesis if differences in early professional development is mediated by time spent in solitary practice.

In the subsample, four domain-specific activities received higher than average relevance ratings: (1) reviewing difficult/challenging cases alone; (2) attending training workshops for specific models of therapy; (3) mentally running through and reflecting on the past sessions in your mind; and (4) mentally running through and reflecting on what to do in future sessions. Given that these ratings were accumulated from clinicians across clinicians of varying effectiveness (e.g., the most and least effective), the meaning and relationship to performance is unclear. Clinicians could, for example, rate a specific activity highly relevant while simultaneously not engaging in it. Moreover, research on learning calls into question the reliability and validity of student and instructor self-ratings of effective methods for enhancing learning (Bjork & Bjork, 2011; Shea & Morgan, 1979).

Studies on deliberate practice highlight the role cognitive effort plays in the acquisition of domain-specific knowledge and skills (Ericsson et al., 1993; Ericsson & Lehmann, 1996). Of the 25 activities listed, one significant correlation was found between outcome and the cognitive effort ratings of "reviewing of therapy recordings alone." In other words, therapists with better outcomes rated the activity of reviewing of therapy recordings alone as requiring more cognitive effort than other activities. Researchers note the important role that reviewing performance recordings can play in the identification and remediation of errors (Abbass, 2004; Binder, 1999; Ericsson, 1996; Ericsson,

2006; Norcross & VandenBos, 2011). It is also possible to speculate that there is with more cognitive demands in a given practice activity, the learner is more likely to benefit in the process (Bjork & Bjork, 2011).

Across the entire cohort, the significant correlation between relevance and cognitive effort ratings of the majority of practice activities is consistent with previous studies (e.g., Ericsson et al., 1993; Starkes et al., 1996), suggesting activities deemed highly relevant also tended to be perceived as requiring high cognitive effort. In addition, two items received higher than average cognitive effort ratings: (1) clinical supervision as a supervisee (review of difficult/challenging cases and/or non-progressing cases; and (2) attending training/workshops for specific models of therapy. Once again, given that these ratings were accumulated from clinicians across clinicians of varying effectiveness (e.g., the most and least effective), the meaning and relationship to performance is unclear. Clinicians could have rated a particular item cognitively demanding while never actually engaging in the activity. A replication of this study with a significantly larger sample is ongoing and will hopefully provide the statistical power necessary for a more robust examination of the relationship between outcome and both the relevance of and cognitive effort required for engaging in specific domain related activities.

While the results are intriguing, and point to a heretofore unexplored topic in the area of professional development, the study does suffer from a number of limitations. Given its exploratory nature, many variables were included in the analyses, thereby increasing the risk of Type I error. In order to address this concern, Bonferroni corrections were employed in the second and smaller study in accordance with the number of GLMMs. Still, the risk remains. These initial findings await replication.

As already noted, the portion of the study related to deliberate practice had 17 participants. While previous studies of therapist effects have a range of sample sizes, from as large as 91 within a university counseling center (Okiishi et al., 2003), to as small as nine therapists in a mental health clinic for male veterans (Luborsky, McClellan, Woody, O'Brien, & Auerbach, 1985), the number of participants in the deliberate practice portion of the limits generalizability. More importantly, perhaps, is the nature of the sample. To begin, top-performers were over-represented, thereby increasing the homogeneity of therapist effectiveness and preventing more robust comparisons with poorly performing clinicians. Not surprisingly perhaps, therapists in the least effective group were less likely to participate in this study. The combination clearly leads to a risk of a self-selection bias limiting the generalizability of the results.

One additional limitation that deserves mention is the use of retrospective methods in assessing the amount of time spent in therapy-related and unrelated activities. While consistent with other deliberate practice research, the validity of such methods is a matter of debate (e.g., Charness et al., 2005; Ericsson et al., 1993; Law, Cote, & Ericsson, 2007). Unlike other performance domains (e.g., sports, music) where confirmation of deliberate practice time is more feasible (e.g., interviews with coaches, teachers, parents), the highly individualized and, to date, private nature of psychotherapy practice makes the cross-validation with independent raters impractical (excepting, perhaps, time spent in supervision, mentoring, staffing). Ratings done by others of an adult's solitary behavior presents obvious challenges. Prospective journaling, where individuals keep log activity over a set period of time may allow for more accurate estimates of engagement in deliberate practice (Shiffman, Stone, & Hufford, 2008). In real world clinical settings, such as this study, any such efforts would need to be carefully balanced with time pressures faced by practitioners. Finally, compared with Clark et al.'s (2009) average number of sessions of 6.35, the limited length of treatment in this study

(M = 4.72), might constrain the implications of the study when considering samples with longer treatments. Nonetheless, the majority of clients had a planned termination of the treatment process, which might suggest a briefer approach among this cohort.

In accounting for the reasons professionals do not spend more time in deliberate practice (aside from the lack of financial compensation for practice time), Ericsson (2009) notes:

Most professionals - such as doctors, nurses, stockbrokers, and accountants - do not receive the constant pressure from performing in front of an audience of paying ticket holders, like actors, musicians, and athletes. The lack of scrutiny and perhaps feedback may be an important difference that explains why many doctors do not spontaneously adopt the best practice methods for treating their patients, and spend a rather modest amount of time engaged in deliberate practice and effortful training to improve and maintain their skills… The greatest obstacle for deliberate practice during work is the lack of immediate objective feedback. (p. 422)

Ericsson (2009) points to one significant practice-related barrier to engaging in deliberate practice: lack of immediate feedback. Research documents that therapists routinely overestimate their effectiveness—on average, by about 65% (Walfish, McAlister, O'donnell, & Lambert, 2012; Miller, Hubble, & Duncan, 2007). Miller et al. (2007) argue persuasively that such findings may in part contribute to less time being devoted to improving performance. As both prior research and the present study confirms, experience is not a reliable predictor of outcomes (c.f., Beutler, 2004). Indeed, despite the early gains of skill-based competencies in their careers, professionals tend to plateau in their development (Ericsson, 2009).

Recently, a number of valid, reliable, and feasible measures for systematically monitoring progress in clinical practice have become available to practitioners for systematically monitoring progress in clinical practice, the routine use of which have been shown in multiple randomized

clinical trials to reduce dropout and deterioration rates (Hannan et al., 2005; Simon, Lambert, Harris, Busath, & Vazquez, 2012) while simultaneously improving treatment outcome (Lambert & Shimokawa, 2011; Miller, Duncan, Brown, Sorrell, & Chalk, 2006). In this regard, Ericsson (2009) indicates that the key aspect of feedback is pushing performers to "seek out challenges that go beyond their current level of reliable achievement-- ideally in a safe and optimal learning context that allows immediate feedback and gradual refinement by repetition" (p. 425). Possibly, deliberate practice in the psychotherapy profession can be specifically targeted the following areas: (a) Improving outcomes of at-risk cases; (b) creating social experiments in naturalistic settings to test, re-calibrate, and improve empathic accuracy (Sripada et al., 2011); (c) enhancing environments for targeted learning of fundamental therapeutic skills, such as rehearsing difficult conversations (Bjork & Bjork, 2011; Burns, 2009); (d) using standardized patients' simulated case vignettes to improve interaction with clients (Issenberg, McGaghie, Petrusa, Gordon, & Scalese, 2005; Ravitz et al., 2013), and (e) setting aside time to reflect and plan ahead (Lemov, Woolway, & Yezzi, 2012; Miller & Hubble, 2011).

The present study provides preliminary evidence for the role deliberate practice plays in the development of highly effective therapists. Parallel to the development of expertise in sports (Starkes et al., 1996), highly effective therapists spent more time engaged in activities outside of practice specifically aimed at improving performance while practicing. Deliberate practice might provide the necessary scaffolding for the development of therapeutic skills beyond a given therapist's current ability. Beyond devoting time to the process, it is not known what moderator variables may influence, enhance, or suppress the engagement in deliberate practice or the consolidation of the effects of deliberate practice on client outcomes. It is entirely possible, for example, that deliberate practice may need to occur in combination with other activities (e.g., traditional classroom training,

18

work-life balance, time off for personal activities), to enhance learning, skills acquisition, and maintenance of expert performance.

No longer treated as a nuisance variable (Garfield, 1997), therapist effects have become a serious focus of clinical trials and naturalistic research (Baldwin & Imel, 2013). The findings reported here are consistent with prior research. Clearly, therapists vary in their ability to engage and help clients. The results from the present study point towards the important role time spent in solitary practice accounts for such differences.

References

Abbass, A. (2004). Small group videotape training for psychotherapy skills development. *Academic Psychiatry, 28*(2), 151-155. doi: 10.1037/a0022427

Andrews, W., Wislocki, A. P., Short, F., Chow, D. L., & Minami, T. (2013). A 5-year evaluation of the human givens therapy using a practice research network. *Mental Health Review Journal, 18*(3), 165-176. doi: 10.1108/MHRJ-04-2013-0011

Baldwin, S. A., & Imel, Z. E. (2013). Therapist Effects: Findings and Methods. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 259-297). New York: Wiley.

Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling & Psychotherapy Research, 6*(1), 3-15. doi: http://dx.doi.org/10.1080/14733140600581218

Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., H., T., Noble, S. et al. (2004). Therapist Variables. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 227-306). New York: Wiley.

Binder, J. L. (1999). Issues in teaching and learning time-limited psychodynamic psychotherapy. *Clinical Psychology Review, 19*(6), 705-719. doi: 10.1016/S0272-7358(98)00078-6

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56-64). New York, NY: Worth Publishers; US.

Blatt, S. J., Sanislow, C. A., Zuroff, D. C., & Pilkonis, P. A. (1996). Characteristics of Effective Therapists: Further Analyses of Data From the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting & Clinical Psychology, 64*(6), 1276-1284.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*(3), 252-260. doi: http://dx.doi.org/10.1037/h0085885

Borkovec, T. D., Echemendia, R. J., Ragusea, S. A., & Ruiz, M. (2001). The Pennsylvania Practice Research Network and future possibilities for clinically meaningful and scientifically rigorous psychotherapy effectiveness research. *Clinical Psychology: Science and Practice, 8*(2), 155-167. doi: http://dx.doi.org/10.1093/clipsy/8.2.155

Budge, S. L., Owen, J. J., Kopta, S., Minami, T., Hanson, M. R., & Hirsch, G. (2013). Differences Among Trainees in Client Outcomes Associated With the Phase Model of Change. *Psychotherapy: Theory, Research, Practice, Training, 50*(2), 150-157. doi: 10.1037/a0029565

Burns, D. D. (2009). Advanced empathy skills. *Evolution of Psychotherapy Conference*, Lecture

    conducted from Anaheim, California.

Charness, N., Tuffiash, M., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The Role of

    Deliberate Practice in Chess Expertise. *Applied Cognitive Psychology, 19*(2), 151-165. doi:

    http://dx.doi.org/10.1002/acp.1106

Chow, D. (2014). *The study of supershrinks: Development and deliberate practices of highly effective*

    psychotherapists. (PhD), Curtin University.

Chow, D., & Miller, S.D. (2012). Retrospective Analysis of Psychotherapists' Involvement in

    Deliberate Practice (RAPIDPractice). Unpublished protocol. Chicago, IL: Authors.

Clark, D. M., Layard, R., Smithies, R., Richards, D. A., Suckling, R., & Wright, B. (2009).

    Improving access to psychological therapy: Initial evaluation of two UK demonstration sites.

    Behaviour Research and Therapy, 47(11), 910-920.

Connell, J., & Barkham, M. (2007). *CORE-10 User Manual, Version 1.1*: CORE System Trust &

    CORE Information Management Systems Ltd.

Cote, J., Ericsson, K., & Law, M. P. (2005). Tracing the Development of Athletes Using

    Retrospective Interview Methods: A Proposed Interview and Validation Procedure for

    Reported Information. *Journal of Applied Sport Psychology, 17*(1), 1-19. doi:

    10.1080/10413200590907531

Crits-Christoph, P., Baranackie, K., Kurcias, J. S., Beck, A. T., Carroll, K. M., Perry, K., . . . Zitrin,

    C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies.

    Psychotherapy Research, 1(2), 81-91.

Crits-Christoph, P., & Gallop, R. (2006). Therapist effects in the National Institute of Mental Health

    Treatment of Depression Collaborative Research Program and other psychotherapy studies.

    Psychotherapy Research, 16(2).

Crits-Christoph, P., Gibbons, M. B., Hamilton, J., Ring-Kurtz, S., & Gallop, R. (2011). The

    Dependability of Alliance Assessments: The Alliance-Outcome Correlation Is Larger Than

    You Might Think. *Journal of Consulting & Clinical Psychology, 79*(3), 267-278.

Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of

    comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59*,

    20-26.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate

    Practice Spells Success. *Social Psychological and Personality Science, 2*(2), 174-181. doi:

    10.1177/1948550610385872

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006a). Rejoinder to commentaries by

    Stephen Soldz and Paul Crits-Christoph on therapist effects. *Psychotherapy Research, 16*(2).

    doi: 10.1080/10503300500282186

Elkin, I., Falconnier, L., Martinovich, Z., & Mahoney, C. (2006b). Therapist effects in the National

    Institute of Mental Health Treatment of Depression Collaborative Research Program.

    Psychotherapy Research, 16(2).

Ericsson, A. K. (1996). The acquisition of expert performance: An introduction to some of the issues.

    In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the

    arts and sciences, sports, and games* (pp. 1-50). Mahwah, N.J.: Lawrence Erlbaum

    Associates.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine, 79*(10 Suppl). doi: 10.1111/j.1553-2712.2008.00227.x

Ericsson, K. A. (2006). The Influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683-703). Cambridge: Cambridge University Press.

Ericsson, K. A. (2009). Enhancing the development of professional performance: Implications from the study of deliberate practice. In K. A. Ericsson (Ed.), *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 405-431). New York, NY: Cambridge University Press; US.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363-406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of Maximal Adaptation to Task Constraints. *Annual Review of Psychology, 47*(1), 273-305. doi: 10.1146/annurev.psych.47.1.273

Ericsson, K. A., Nandagopal, K., & Roring, R. W. (2009). Toward a science of exceptional achievement: Attaining superior performance through deliberate practice. In W. C. Bushell, E. L. Olivio & N. D. Theise (Eds.), *Longevity, Regeneration, and Optimal Health: Integrating Eastern and Western Perspectives (Annals of the New York Academy of Sciences)* (pp. 199-217). River Street, Hoboken: Wiley-Blackwell.

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of

the CORE-OM. *British Journal of Psychiatry, 180*(1), 51-60. doi:
http://dx.doi.org/10.1192/bjp.180.1.51

Evans, C., Mellor-Clark, J., Margison, F., & Barkham, M. (2000). CORE: Clinical outcomes in routine evaluation. *Journal of Mental Health, 9*(3).

Garfield, S. L. (1997). The Therapist as a neglected variable in psychotherapy research. *Clinical Psychology: Science & Practice, 4*(1), 40-43.

Gobet, F., & Charness, N. (2006). Expertise in Chess. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The cambridge handbook of expertise and expert performance* (pp. 523-538). Cambridge: Cambridge University Press.

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

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70. doi: 10.2307/4615733

Horvath, A. O., Del Re, A., Fluckiger, C., & Symonds, D. (2011). Alliance in Individual Psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 48*(1), 9-16. doi: 10.1037/a0022186

Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., Lee G. D., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher, 27*(1), 10-28. doi: 10.1080/01421590500046924

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19. doi:10.1037/0022-006X.59.1.12.

Kahn, J. H. (2011). Multilevel Modeling: Overview and Applications to Research in Counseling Psychology. *Journal of Counseling Psychology April, 58*(2), 257-271. doi: 10.1037/a0022680

Keith, N., & Ericsson, K. (2007). A deliberate practice account of typing proficiency in everyday typists. *Journal of Experimental Psychology: Applied, 13*(3), 135-145. doi: http://dx.doi.org/10.1037/1076-898X.13.3.135

Keppel, G., & Wickens, T. D. (2004). *Design and analysis: A researcher's handbook* (4th ed.). Upper Saddle River, N. J.: Pearson Prentice Hall.

Klockars, A. J., & Hancock, G. R. (1992). Power of Recent Multiple Comparison Procedures as Applied to a Complete Set of Planned Orthogonal Contrasts. *Psychological Bulletin, 111*(3), 505-510.

Krampe, R., & Ericsson, K. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General, 125*(4), 331-359.

Lambert, M. J., & Shimokawa, K. (2011). Collecting Client Feedback. *Psychotherapy: Theory, Research, Practice, Training March, 48*(1), 72-79. doi: 10.1037/a0022238

Law, M. P., Cote, J., & Ericsson, K. (2007). Characteristics of expert development in rhythmic gymnastics: A retrospective study. *International Journal of Sport and Exercise Psychology, 5*(1), 82-103. doi: http://dx.doi.org/10.1080/1612197X.2008.9671814

Lemov, D., Woolway, E., & Yezzi, K. (2012). *Practice perfect: 42 rules for getting better at getting better*. San Francisco, CA: Jossey-Bass.

Levitt, H. M., & Williams, D. C. (2010). Facilitating client change: Principles based upon the experience of eminent psychotherapists. *Psychotherapy Research, 20*(3), 337 - 352.

Luborsky, L., McClellan, A. T., Woody, G. E., O'Brien, C. P., & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry, 42*(6), 602-611.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science, 1*(4), 173-181. doi: http://dx.doi.org/10.1023/A:1026595011371

McCulloch, C. E., Searle, S. R., & Neuhaus, J. M. (2008). *Generalized, linear, and mixed models* (2nd ed.). New York: John Wiley & Sons, Inc.

McMillen, C. J., Lenze, S. L., Hawley, K. M., & Osborne, V. A. (2009). Revisiting practice-based research networks as a platform for mental health services research. *Administration and Policy in Mental Health and Mental Health Services Research, 36*(5), 308-321. doi: http://dx.doi.org/10.1007/s10488-009-0222-2

Miller, S. D., Duncan, B. L., Brown, J., Sorrell, R., & Chalk, M. B. (2006). Using formal client feedback to improve retention and outcome: Making ongoing real-time assessment feasible. *Journal of Brief Therapy, 5*(1), 5-22.

Miller, S. D., & Hubble, M. (2011). The Road to Mastery. *Psychotherapy Networker, 35*.

Miller, S. D., Hubble, M. A., & Duncan, B. L. (2007). Supershrinks. *Psychotherapy Networker, 31*(6).

Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The outcome of psychotherapy: Yesterday, today, and tomorrow. *Psychotherapy: Theory, Research, Practice, Training, 50*(1), 88-97. doi: 10.1037/a0031097

Norcross, J. C., & Guy, J., D. (2007). *Leaving It at the office: A guide to psychotherapist self-care*. New York: Guilford Press.

Norcross, J. C., & VandenBos, G. R. (2011). Training audiotapes and videotapes. In J. C. Norcross,

    G. R. Vandenbos & D. K. Freedheim (Eds.), *History of psychotherapy: Continuity and*

    change (2nd ed., pp. 693-702). Washington, DC: American Psychological Association; US.

Norman, G., Eva, K., Brooks, L., & Hamstra, S. (2006). Expertise in medicine and surgery. In K. A.

    Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of*

    expertise and expert performance (pp. 339-353). Cambridge: Cambridge University Press.

Okiishi, J. C., Lambert, M. J., Nielsen, S. L., & Ogles, B. M. (2003). Waiting for supershrink: An

    empirical analysis of therapist effects. *Clinical Psychology & Psychotherapy, 10*(6), 361-373.

    doi: http://dx.doi.org/10.1002/cpp.383

Orlinsky, D., Ambuhl, H., Ronnestad, M. H., Davis, J. D., Gerin, P., Davis, M., & et al. (1999).

    Development of psychotherapists: Concepts, questions, and methods of a collaborative

    international study. *Psychotherapy Research, 9*(2), 127-153.

Orlinsky, D. E., & Ronnestad, M. H. (2005). *How psychotherapists develop: A study of therapeutic*

    work and professional growth. Washington: American Psychological Association.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data*

    analysis methods (2nd ed.). Newbury Park, CA: Sage.

Ravitz, P., Lancee, W. J., Lawson, A., Maunder, R., Hunter, J. J., Leszcz, M., . . . Pain, C. (2013).

    Improving physician-patient communication through coaching of simulated encounters.

    Academic Psychiatry, 37(2), 87-93. doi: 10.1176/appi.ap.11070138

Ricks, D. F. (1974). Supershrink: Methods of a therapist judged successful on the basis of adult

    outcomes of adolescent patients. In D. F. Ricks & M. Roff (Eds.), *Life History Research in*

    Psychopathology (pp. 275-297). Minneapolis: University of Minnesota Press.

Saxon, D., & Barkham, M. (2012). Patterns of Therapist Variability: Therapist Effects and the

    Contribution of Patient Severity and Risk. *Journal of Consulting & Clinical Psychology*

    August, 80(4), 535-546.

Seidel, J. A., Miller, S. D., & Chow, D. L. (2013). Effect size calculations for the clinician: Methods

    and comparability. *Psychotherapy Research*, 1-15. doi: 10.1080/10503307.2013.840812

Serlin, R. C., Wampold, B. E., & Levin, J. R. (2003). Should Providers of Treatment Be Regarded as

    a Random Factor?: If It Ain't Broke, Don't "Fix" It: A Comment on Siemer and Joormann

    (2003). *Psychological Methods December, 8*(4), 524-534.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and

    transfer of a motor skill. *Journal of Experimental Psychology: Human Learning & Memory*

    March, 5(2), 179-187.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual*

    Review of Clinical Psychology, 4(1), 1-32. doi: 10.1146/annurev.clinpsy.3.022806.091415

Simon, W., Lambert, M. J., Harris, M. W., Busath, G., & Vazquez, A. (2012). Providing patient

    progress information and clinical support tools to therapists: Effects on patients at risk of

    treatment failure. *Psychotherapy Research, 22*(6), 1-10. doi: 10.1080/10503307.2012.698918

Skovholt, T., & Jennings, L. (2005). Mastery and Expertise in Counseling. *Journal of Mental Health*

    Counseling, 27(1).

Sonnentag, S., & Kleine, B. M. (2000). Deliberate practice at work: A study with insurance agents.

    Journal of Occupational and Organizational Psychology, 73(1), 87-102. doi:

    http://dx.doi.org/10.1348/096317900166895

Sripada, B. N., Henry, D. B., Jobe, T. H., Winer, J. A., Schoeny, M. E., & Gibbons, R. D. (2011). A

    randomized controlled trial of a feedback method for improving empathic accuracy in

psychotherapy. *Psychology and Psychotherapy: Theory, Research and Practice*, no-no. doi:

    10.1348/147608310x495110

Starkes, J. L., Deakin, J. M., Allard, F., Hodges, N., & Hayes, A. (1996). Deliberate practice in

    sports: What is it anyway? In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of*

    expert performance in the arts and sciences, sports, and games (pp. 81–106). Mahwah, NJ:

    Erlbaum.

Tasca, G. A., & Gallop, R. (2009). Multilevel modeling of longitudinal data for psychotherapy

    researchers: I. The basics. *Psychotherapy Research, 19*(4-5), 429-437. doi:

    http://dx.doi.org/10.1080/10503300802641444

Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in

    psychotherapy: An elusive goal? *American Psychologist*, Advanced online publication. doi:

    http://dx.doi.org/10.1037/a0035099

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. J. (2012). An investigation of self-

    assessment bias in mental health providers. *Psychological Reports, 110*(2), 639-644. doi:

    http://dx.doi.org/10.2466/02.07.17.PR0.110.2.639-644

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

    NJ Erlbaum.

Wampold, B. E. (2005). Establishing Specificity in Psychotherapy Scientifically: Design and

    Evidence Issues. *Clinical Psychology: Science & Practice Summer, 12*(2), 194-197.

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists:

    A naturalistic study of outcomes in managed care. *Journal of Consulting & Clinical*

    Psychology, 73(5), 914-923. doi: 10.1002/jclp.20110

Wampold, B. E., & Serlin, R. C. (2000). The consequence of ignoring a nested factor on measures of effect size in analysis of variance. *Psychological Methods, 5*(4), 425-433.

Table 1

Therapists Grouped into Quartiles of Adjusted Client Outcomes (N = 69).

Q [a]	Initial score M (SD)	Adjusted CORE Score[b]	Raw ES[c]	RCI Mean (%)[d]	CS Mean (%)[e]	Deterio-ration (%)	No Change (%)	M No. of Sessions	Planned Ending M (%)	Un-planned Ending M (%)
1	19.57 (7.17)	8.75	1.50	76.6	58.9	1.6	22.0	4.52	73.4	12.2
2	20.10 (7.37)	10.75	1.25	67.6	46.5	2.2	30.2	4.44	60.5	15.7
3	19.02 (7.45)	12.07	0.99	57.8	34.2	3.4	38.9	5.74	51.7	23.4
4	20.05 (7.57)	14.17	0.76	50.3	25.2	4.7	44.5	6.30	27.7	22.5

Note. [a] Quartiles: Quartile 1= best performing therapists, Quartile 4 = poorest performing therapists. Sample sizes for each of the four quartiles were 17, 17, 17, and 18, respectively.

[b] Adjusted CORE score = posttreatment CORE scores, adjusted for initial (pretreatment) CORE scores. Lower scores indicate better outcomes.

[c] Raw ES = Single-group pre-post effect size using the SD of the pre-treatment scores.

[d] % of clients with a score change ≥ RCI (reliable change index; Jacobson & Truax, 1991)

[e] % of clients with a score change ≥ CS (clinical significance; Jacobson & Truax, 1991)

Table 2

Mean Relevance and Cognitive Effort Ratings for 20 Therapy-Related and 5 Non-Therapy-Related

Activities for Participating Therapists

Activities		Relevance		Cognitive Effort	
	N	M	SD	M	SD
1. General clinical supervision as a supervisee (without review of Audio/Visual recordings of sessions.	13	6.92	2.63	7.46	2.63
2. Clinical Supervision as a supervisee (with review of Audio/Visual recordings of sessions).	10	4.10	3.93	5.30	4.60
3. Clinical Supervision as a supervisee (review of difficult/challenging cases and/or cases with nil improvement).	12	7.67	2.77	8.00 **H**	2.76
4. Live supervision provided during sessions (e.g., supervisor as co-therapist, one-way mirror/reflecting team, etc.)	10	3.40 **L**	4.01	5.80	5.01
5. Reading of journals pertaining to psychotherapy and counselling.	14	7.21	1.67	6.71	1.68
6. Reading/Re-reading of core counselling and therapeutic skills in psychotherapy.	11	6.55	3.24	6.00	2.79
7. Focused learning in specific model(s) of psychotherapy.	11	7.27	2.87	7.45	2.98
8. Reviewing therapy recordings alone.	10	4.00	3.71	4.40	3.98
9. Reviewing of therapy recordings with peers.	10	4.30	3.83	4.50	4.04
10. Reviewing difficult/challenging cases alone.	14	8.00 **H**	2.77	7.43	2.59
11. Attending training workshops for specific models of therapy.	12	8.00 **H**	2.17	8.25 **H**	1.96
12. Case discussion/ conceptualization/ formulation	12	6.25	3.49	6.50	3.26

|					
with a mentor/clinical supervisor.					
13. Mentally running through and reflecting on the past sessions in your mind.	15	8.20 **H**	2.65	7.13	2.30
14. Mentally running through and reflecting on what to do in future sessions.	15	8.40 **H**	2.38	7.07	2.05
15. Writing down your reflections of previous sessions.	12	7.92	3.00	6.92	3.20
16. Writing down your plans for future sessions.	13	7.00	2.89	6.15	2.97
17. Case discussion/ conceptualization/ formulation with peers.	12	6.67	2.64	6.75	2.45
18. Viewing master therapist videos, with the aims of developing specific therapeutic skills as a therapist.	11	4.36	3.23	4.45	3.70
19. Reading case examples (e.g., narratives, transcripts, case studies).	12	4.33	3.23	4.92	3.40
20. Discussion of psychotherapy related subjects with contemporaries/peers/mentors.	13	6.85	3.02	5.69	2.66
21. Tending to self-care activities and emotional needs.	13	7.31	2.29	4.85	3.58
22. Socializing.	15	6.80	2.83	5.13	2.83
23. Exercising.	15	6.27	3.63	4.40	3.68
24. Rest (e.g., naps in the day, going for a walk, engaging in a non-therapeutic activity that is enjoyable)	15	7.40	2.59	4.13	3.54
25.Others (Please specify):	7	5.00	4.00	5.57	4.43
Grand Mean		**6.41**		**6.04**	

Note. *H = significantly higher than the grand mean; L = significantly lower than the grand mean at p < 0.05. Some therapists did not provide the rating of particular activities, as they did not engage in the activity in question.*

"Others" included musical activities, meditation, spiritual practices, child-rearing, reading non-therapy related topics, etc.
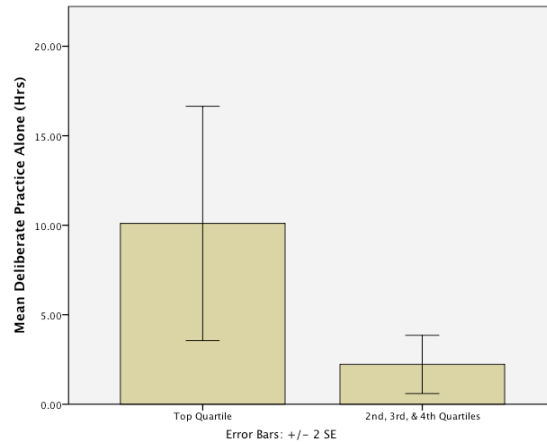
Figure 1. Comparing therapists from the top quartile with the others in the lower quartiles based on their adjusted client outcomes, as a function of estimated time spent on "deliberate practice alone" per typical work-week.
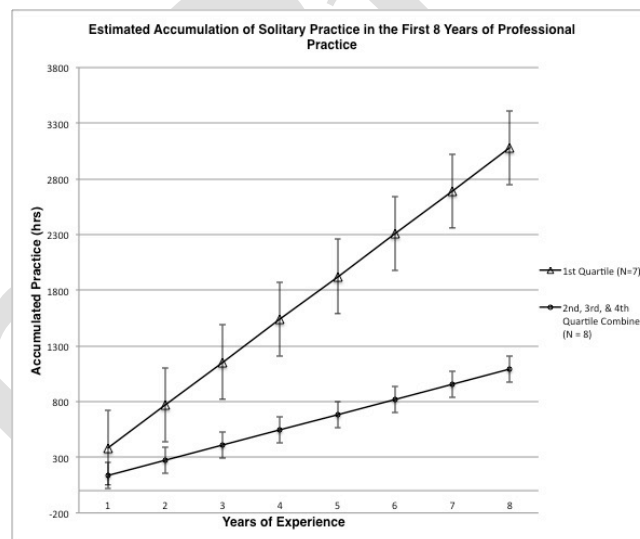
Note. Error bars = standard error of the mean (SE)

Figure 2. Comparing therapists from the top quartile with the others in the lower quartiles based on their adjusted client outcomes, as a function of their accumulative time spent on "deliberate practice alone" in the first eight years of clinical practice.

Online Appendix for

The Role of Deliberate Practice in the Development of

Highly Effective Psychotherapists

For the Main Sample (n = 4580 clients; n = 69 therapists)

Additional Client Variables:

When client demographic variables (gender and age) were added into the regression model, aside from the pre-test score, none reached significance as a predictor of the adjusted client outcomes (Pre-test score: β = .47, SE = .02, $F[1, 3946.35]$ = 987.23, $p < .001$; gender: β = .45, SE = .23, $F[1, 3794.95]$ = 3.76, $p = .053$; age: β = .02, SE = .01, $F[1, 3940.25]$ = 3.31, $p = .077$). Four other client variables relating to initial psychological functioning were added to the MLM (clients' severity ratings of their presenting concerns, duration of presenting problem, the number of presenting concerns, and whether they were on prescribed psychotropic medications). Three of these client variables were not significant predictors of the adjusted client outcomes (Pre-test CORE score: β = .47, SE = .02, $F[1, 1974.01]$ = 431.98, $p < .001$; severity: β = -.23, SE = .19, $F[1, 1583.33]$ = 1.43, p = .231; duration: β = .09, SE = .11, $F[1, 1732.74]$ = .55, p = .459; medication: β = .48, SE = .32, $F[1, 1969.14]$ = 2.30, p = .130), although the number of presenting concerns reached significance (β = .26, SE = .13, $F[1, 1522.20]$ = 4.03, p = .045). However, after including client variables about initial psychological functioning into the MLM, the ICC indicated that 6.60% of the total variance in client outcomes was attributable to the therapist $(3.055/[3.055+43.184] = 0.066)$. This is still a statistically significant proportion of variance (Wald Z = 3.04, p = .002). Furthermore, as these variables

collectively failed to reduce the ICC from the base model (5.1%), and previous research has also

failed to establish any link between these client variables and client outcomes (e.g., Wampold &

Brown, 2005), they were not included in the MLM.

Additional Therapist Variables:

Therapist caseload, gender, and age category were simultaneously added to the model. None of

these therapist variables, however, turned out to be a significant predictor of the adjusted client

outcomes (caseload: β = .0036, SE = .0025, $F[1, 45.95]$ = 2.12, p = .153; gender: β = .4209, SE =

.5582, $F[1, 53.12]$ = 0.5582, p = .454; age category: β = -.6555, SE = .3328, $F[1, 47.44]$ = 3.88, p =

.055).

For the Sub-Sample (n = 1632 clients; = 17 therapists)

Additional Therapist Variables:

Demographics. Years of experience, gender, age (three categories), profession (five categories),

highest qualification (seven categories) were initially analysed with *separate* GLMMs in order to

identify the predictors that were correlated with adjusted client outcome. None of the five predictors

was significantly related to client outcome: years of experience (b = .003, SE = 0.008, $t[1599]$ = 0.43,

p = .667, η^2 = 0.0001); gender (b = -.021, SE = 0.087, $t[1599]$ = -0.246, p = .806, η^2 = 3.7522E-05);

profession ($F[4, 1596]$ = 0.73, p = .571, η^2 = 0.0005); age ($F[2, 1598]$ = 2.73, p = .066, η^2 = 0.001)

and qualification ($F[6, 1594]$ = 1.020, p = .410, η^2 = 0.0006). No further analyses were conducted on

these predictors.

Caseload. The number of clients seen by a given therapist was indicated as *caseload*. This was

entered into a new regression model and was not related to client outcome (b = 0.000, SE = 0.000,

t[1599] = 0.998, p = .381, η^2 = 0.0006). No further analyses were conducted on this predictor.

Integration. Therapists in Study II were asked to rate the degree to which they regarded their theoretical orientation as integrative on a 6-point likert-scale from "*not at all*" to "*very greatly*". Integration was entered into a new regression model and was not significantly related to client outcome (b = 0.03, SE = 0.03, t[1599] = 1.07, p = .283, η^2 = 0.0007). No further analyses were conducted on this predictor.

The Retrospective Analysis of Psychotherapists' Involvement in Deliberate Practice (RAPIDPractice). None of the three subdomains of deliberate practice, Time in Solitary Activities, Time in Non-Solitary Activities, and Time in Non-Therapy Activities, were significantly related to the client outcome (b = -0.001, SE = 0.002, t[1549] = -0.34, p = .73; b = -0.001, SE = 0.004, t[1403] = -0.18, p = .86; b = 0.001, SE = 0.004, t[783] = 1.38, p = .17, respectively).

Feedback Informed Treatment

David S. Prescott, Cynthia L. Maeschalck, and Scott D. Miller

Contact information for corresponding author:

David S. Prescott, LCSW, LICSW

Director, Continuing Education Center

Safer Society Foundation

davidprescottvt@gmail.com

1

Learning Objectives

This chapter provides an overview of Feedback Informed Treatment (FIT), including its historical evolution, guiding principles, central tenets, and practical application. As a result of reading this chapter, readers will:

1) Learn about the research foundations of Feedback-Informed Treatment (FIT)
2) Learn about measuring therapeutic outcomes and the alliance
3) Consider diversity in FIT implementation
4) Lean about *deliberate practice* – individualized activities aimed at improving performance.

Keywords: psychotherapy; psychotherapy outcome; therapeutic alliance; feedback; deliberate practice; routine outcome measurement; psychotherapy integration

Introduction

A friend and colleague of the authors, Birgit Valla, recounts an unforgettable story:

> *I worked with a young boy accompanied by his mother. I read up on different therapeutic approaches for children and tried to follow the most appropriate method to the best of my abilities. With a supervisor and three fellow students monitoring everything I did from behind a one-way mirror, I was certain that the therapy was carried out in the best possible way. After every session, we discussed the therapy in a tutorial. I received good feedback on my relationship with the boy and everything of interest that emerged in our*

sessions. His statements were analysed in light of the issue in question along with my interventions as a therapist. The boy's mother sometimes attended the sessions, while on other occasions she sat in the waiting room. The therapy was carried out as prescribed, the boy and his mother seemed satisfied, and eventually we approached the final session. As a conclusion to the therapy, I had a session with the mother to sum up and ask her about how she had experienced the contact…

The mother arrived for our appointment and … seemed a bit uncomfortable and I asked her how she thought the therapy had been. She … was quite disappointed with all of it and did not think it had helped very much with the boy's problems. She'd had completely different expectations for the help they would receive and described what she had envisioned. It turned out that she had very clear ideas. I listened to her feedback, resisted the need to defend myself and hoped that I demonstrated understanding for her experience. We then ended the session and said our goodbyes. (Valla & Prescott, 2019, p. 2)

It seemed apparent that Birgit was providing good treatment with ample guidance from her colleagues, but clearly something was missing. Far from a unique situation, is the case represents a kind of outcome that could happen to anyone. Ironically, truly "bad" clinical experiences are easily observed ~~and scrutinized~~, while mediocrity can continue unnoticed for long periods of time. This can happen even to clinicians, therapists, and other professional who aspire to do their very best.

3

Coming out of this experience, Birgit started incorporating client feedback at the beginning and end of every session. Like many others, she has often been amazed by what does – and doesn't – make a difference in therapy. Feedback Informed Treatment (FIT) grew out of the discovery which, as Birgit experienced, when so many treatment approaches focus on models and techniques have has positive results across wide swaths of people, it's easy to forget how the smallest events can alter the course of treatment for better or worse at the individual level (Giorgi, 2011).

What Is Feedback Informed Treatment?

Since the late 1990s, FIT practice has evolved out of the recognition that not all therapies work for all clients, and that therapists do not always identify clients at risk of therapy failure. FIT is a transtheoretical approach that uses ongoing administration of outcome and alliance measures to collect real-time client feedback about their experience in therapy. "Transtheoretical" means that FIT can be applied across disciplines, no matter what treatment approach the therapist is using. The goals of FIT target two fundamental aspects of mental health treatment: (1) improving individual client outcomes by identifying clients at risk of treatment failure so adjustments to treatment can be made to get treatment on track for success and (2) increasing the effectiveness of individual therapists over time.

FIT is an evidence-based practice that merges "the best available research with clinical expertise in the context of patient characteristics, culture, and preferences" (APA

4

Presidential Task Force on Evidence-Based Practice, 2006, p. 273). It offers a practical system for the "monitoring of patient progress (and of changes in the patient's circumstances – e.g., job loss, major illness) that may suggest the need to adjust the treatment… (e.g., problems in the therapeutic relationship or in the implementation of the goals of the treatment)" (APA Presidential Task Force on Evidence-Based Practice, 2006, pp. 276-277).   FIT can improve the likelihood of positive therapy effects and encourage therapists to focus on growth and development as they attempt to improve their effectiveness throughout their careers.

Guiding Principles of FIT

FIT practitioners believe treatment should offer ~~some~~ benefits for ~~participating~~ clients. Continuing to provide service when there is no measurable improvement or where there is evidence of deterioration may be costly and unethical. For these reasons, clinicians should monitor the effectiveness of treatment with each client and make adjustments when clients report a lack of progress (e.g. change approach, intensity, or provider).

Since client engagement and participation are vital to positive treatment outcomes (Prescott, Maeschalck, & Miller, 2017), FIT therapists use empirically validated alliance and outcomes measures to monitor two factors that have been demonstrated to improve engagement in real time:

5

(1) the client's subjective view of the therapeutic alliance;

(2) the client's subjective experience of change.

The therapeutic alliance relies on client preferences and agreement between client and therapist about the goals of treatment, the nature of the relationship, and the means and methods used to reach those goals (Bordin, 1979).

Research into the power of the therapeutic alliance has been thoroughly established in over 1100 research findings (Horvath et al., 2011).

The therapeutic alliance is a strong predictor of the eventual outcome of therapy. A strong alliance generally leads to a positive outcome, while a weak alliance often ends with drop-outs and null outcomes (Prescott, Maeschalck, & Miller, 2017).

The client's perception and feedback about the therapeutic alliance offers a better forecast of outcomes than assessments made by therapists (Horvath et al., 2011). Research consistently shows that clients and therapists experience therapeutic relationships differently: clients may not necessarily perceive the therapist's actions in the way the therapist may have intended. The research suggests that clients and therapists often attribute therapeutic change to different experiences and influences (Horvath, Del Re, Flückiger, & Symonds, 2011). In order to ensure their effectiveness, FIT practitioners administer empirically validated alliance measures specifically designed to elicit the client's subjective view of the alliance on an ongoing basis. Formally monitoring the client's perception of the alliance through feedback measures at each session helps

6

clinicians identify when treatment is not going in the intended direction and clients may be at risk of dropping out, even when they appear to be engaged. ~~Early identification of flaws in the therapeutic alliance provides clinicians the opportunity to clarify client preferences and make adjustments to meet the client's needs, thereby strengthening the alliance and client engagement.~~

Monitoring the clients experience of change, especially early change, is the second factor routinely monitored by FIT practitioners. Early measurable change is a strong predictor of client engagement and, in turn, of therapeutic outcome. Several ~~large~~ highlight the importance of monitoring early change, suggesting that clients who experience little or no change in the first handful of sessions are likely to drop out or continue in treatment in the absence of change, while losing hope that therapy can help (Owen, Adelson, Budge, et al., 2015; Stulz, Lutz, Lucock, & Barkham, 2007). Lambert (2013) found that as many as 30% of people remain in treatment with no measurable benefit. In the same study, Lambert (2013) found a 90% chance of treatment failure if there is no change between the second and eighth visits. Thus, FIT practitioners strive to maintain engagement and ensure that change is ongoing.

In sum, FIT is an evidence-based practice that involves regularly collecting client feedback to measure the alliance and outcome, and discussing progress in these areas with the client. This information can help guide practice decisions. Maintaining engagement and ensuring that change continues is essential to successful therapy

7

outcomes. Routinely monitoring client feedback about their experience and progress in therapy:

1. Helps to inform and refine service delivery to best meet each clients needs and improve the chances of positive therapy effects;
2. Serves as an early-warning system for when treatment is going off-track as well as providing ideas about the best ways forward;
3. Assists professionals in improving their skills at helping people; and
4. Helps agencies to reduce the variability between clinicians with respect to their effectiveness.

Core Competencies of Feedback Informed Treatment

Four core areas of competence of Feedback Informed Treatment guide FIT practitioners (Miller, Maeschalck, Axsen, & Seidel, 2011; Prescott, 2017). ~~These are:~~

Competency 1: Research foundations~~, includes familiarity with:~~

- the research on the therapeutic alliance,
- research on behavioral healthcare outcomes,
- the general research on expert performance and its application to clinical practice, and
- the properties of valid, reliable, and feasible alliance and outcome measures.

Competency 2: Implementation~~, includes:~~

- integration of consumer-reported outcome and alliance data into clinical work,
- collaboration and transparency in interactions with consumers about collecting feedback,
- using the resulting information to inform and tailor service delivery, and
- ensuring that the course and outcome of behavioral healthcare services are informed by consumer preferences.

Competency 3: Measurement and reporting~~, includes:~~

8

- measuring the therapeutic alliance and the outcome of clinical services on an ongoing basis with consumers, and
- providing details in reporting outcomes sufficient to assess the accuracy and generalizability of the results.

Competency 4: Continuous professional improvement, ~~includes clinicians:~~

- determining their baseline level of performance,
- comparing their level of performance to the best available norms, standards, or benchmarks,
- developing and executing a plan for improving their performance, and
- seeking performance excellence by deliberate practice.

When and where does FIT apply?

For any treatment setting, modality, or client type, FIT offers valuable guidance that can be used in almost every situation. For example, FIT can be used in residential programs, intensive day treatment, outpatient and outreach settings with individuals, families, couples, groups, adults, adolescents, children, and even young children. FIT works with voluntary clients as well as clients who are mandated. FIT is also used successfully in many different countries and with many different cultural groups around the world.

Background Information: Why all therapists should use FIT in their practice

Psychotherapy is effective for reducing distress and improving well-being. ~~This has been solidly established through decades of research.~~ Consistently, clients who engage in treatment are better off than people in control or comparison conditions, or those who receive no treatment: between 0.8 and 1.2 standard deviations above the mean

of no treatment or control conditions (Lambert & Ogles, 2004; Smith, Glass, & Miller, 1980; Wampold, 2001). The effectiveness of psychotherapy is often equal to or greater than many medical treatments (i.e. chemotherapy for breast cancer; coronary bypass surgery for heart problems). Psychotherapy is also far more cost effective than many medical treatments and with fewer negative side effects (Schuckard, Miller, & Hubble, 2017; Wampold, 2007).

Even with the obvious benefits of psychotherapy, clients who no-show, dropout, and ~~clients who~~ do not progress despite the therapist's best efforts present an ongoing concern. The dropout rate for therapy remains stubbornly high -about 25% for adults and 35% for adolescents and children (Swift & Greenberg, 2012; Weisz, Sandler, Durlak, & Anton, 2005) and many clients who seek therapy do not benefit, or get worse. This can only create an enormous financial burden for the health care systems.

To make matters worse, clinicians often fail to identify clients at risk for premature termination or negative outcomes. Hannan and colleagues found that clinicians anticipated negative client outcomes in less than 5% of cases (Hannan et al., 2005). Clinicians' and clients' views of how therapy is progressing can stray widely, begging the question: What can therapists do to more accurately identify when clients are at risk of dropping out or when treatment is failing?

Routine outcome measurement (ROM) systems such as FIT can help. Over the last two decades, substantial research suggests that using client feedback about the

10

alliance and progress in treatment to inform and guide practice decisions provides significant improvements in the effectiveness of therapy. Lambert, Whipple, and Kleinstäuber's 2018 meta-analysis of 15 randomized clinical trials (RCTs) found that therapists who received ongoing feedback and clinical support tools showed outcomes with deterioration rates in at-risk patients improving from an average of 30% to 12% and increased by half the likelihood of at-risk clients experiencing positive outcomes.

Studying an additional nine RCT's including 2,272 clients Lambert et al. (2018) found that feedback produced an average effect size of .40 on client outcomes. When Miller, Duncan, Brown, Sorrell, and Chalk (2006) provided therapists with ongoing, real-time feedback regarding the client's experience of the therapeutic alliance and progress in therapy, they found higher retention rates and a doubling of the effect size (baseline ES = .37 v. final phase ES = .79). Ongoing client feedback regarding the working relationship and progress in treatment not only increased success rates but also improved the cost-effectiveness of services.

Insert graph

Frequency and immediacy of feedback also play a key role in improving the outcome of therapy. Slade, Lambert, Harmon, Smart, and Bailey (2008) found that immediate feedback had a larger impact on outcomes than feedback that was delayed by two weeks.

11

Some studies suggest that the impact of FIT may not have a positive impact on outcome or the effect may not be as pronounced as was first thought (Davidsen, Poulsen, Lindschou, et al., 2017; Mikeal, Gillaspy, Scoles, & Murphy, 2016; Østergård, Randa, & Hougard, 2020). However, these recent studies show that the impact of FIT is tied to differences in therapists and that it can take several years to implement successfully (often two to four years). Therapists who implement FIT require ongoing training and supervisory support to achieve the full benefit that FIT offers in terms of outcome. For example, Heidi Brattland and her colleagues (2018) studied the effects of FIT in an outpatient mental health setting. They found:

- Differences between therapists accounted for 9%–10% of the variability in outcomes.
- ROM was associated with better treatment outcomes independent of clients' initial distress levels.
- Over the four-year duration of the study, ROM had increasing superiority over treatment as usual.
- No effect in the first two years.

What are the implications of these findings?

- Therapist effects are essential to outcomes.
- It takes time to implement an effective ROM program.
- Training and supervision of therapists must be sustained over time.

What does FIT look like in practice?

FIT does not designate specific outcome and alliance measures that practitioners must use to gather client feedback (although it requires that they be valid and reliable

measures). Although many options are available to measure alliance and outcomes, longer formats take more time to administer, making them impractical for regular use. ~~Indeed,~~ The use of feedback measures decreases when they require more than five minutes to administer (Brown, Dreis, & Nace, 1999).

The Outcome Rating Scale (ORS; Miller & Duncan, 2000) and the Session Rating Scale (SRS; Johnson, Miller, & Duncan, 2000) are commonly used by therapists implementing FIT. These scales are brief measures that have proven feasible for regular use and demonstrated reliable and valid results. Designed to minimize complexity, the ORS and SRS are visual analog scales with no numbers or markers on the scale. Visual analog scales remove the preconceived values that people attach to numbers. Consequently, they also tend to have good face validity. So, rather that asking the client to use a numerical value to rate their experience or progress, the client is asked to mark a line at a point nearest to whichever end (positive or negative) is closest to their experience.

The ORS and SRS provide a snapshot of client well-being and functioning along with their experience of the alliance within each session that can be monitored over time. The ORS measures change in three areas of client functioning: individual (or symptomatic) functioning, interpersonal relationships, and social role performance (including work and quality of life). The SRS measures the major elements of a good therapeutic alliance including four interrelated domains of the therapeutic alliance:

13

- the quality of the relationship between the client and therapist,
- the degree of agreement on the goals of treatment,
- agreement on the methods used,
- and agreement on overall approach to therapy.

Age-specific versions of the ORS and SRS are available for use with adults, adolescents, and children. This makes FIT applicable to diverse client populations. A group session rating scale (GSRS) measures the client's alliance with other group members as well as the group facilitator. Oral versions of ORS and SRS are available for use over the phone and can be useful with clients who have problems with literacy, eyesight, or a strong preference against paper or electronic versions of the measures. As of 2020, the measures have been translated into 19 different languages and are in use around the world.

The adult versions of the ORS and SRS are designed for clients who are able to read at a grade six level or higher, and are normed for people 13 years and older. The Child Outcome Rating Scale (CORS) and Child Session Rating Scale (CSRS) can be used by children aged 6-12 with a grade two (approximately) or higher reading level. The simplified language in the measures means that they are also applicable to adolescents or with adults who have challenges in the areas of literacy or language. Adaptations of the measures are also available for use with very young children lacking literacy skills (typically under six years old). These are called the Young Child Outcome Rating Scale (YCORS) and the Young Child Session Rating Scale (YCSRS). They use three pictures representing a happy or smiley face, a neutral face, and a frowning or unhappy face,

which young clients choose to represent their experience. A fourth blank feeling face provides an option to draw their feeling if the other three faces do not adequately capture the child's experience. Because these measures do not have any numeric assignment or measurable scale, they do not provide a purely arithmetic measure of change. However, they engage young children in providing feedback, and that matters.

FIT practitioners gather feedback in real time. To be most effective, outcome measures need to be administered as early as possible at the initial session and each session thereafter to capture a baseline of the client's distress and measure change over time. The alliance measures should be administered at the end of each session. The initial alliance score provides a baseline from which the therapist works to improve engagement over time.

Tracking and graphing outcome and alliance scores allows the therapist and client to observe changes over time, providing an excellent basis for clinical discussion. Initially, this was done by plotting the scores on a paper graph. Since then, more technical solutions have been developed to facilitate tracking scores. These computerized applications allow clients to complete the measures on a tablet or other device. The software then scores and tracks client data, providing alerts when data indicates potential problems with the alliance or lack of progress. These systems are highly efficient, user friendly, and quick.

Typically, the session schedule and length of treatment depends on client progress. Psychotherapy is most effective when the client wants to continue with their therapy and meaningful change is evident. When positive change occurs rapidly, increasing the frequency and intensity of treatment can maximize gains. When positive change slows or diminishes, it may indicate a need to focus on consolidating change and transitioning out of therapy. The frequency of sessions may taper off with treatment lowering in intensity.

Therapists' views on collecting and using patient feedback can vary dramatically. Their attitudes can actually influence change. Therapists who value feedback achieve better outcomes (Miller, 2014). FIT measures change by outcome data and client feedback, not simply by the therapist's impression of improvement. Therapists can use any theoretical approach to achieve patient change, but if the data indicate that the approach is not working, the therapist is best positioned to make adjustments. Lack of change within the first few sessions (typically by the third session) merits evaluating the frequency, intensity, or method of treatment. Therapists and client working together can examine the alliance in detail and adjust their approach as needed. If the lack of significant change continues despite these adjustments, it may be time for bigger changes to the treatment approach or a referral to a different therapist.

16

Creating a Culture of Feedback

Even though FIT is being used in various settings using many kinds of treatment approaches and with diverse client populations, therapists sometimes doubt the relevance of FIT or the validity of client feedback. A common question is whether clients who have been diagnosed with severe and persistent mental illness or have cognitive impairment are capable of providing valid feedback. Validation studies do not support such conjecture (Prescott et al., 2017). In fact, research consistently indicates that it is therapist and not client variables that impact outcomes the most (Schuckard et al., 2017). In all cases, it is useful to obtain feedback regarding the client's experience. Besides providing a measure of change, FIT uses feedback to engage clients in dialog. Given the subjective nature of the measures, each client's scores hold unique meaning for them. The job of the therapist is to interpret the results and work with clients to find a way forward that works for each client. Outcome and alliance measures provide an overview a client's situation. Clarification of the meaning of the scores comes through dialogue with the client.

Using feedback to engage clients in discussions about their experience and progress in therapy contributes to improved outcomes (Hawkins, Lambert, Vermeersch, Slade, & Tuttle, 2004; Schuckard, et al., 2017). Therapists should create a transparent and open environment that encourages clients to provide detailed and honest feedback. Creating a "culture of feedback" from the onset of therapy is key.

For example:

17

Therapist (T): It's important that we make sure our time together has the best chance for success. It will help if you'd be willing to complete two quick measures each time we meet – one at the start and one at the end of our sessions. Most of the time, when treatment works, we should start seeing progress sooner rather than later. These tools help us see if the time we spend together is helping you. If things aren't getting better, then we'll talk about what changes we might make to get us going in the right direction. If you fill in the first measure at the beginning of each session, it will tell us if things are changing for you. The second one is for the end of each session. It will ask you how you think the session went and whether you believe I understand what you need. It's kind of like when your doctor gives you medicine for a fever and then checks your temperature to make sure the medicine is working. Would you be willing to do this with me?

Creating an environment where clients are at ease providing feedback is easier said than done. The desire to be viewed positively can leave clients fearful of the consequences of honest feedback and reluctant to give negative responses. Blanchard and Farber (2015) found that over 90% of adult psychotherapy patients report having lied to their therapists. The study reported one of the most common lies was that the client likes or benefits from their therapist's interventions more than they actually do. Unfortunately, fixing hidden problems is nearly impossible. Soliciting negative feedback is essential because it can guide efforts to meet client needs and align more closely with their

18

preferences. Ironically, the most effective practitioners receive more negative feedback than their more average counterparts (Miller, Duncan, & Hubble, 2007; Owen, Miller, Seidel, & Chow, 2016), clearly demonstrating how an open and forthright environment where clients feel confident giving honest feedback will benefit the therapy process.

FIT Methods and Techniques

Because clients are diverse, clearly no single treatment approach will work for everyone. When practicing FIT, outcome outweighs technique. The treatment approaches should depend on the probable effect they will have on client outcomes. When feedback scores indicate problems with the alliance and/or a lack of progress, therapists can tailor or change their treatment plan (for example, the frequency, type, or intensity of treatment method).

One of the biggest mistakes in implementing FIT happens when therapists don't value feedback measures adequately, or use them carelessly. This can include administering the measures without ~~actually~~ discussing the scores with clients or not using the feedback to guide therapy. Feedback summaries, including charts that display outcomes, should be shared with clients and form the basis of discussion around the alliance and therapeutic progress. Based on clients' input, therapists can adjust their approach to align with client preferences and ultimately better meet their needs. Once clients understand the purpose of feedback and its value in improving treatment effectiveness, therapists can administer the ORS as early as possible in each session

19

thereafter unless there is a specific reason not to, such as frequency of sessions. In order to prevent measurement fatigue, therapists should not expect clients to fill out the measure any more often than once per week. Administering the ORS as early as possible (ideally, in the first session) is useful because it provides a baseline of client functioning.

The SRS is best used close to the end of each session, leaving time to discuss the client's scores with them. This timing is essential so that the therapist and client can discuss elements of feedback that require attention. Since the alliance between therapist and client carries such influence and is predictive of treatment outcome, therapists should address any indication of an alliance problem or rupture as soon as possible. Acting quickly can prevent the premature termination or worsening of treatment.

Some outcomes patterns can raise red flags. When a client's initial score points to functioning in the normal rage of distress, therapists should proceed with caution. For example, in the case of a voluntary client seeking help with a specific concern, long-term, depth-oriented techniques may not be necessary, while problem-solving techniques focused on the immediate presenting concern may be more appropriate. Alternatively, when dealing with **involuntary or mandated clients**, asking them to complete the ORS based on how the party that sent them to treatment would score the measure can yield interesting results. This provides an opportunity to compare the clients' self-perception to that of the referral source (which may include judges or probation officers). It also fosters discussion about the kind of change the referral source would want to see in order to feel

satisfied that client has addressed the issues that led to treatment. In other cases, such as with adolescents, outcome scores that increase and decrease dramatically can indicate that the young person experience rapidly changing circumstances where they do not feel in control.

While having clients stay in therapy for longer periods of time can consolidate treatment gains (and prevent reversion to past behavior), diminishing returns can happen when the central goals are met, feedback scores plateau, and clients continue treatment without further demonstrable change. When clients have made positive changes and their scores level off, it is time for therapists and clients to explore reducing the frequency of sessions and begin planning to end services. If maintenance sessions retain the same frequency and intensity of earlier therapy, client scores can begin to fluctuate noticeably. This pattern can signal the ups and downs of the client's daily life rather than meaningful changes over time.

Sometimes, ORS scores will plummet after positive change has taken place – a phenomenon known as "ditching." Often, this is due to an extra-therapeutic variable – some event that takes place outside of therapy. If this happens, clinicians can explore each domain of the ORS in detail to establish whether external variables could be affecting the scores. Monitoring change over the ensuing weeks creates opportunity for quick recovery of these treatment gains.

21

Considering Diversity

People have been marginalized on the basis of their race, ethnicity, nationality, gender, age, sexual orientation, gender identity, religion, physical ability, socioeconomic status, and body size. Failing to consider the impact of marginalization can lead to power imbalances that contribute to clients' distrust and discourage them from providing honest feedback regarding their psychotherapy. Yet, cultural competence is rarely as straightforward as it might seem. In a recent review of the literature, Hayes, Owen, and Nissen-Lie (2017) debunk three prevailing myths about cultural competence, including that it involves a high level of education about different cultures, that therapists need to develop a therapy-specific language to work with cultural minority clients, and that therapists do not need to concern themselves with matters of culture.

Numerous studies have demonstrated significant differences both between therapists and within individual therapist's practice outcomes influenced by culture, race, and ethnicity (Flückiger, Horvath, Symonds, et al., 2013; Fuertes, Mislowack, Bennett, et al., 2007; Hayes, McAleavey, Castonguay, & Locke, 2016; Hayes, Owen, & Bieschke, 2015; Imel et al., 2011; Larrison, Schoppelrey, Hack-Ritzo, & Korr, 2011; Worthington, Mobley, Franks, & Tan, 2000). Such results commonly show that variations in effectiveness are unrelated to therapist gender, years of experience, self-perceived cultural competence, professional degree, theoretical orientation, or their own race or

ethnicity. As Hayes et al. (2017) emphasized in a review of the evidence, "No therapist is immune from having [cultural] disparities in their caseload" (p. 162).

Central to FIT is gathering feedback that can help clinicians shape services to fit the individual client. The very nature of FIT offers an opportunity to build specific guidelines for creating a culture of feedback that maximizes opportunities for attending to diverse client backgrounds, experiences, and ways for making sense of the world (Bertolino & Miller, 2013). In particular, the SRS is designed to alert therapists to:

- differences in understanding,
- goal consensus,
- and preferences related to identity, so they may be discussed and addressed directly.

This means that FIT, used properly, can be an excellent means for ensuring that cultural differences are respected in therapeutic practice, as well as for providing avenues for deliberate practice on the part of therapists.

Although FIT has proven versatile across many countries and cultures, providing direct feedback to health professionals is not always the norm. In some cultures, it may conflict with deeply held cultural as well as personal values. Even the outcome being measured may present challenges. One especially piquant example is the domain of personal well-being, which is included on many outcome measures. This domain may take on different meanings and be of less personal relevance to the client in cultures where people experience well-being within the context of relationships (Bertolino & Miller, 2013; Koo, Dion, & Rice, 2016). Creating space for the client's perspective in this

23

example could involve choosing a method and measure for seeking feedback in which well-being refers to the experience of the people closest to the client.

In another example, from the United Stated, older adults were frequently taught in their youth not to critique or evaluate a professional's activities, especially when they are in an expert position. Still other clients can be unwilling to provide feedback in every session. For some US adolescents, rejecting the measures can be an act of asserting independence, which can be developmentally appropriate. Cultural adaptations in the use of feedback can take place within the basic FIT principles of therapist flexibility amid strongly held client preferences. To this end, Miller and colleagues have produced a comprehensive and detailed series of FIT manuals. Manual 5 guides therapists on how to apply FIT across diverse settings, cultures, and clients (Bertolino & Miller, 2013).

In still another example of how culture can play a role in routine outcome monitoring, Rodriguez, Terrones, Brookman-Frazee et al., (2019) found that "stronger cultural heritage identity was related to potential harm with routine progress monitoring" (p. 1). The authors found that this concern about potential harm had little to do with the therapist's cultural background. They further suggest that their findings may be due to the potential for cultural misapplication of otherwise evidence-based assessment and treatment practices. These findings point to the importance of a nuanced understanding of FIT and how it can be used in the most helpful way possible.

So, how does FIT practice address cultural diversity more directly? After the translation of the ORS and SRS measures available in multiple languages, several FIT implementation efforts are now underway in various settings involving diverse populations. Unlike other measures to date, comparisons of clients from different countries and cultures have not shown differences in either the psychometric properties of the measures or the predictive trajectories (Koo et al., 2016; Schuckard et al. 2017; Miller, Bargmann, & Wampold, in preparation).

Next, the current algorithms are based on over a million cases. These have been tested and found not to vary across major ethnic and minority groups. These algorithms are in use with FIT software packages such as www.myoutcomes.com. Finally, in a recent publication, Miller, Hubble, and Chow (2020) describe methods for using aggregate data to check for outcome differences when working with various kinds of people, including along cultural lines, and offer suggestions for how to improve their performance using deliberate practice.

Case Example

As an example of FIT, consider the following experience from the first author's practice. It involves the integration of ongoing feedback into clinical services as well as using measures of outcome and the alliance to track outcomes. Finally, it also involves what the first author did to improve his clinical practice as a result:

> *Derek was 17 years old when his adoptive parents referred him for*

treatment. Adopted at the age of ten, Derek had previously received treatment for

concerns related to trauma and neglect in his family of origin. Having received

diagnoses from previous providers (such as Conduct Disorder and Attention

Deficit Hyperactivity Disorder), his problems revolved around attempting to view

his mother naked and in various stages of undress as she showered and got

dressed.

Treatment took place in an outpatient setting. Derek's mother had sought

out the author due to concerns that when he wasn't attempting to spy on her, he

sometimes hugged her to an inappropriate extent (e.g. ~~very~~ frequently and often

while he had an erection). Further, his mother was concerned that he seemed to

have difficulty negotiating conversations with young women in his school and was

fearful that he would act provocatively and inappropriately with them. Our first

session was straightforward, and Derek took the Outcome Rating Scale (ORS)

~~very~~ seriously (Derek would do this throughout our work). His initial ORS

scorewas 22 out of a possible 40, well below the cutoff for adolescents of 28. As

his therapist, I used motivational enhancement and cognitive-behavioral therapy

to explore the difference between Derek's current and desired future states.

At the end of the first session, Derek thought carefully about the SRSand

returned a score of 37, nearly perfect except for a score of 7 out of 10 on the item

related to the therapist's approach. Suspecting that this is part of a typical pattern

of responding on the SRS in which clients provide very high scores and the challenge for therapists is to elicit negative feedback, I asked for his thoughts on my approach. Derek stated bluntly, "You need to ask me a lot more questions. You need to be really hard on me or else I'll just lie to you, and I don't want to do that."

This feedback came as a surprise as I had specifically been using reflective statements in place of questions to avoid seeming like a criminal interrogator. More to the point, I have worked to develop this approach over many years, and here was Derek telling me to do the oppositeOf course, self-identifying as a provider of treatment X or Y may come at the cost of actual helpfulness with clients like Derek who don't like those particular approaches. Had I not explicitly asked for feedback, he may well have come to view me as naïve or unable to connect with him.

~~Of course, while~~ Derek may have been the expert on what would work for him, but I still possessed considerable knowledge about adolescents with problems related to sexual behavior. It was clear that if all I did was become brusque, I would not only lose our alliance but provide Derek with further opportunities for not engaging in treatment. I've often found that professionals in my field are who others turn to when directly confrontational approaches haven't worked.

27

Adapting my style with Derek required consistent checking in on the alliance to ensure that it was working. Indeed, Derek responded well to a combination of very direct feedback on my part mixed with a steady stream of validation. As soon as Derek began to offer excuses or externalize blame of his actions, I would say something like: "Look, I'm your biggest admirer, but nothing seems to be changing. With all due respect, if this behavior continues, what is going to happen?" This delivery turned out to be critically important. By being direct but not harsh, Derek still ended up, in essence, confronting himself.

Derek's SRS scores went up and generally hovered between 38 and 39 on this 40-point measure. He never provided a full 40, saying that absolutely no one is perfect, but offering no specific feedback beyond this. He did, however, come to take a diligent approach towards improving his relationships with others, most specifically his adoptive mother.

Ultimately, Derek's ORS scores rose to 32 and stayed there, despite small conflicts with his mother over keeping his room clean and attending to household obligations. His inappropriate behaviors with sexual overtones had stopped altogether. Statistically, his ORS scores indicated a reliable level of change.

In the end, Derek was able to make significant changes to his life on his own terms and not because of his therapist's preferred approaches. Being Derek's therapist meant stretching beyond my familiar approach at the time and

reaching for what would work for him in the long and short terms. What I learned

from Derek (and what other colleagues have observed informally) is that the

concept of confrontation is best considered in conjunction with a goal that the

client can accomplish; in the long run, it is not an effective technique on its own.

I developed new skills, such as being highly directive while remaining affirming,

and I have been able to use these refinements to my skill-set in other venues with

success.

Outcome Research

In 2012, the U.S. Substance Abuse and Mental Health Services Administration formally recognized FIT as an evidence-based practice and listed it on their National Registry of Evidence-Based Programs and Practices (https:/www.samhsa.gov/nrepp). Since then, the number of RCTs on FIT (using the ORS/SRS or some other combination of measures) has expanded with clinically, culturally, and economically diverse clients. FIT's effectiveness has been impressive: feedback and routine outcome monitoring reliably increases the rate of clinically significant change, consistently decreases deterioration rates and dropouts from treatment, and significantly reduces the cost of mental health care (Lambert et al., 2018). Schuckard et al. (2017) noted that the costs increased in non-feedback groups. Questions remain, however. For example, is it simply the use of measures to assess outcome and alliance that contributes to improved treatment

outcomes, or does the process of engaging people in their own care produce the most improvement? Can you even have the second factor without the first?

Focusing too closely on the measures themselves comes with the risk of distracting from outcome (Miller, Duncan, & Hubble, 2004). One ~~dismantling~~ study found that using both alliance and outcome measures did not translate into any significant increase in feedback effects when compared to using just one or the other to solicit feedback (Mikeal, Gillaspy, Scoles, & Murphy, 2016). These findings suggest that the process of asking clients about their experience of therapy may prove more important than what is being asked (i.e. which measures are used to gather feedback). These findings also highlight the importance of developing a collaborative treatment culture that engages clients in dialog about the treatment process as well as the change process itself. Further, it points to a reality that can be unpleasant for clinicians and administrators to realize: In some cases, improving treatment skills and services can take years.

Therapy without both outcome and alliance feedback presents limitations. For example, opting to administer only alliance measures hinders the ability to track client progress along with valuable aggregate data, including session-to-session change. While studies to date have not proven which measure – alliance or outcome – actually makes the critical difference, it's virtually impossible to know if the client is making progress without using both. Using outcome measurement alone may limit insight into fluctuations

in the strength of the alliance, which hold predictive value in terms of engagement and retention.

Future Directions

There are three principal future directions for FIT. These are: 1) increased dissemination and implementation; 2) more investigation into understanding when FIT is and isn't effective, and 3) improving the deliberate practice of therapists in response to feedback and routine outcome monitoring. Given its evidence-base and portability into any therapist's practice, it is no wonder many psychotherapists are implementing formal feedback systems to monitor their clients' progress and experience of the alliance. Still, FIT continues to evolve and remains a relatively new approach. Not all mental health professionals are using routine outcome monitoring in their practices. Some are not familiar with FIT, while others are hesitant to implement FIT. In one such instance, Babins-Wagner (2017) presented research on the value of outcome and alliance measurement and found that only 60% of therapists in an agency opted to use the measures. When asked, therapists typically offer many reasons for this (Babins-Wagner, 2017). Objections such as "it will take too much time," "it will conflict with my style," or "my clients won't like it" too often prevent therapists from implementing FIT. Even when therapists administer feedback measures, they often don't act on the information they receive. As just one example, Lutz (2014) found that when the feedback suggests a client's condition is worsening, they do not discuss it with clients about 60% of the time.

31

Furthermore, therapists provide clients with other treatment resources only about 27% of the time. Studies have also found that therapists only adjusted therapeutic interventions 30% of the time, varying the intensity or dose of services 9% of the time, and consulted with others (for example, through supervision or education) about 7% of the time (Lutz, 2014; Miller, 2014).

While some mental health professionals have reservations about implementing formal measurement processes, 92% of people in health care say they like the use of outcome measures (Lutz, 2014). Thus, therapists' attitudes are more likely than clients to prevent full implementation of FIT. Such findings suggest a need to help therapists explore and resolve their reluctance to implement FIT. Ultimately, the goal of dissemination and implementation is not only to share the value and methods of FIT, but to ensure that it is conducted appropriately and that clients receive the most effective services.

As encouraging as it may have been when early evidence showed client feedback and routine outcome monitoring to improve therapy, recent studies have returned less robust results (Lambert et al., 2018). Chow (2017) notes that "ROM and feedback studies are not immune to the decline effect … earlier studies have demonstrated therapeutic benefits of using feedback measures, but more recent studies have shown contrary results" (Chow, 2017, p. 325). Several recent studies have produced evidence of this decline effect. Other methods for providing therapy have been susceptible to this effect. It

may result from several factors, including the enthusiasm and allegiance of its developers

and early adherents, regression toward the mean in scientific studies, probable

publication bias, and the lack of treatment fidelity in later studies (Maeschalck, Prescott,

& Miller, 2019). Future researchers may wish to examine the boundary conditions of the

efficacy of FIT more closely, including when, how, and with whom it works.

Concerned that on its own, feedback is not enough to produce acceptable,

sustainable gains in therapy outcomes, Miller, Hubble, Chow, and Seidel (2013) have

looked more closely at the emerging empirical consensus that therapist factors influence

outcomes. "Available evidence documents that the therapist is one of the most robust

predictors of outcome among factors studied. Indeed, the variance of outcomes

attributable to therapists (5%–9%) is larger than the variability among treatments (0%–

1%), the alliance (5%), and the superiority of an empirically supported treatment to a

placebo treatment (0%– 4%)" (Miller et al., 2013, p. 90). They examined research

literature on experts and expertise in fields such as music, medicine, and sports. This

large body of research outside of psychotherapy provides clearer direction that can

improve outcomes.

They concluded that the most common avenues to superior performance consist

of three steps: (1) determining one's baseline level of effectiveness; (2) actively obtaining

systematic, ongoing feedback on performance; and (3) engaging in deliberate practice to

improve outcomes. While this can seem simple when described in the pages of a book, it can require very significant effort.

Deliberate practice involves individualized activities especially designed to improve specific aspects of an individual's performance through repetition, reflection, and successive refinement. To achieve the maximum benefit from feedback, individuals have to monitor their training with active concentration on a regular basis (Ericsson & Lehmann, 1996).

FIT can enhance outcomes at two levels. First, it informs the work one client at a time. Second, aggregate outcome data informs the therapist about their level of effectiveness compared to national norms via their relative effect size. By establishing a baseline of performance, and by analyzing outcome and alliance data, therapists can identify weak spots and practice skills for professional growth and improvement. Deliberate practice includes creating and executing a plan for improving one's performance. The emerging research indicates that therapists with superior outcomes experience professional self-doubt. They are inclined to dwell less on their successes, focusing instead on what they don't do well and how they can do better (Chow, 2017). In fact, the most effective therapists spend, on average, two to three times more hours per week engaged in deliberate practice activities than other therapists do (Chow, 2017). Recent developments recommend deliberate practice as a clear pathway to improved therapeutic outcomes.

34

Contraindications/Critique

Research studies to date have not identified clinical contraindications, at least not in the traditional sense of a particular treatment being inappropriate for a certain condition. People generally like to be asked about their experiences and whether things are getting better for them. Rather than saying that there are some clients for whom FIT won't be effective, it is more likely that some practitioners are not a good match for FIT. As seen earlier in this chapter, there are a number of conditions under which FIT may not be wholly effective. At the same time, it can take practitioners years to become more effective as a result of being feedback-informed. Therefore, the greatest contraindications are when:

- Clinicians do not want to use FIT or use it half-heartedly.
- Clinicians are unable, unready, or unwilling to create a culture of feedback in their practice.
- Clinicians do not use the measures correctly or for the purposes intended.
- Clinicians do not act on the feedback they do receive (thereby communicating to the client that they are eliciting feedback only in a pro forma fashion).
- Clinicians are unwilling to engage in deliberate practice and use the feedback they receive for the purpose of improving services.
- Clinicians believe that their outcomes may be used against them.
- Practice settings/agencies are uninterested in being feedback-informed or are not invested in supporting clinicians in using FIT.

Societal Example

Placing all of the above information in a broader context, what can we learn about feedback's role in other circumstances, across society? And how can this shape our

responses to numerous circumstances within and outside of clinical work? To start, many readers will have had the experience of making a complaint to a business. When the business responded favorably and took some action to improve the situation, it is likely that we became more loyal or regular customers. In the authors' experience, this is something that many business owners know, and yet clinicians don't: when one gets a complaint and addresses it appropriately, that business gains a loyal customer.

Of course, getting meaningful and helpful feedback can be harder than it seems, which is why market research is a highly specialized business. Ask yourself: Have you ever been in a restaurant where the meal was less than perfect? When the server came over and asked the familiar question, "How is everything?" were you completely honest? Or did you simply want to resume your mealtime conversation? The way one asks for feedback matters, as does the broader spirit of inquiry.

At a broader level, FIT practitioners will begin to notice all the ways that feedback is there for the taking, if only we are all willing to listen to and act upon it. The implications for human interaction are profound; it often seems that people do not listen to one another with a goal of understanding nearly as much as they think.

Conclusion

Increasingly, it appears to the authors that professionals feel obliged to adhere to one therapeutic model or another. In many circles, it is common to hear people describe not only the clientele they treat but the methods they use (for example, "I treat trauma

survivors using EMDR"). Given the effort and expenses involved in learning particular models and techniques, this is perhaps unsurprising. However, it might also come with a different price tag: **In psychotherapy research and practice, one size never fits all.** As useful as deep immersion in specific models and techniques can be, FIT enables the individual therapist, working with diverse clients, to foster change and become more effective one client at a time.

Additional Readings/Resources/Websites

Miller, S. D., Hubble, M. A., & Chow, D. L. (2020). *Better results: Using deliberate practice to improve therapeutic effectiveness.* Washington, DC: American Psychological Association Press.

Prescott, D.S., Maeschalck, C.M., & Miller, S.D. (2017). *Feedback-Informed Treatment in clinical practice: Reaching for Excellence.* Washington, DC: American Psychological Association Press.

www.myoutcomes.com (web-based platform for using the ORS and SRS that offers real-time feedback, expected trajectories of progress, and a number of other features)

www.centerforclinicalexcellence.com (website for the International Center for Clinical Excellence, a platform for training as well as sharing resources and knowledge about FIT)

Activity

Before diving head first into FIT, you may wish to reflect on some key elements within it. For example:

Ask yourself how effective you are compared to your peers. Are you average? Are you better than or less than average?

What factors led you to arrive at your conclusion? Was it purely the percentage of your clients that appear to improve, or are there other factors that influenced your decision?

How effective are you at identifying clients that are getting worse while in treatment? Given how poorly clinicians fare in this area in the extant research, what confidence do you place in your response?

What part of your clinical sessions do you think you could improve in (for example, how you work with strong client affect, how you begin or close a session, work with a model, etc.)? What are some small steps you can take to improve in these areas?

What kind of coaching is available to you as you try to improve in the areas that you just identified?

Chapter Summary in Bullet-Point Form

- The therapeutic alliance is central to all bona fide forms of psychotherapy
- The client's perception and feedback about the alliance is more predictive of outcomes than the therapist's perception.
- Monitoring the client's experience of change, especially early change, is crucial to ensuring client engagement and, ultimately, a successful treatment outcome.
- FIT is an evidence-based practice that involves regularly collecting client feedback to measure the alliance and outcome, and discussing these areas with the client.

38

- Routinely monitoring client feedback about their experience and progress in therapy:
    - helps to inform and refine service delivery to best meet each client's needs thereby, improving the chances of positive therapy effects;
    - serves as an early-warning system for when treatment is going off-track as well as providing ideas about the best ways forward;
    - assists professionals in improving their skills at helping people; and
    - helps agencies to reduce the variability between clinicians with respect to their effectiveness.
- There are four core competencies in FIT. These include
    - Research foundations
    - Implementation
    - Measurement and reporting
    - Continuous professional development
- Clinicians often fail to identify clients at risk for premature termination and other negative outcomes.
- Two measures that can be used in FIT are the Outcome Rating Scale (ORS) and the Session Rating Scale (SRS).
- FIT practitioners work to build a "culture of feedback" with their clients.
- It is crucial to use measures of outcomes and the alliance in the spirit in which they are intended.
- Current FIT algorithms are based on over a million cases. These have been tested and found not to vary across major ethnic and minority groups.
- FIT involves *deliberate practice* to improve outcomes. Deliberate practice involves individualized activities especially designed to improve an individual's performance through repetition, reflection, and successive refinement

NCE-Style Test Questions

1) In FIT, the therapeutic alliance:

    a) Involves agreement on the goals and tasks of treatment as well as the nature of the relationship, in accordance with strongly held personal values.
    b) Means that the therapist has an excellent relationship with clients and their supervisors alike.

c) Is central to all bona fide forms of psychotherapy.
d) A and C only

2) A "culture of feedback" involves:

   a) Firm but fair administration of positive and negative reinforcement to underperforming clinicians.
   b) Ensuring that all players have a voice in team meetings
   c) The client feeling free to discuss their experiences without retribution and with the hope of having an impact. (true)
   d) Skillful use by the therapist of metaphors involving feedback loops

3) Research finds that when change in therapy is going to happen:

   a) It mostly likely begins to happen early in treatment (true)
   b) It most likely happens in response to open-ended questions
   c) The client first needs to get worse before they get better
   d) It occurs quickly

4) Deliberate practice involves:

   a) Mindful awareness within agency settings
   b) Individualized activities designed to improve performance (true)
   c) Processing trauma slowly and deliberately
   d) Deliberately discussing unsuccessful cases in individual and group supervision

5) Successfully eliciting client feedback:
   a) Can require a fundamental shift in how one thinks about therapy and the roles of the client and therapist.
   b) Can be more difficult to accomplish than it seems
   c) Involves much more than gentle inquiries into how things are going in treatment.
   d) All of the above (true)

References

APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist, 61,* 271–285.

Babins-Wagner, R. (2017). FIT in Agency and Clinical Settings. In D. S. Prescott, C. L. Maeschalck, & S. D. Miller (Eds.), *Feedback Informed Treatment in Clinical Practice: Reaching for Excellence* (pp. 167-185)*.* Washington D.C. American Psychological Association Press.

Bertolino, B. & Miller, S.D. (Eds.) (2013). The ICCE Feedback Informed Treatment Manuals (6 Volumes). ICCE: Chicago, Illinois.

Blanchard, M. & Farber, B. (2015). Lying in psychotherapy: Why and what clients don't tell their therapist about therapy and their relationship. *Counselling Psychology Quarterly, 29,* 1-23. 10.1080/09515070.2015.1085365

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*, 252-260.

Brattland, H., Koksvik, J.M., Burkeland, O., Gråwe, R., Klöckner, C., Linaker, O., Ryum, T., Wampold, B., Lara-Cabrera, M., & Iversen, V. (2018). The Effects of Routine Outcome Monitoring (ROM) on Therapy Outcomes in the Course of an Implementation Process. A Randomized Clinical Trial. *Journal of Counseling Psychology*, *65*, 641-652. DOI: 10.1037/cou0000286.

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, and S. D. Miller (Eds.), *The heart and soul of change:*

   What works in therapy (pp. 389-406)*.* Washington, D.C.: American Psychological

   Association Press.

Chow, D., (2017). The practice and the practical: pushing your clinical performance to

   the next level. In Prescott, D., Maeschalck, C., & Miller, S.D. (eds.), *Feedback*

   Informed Treatment in Clinical Practice: Reaching for Excellence (pp. 323-355)*.*

   Washington D.C. American Psychological Association Press.

Davidsen, A. H., Poulsen, S., Lindschou, J., Winkel, P., Tróndarson, M. F.,

   Waaddegaard, M., & Lau, M. (2017). Feedback in group psychotherapy for eating

   disorders: A randomized clinical trial. *Journal of Consulting and Clinical*

   Psychology, 85(5), 484–494.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance:

   Evidence of maximal adaptation to task constraints. *Annual Review of*

   Psychology, 47(1), 273-305.

Flückiger, C., Del Re, A.C., Horvath, A. O., Symonds, D., Ackert, M. & Wampold, B. E.

   (2013). Substance use disorders and racial/ethnic minorities matter: A meta-

   analytic examination of the relation between alliance and outcome. *Journal of*

   Counseling Psychology, 60, 610-616. doi: 10.1037/a0033161.
42

Fuertes, J.N., Mislowack, A., Bennett, J., Paul, L., Gilbert, T.C., Fontan, G., & Boylan, L.S. (2007). The physician-patient working alliance. *Patient Education and Counseling, 66,* 29-36.

Giorgi, B. (2011). A phenomenological analysis of the experience of pivotal moments in therapy as defined by clients. *Journal of Phenomenological Psychology, 42, 61-106.*

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

Hawkins, E.J., Lambert, M.J., Vermeersch, D.A., Slade, K.L., & Tuttle, K.C. (2004). The therapeutic effects of providing patient progress information to therapist and patients. *Psychotherapy Research, 14,* 308-327.

Hayes, J.A., McAleavey, A.A., Castonguay, L.G., & Locke, B.D. (2016). Psychotherapist effects with White and racial/ethnic minority clients: First, the good news. *Journal of Counseling Psychology, 63,* 261-268.

Hayes, J.A., Owen, J.J., & Bieschke, K.J. (2015). Therapist differences in symptom change with racial/ethnic minority clients. *Psychotherapy, 52,* 308-314.

Hayes, J. A., Owen, J., & Nissen-Lie, H. A. (2017). The contributions of client culture to differential therapist effectiveness. In L. G. Castonguay & C. E. Hill (Eds.), *How*

and why are some therapists better than others?: Understanding therapist

effects (p. 159–174). American Psychological Association.

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in

individual psychotherapy. *Psychotherapy, 48*(1), 9-16.

Imel, Z. E., Baldwin, S., Atkins, D. C., Owen, J., Baardseth, T., & Wampold, B. E.

(2011). Racial/ethnic disparities in therapist effective- ness: A conceptualization

and initial study of cultural competence. *Journal of Counseling Psychology, 58,*

290–298. http://dx.doi.org/10 .1037/a0023284.

Johnson, L. D., Miller, S. D., & Duncan, B. L. (2000). *The Session Rating Scale 3.0*.

Chicago: Authors.

Koo, M., Dion, D., & Rice, A. (2016). *FIT and cultural competency*. Unpublished

Manuscript.

Lambert, M.. (2013). Outcome in Psychotherapy: The Past and Important Advances.

Psychotherapy (Chicago, Ill.). 50. 42-51. 10.1037/a0030682.

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy.

In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and*

behavior change (5th ed.) (pp. 139–193). New York: Wiley.

Lambert, M.J, Whipple, J.L. & Kleinstäuber, M. (2018). Collecting and delivering

progress feedback: A meta-analysis of routine outcome monitoring.

Psychotherapy, 55, 520-537.

Larrison, C.R., Schoppelrey, S.L., Hack-Ritzo, S. & Korr, W.S. (2011). Clinician factors

　　related to outcome differences between black and white patients at CMHCs.

　　Psychiatric Services, 62, 525-531.

Lutz, Wolfgang, (2014, December). Why, when and how do patients change? Identifying

　　and predicting outcome in therapy. http://kpplutz.uni-trier.de;

　　https://www.slideshare.net/scottdmiller/lecture-wolfgang-lutz-calgary2014-send.

Mikeal, C., Gillaspy, J.A., Scoles, M.T., & Murphy, J.J., (2016). A Dismantling Study of

　　the Partners for Change Outcome Management System. *Journal of Counseling*

　　Psychology, 63, 704-709.

Miller, S.D. (2014). Dinner with Paul McCartney. Blog post;

　　http://www.scottdmiller.com/feedback-informed-treatment-fit/1327/) December

　　11, 2014.

Miller, S.D., Bargmann, Minami, T. & Wampold, B. (in preparation). Feedback

　　informed treatment: Developing predictive trajectories of change from a

　　naturalistic international data set.

Miller, S. D., & Duncan, B. L. (2000). *The Outcome Rating Scale*. Chicago: Authors.

Miller, S. D., Duncan, B. L., Brown, J., Sorrell, R., & Chalk, M. B., (2006). Using formal

　　client feedback to improve retention and outcomes. *Journal of Brief Therapy, 5,*

　　(1), 5-22.

45

Miller, S. D., Duncan, B. L., & Hubble, M. A. (2004). Beyond integration: The triumph of outcome over process. *Psychotherapy in Australia, 10,* 2-19.

Miller, S. D., Duncan, B. L., & Hubble, M. A. (2007). Supershrinks. *Psychotherapy Networker, 31*(6), 26-35, 56.)

Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The Outcome of Psychotherapy: Yesterday, Today, and Tomorrow. *Psychotherapy, 50,* 88-97.

Miller, S. D., Hubble, M. A., & Chow, D. L. (2020). *Better results: Using deliberate practice to improve therapeutic effectiveness.* Washington, DC: American Psychological Association Press.

Miller, S. D., Maeschalck, C., Axsen, R., & Seidel, J. (2011). The international center for clinical excellence core competencies. Retrieved from http://centerforclinicalexcellence.com/wp-content/ plugins/buddypress-group-documents/documents/1281032711-Core Competencies.pdf.

Østergård, O.K., Randa, H., & Hougaard, E. (2020). The effect of using the Partners for Changes Outcome Management System as a feedback tool in psychotherapy – A systematic review and meta-analysis. *Psychotherapy Research, 30,* 195-212.

Owen, J., Adelson, J., Budge, S., Wampold, B. E., Kopta, M., Minami, T., & Miller, S. D. (2015). Trajectories of change in short-term psychotherapy. *Journal of Clinical Psychology, 71,* 817-827.

Owen, J. Miller, S.D. Seidel, J. & Chow, D. (2016). The working alliance in treatment of

   military adolescents. *Journal of Consulting and Clinical Psychology, 84,* 200-210.

Prescott, D.S. (2017). Core competencies in Feedback-Informed Treatment. In D. S.

   Prescott, C. L. Maeschalck, & S. D. Miller (Eds.), *Feedback Informed Treatment*

   in Clinical Practice: Reaching for Excellence (pp. 37-52)*.* Washington D.C.

   American Psychological Association Press.

Prescott, D.S., Maeschalck, C.M., & Miller, S.D. (2017). *Feedback-Informed Treatment*

   in clinical practice: Reaching for Excellence. Washington, DC: American

   Psychological Association Press.

Rodgriguez, A., Terrones, L., Brookman-Frazee, L., Regan, J., Smith, A., & Lau, A.S.

   (2019, February 4). Associations Between Cultural Identity and Attitudes

   Toward Routine Progress Monitoring in a Sample of Ethnically Diverse

   Community Therapists. *Psychological Services.* Advance online publication.

   http://dx.doi.org/10.1037/ser0000327.

Schuckard, E., Miller, S. D. & Hubble, M. A. (2017) Feedback Informed Treatment:

   Historical and empirical implications. In D. S. Prescott, C. L. Maeschalck, & S.

   D. Miller (Eds.), *Feedback-Informed Treatment in clinical practice: Reaching for*

   Excellence (pp. 13-35). Washington DC: American Psychological Association

   Press.

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

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). The benefits of psychotherapy. Englewood Cliffs, NJ: Prentice-Hall.

Stulz, Lutz, Lucock, & Barkham, (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology, 75,* 864-74.

Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*(4), 547–559. http://doi.org/10.1037/a0028226

Valla, B. & Prescott, D.S. (2019). Beyond best practice: How mental health services can be better. London: Routledge.

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

Wampold, B. E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist, 62,* 857-873.

Weisz, J.R., Sandler, I.N., Durlak, J.A., & Anton, B.S. (2005). Promoting and protecting youth mental health through evidence-based prevention and treatment. *American Psychologist. 60(6),* 628-648.

Worthington, R.L., Mobley, M., Franks, R.P., & Tan, J.A. (2000). Multicultural

counseling competencies: Verbal content, counselor attributions, and social

desirability. *Journal of Counseling Psychology, 47*(4), 460–

468. https://doi.org/10.1037/0022-0167.47.4.460