16

THE IMPLICATIONS OF THERAPIST EFFECTS FOR ROUTINE PRACTICE, POLICY, AND TRAINING

JAMES F. BOSWELL, DAVID R. KRAUS, MICHAEL J. CONSTANTINO, MATTEO BUGATTI, AND LOUIS G. CASTONGUAY

Despite considerable research demonstrating the significance of the individual psychotherapist for mental health treatment outcomes (Baldwin & Imel, 2013), we have witnessed some researchers reacting to the findings on therapist effects with mild amusement, taking the position that these findings are interesting only insofar as they exemplify what can be done statistically with a large enough sample. Others have stated that the fact that there are differences between therapists is plainly obvious and not particularly interesting. Although admittedly anecdotal, the common thread between these viewpoints appears to be: "So what? I don't see how this is clinically useful."

We believe that efforts to address the implications of therapist effects are long overdue. In this chapter, we attempt to address the "so what?" question of therapist effects by discussing (a) what can and should (or should not) be done with process and outcome information collected in routine clinical practice, and (b) the challenges and potential solutions in conducting

http://dx.doi.org/10.1037/0000034-017
How and Why Are Some Therapists Better Than Others? Understanding Therapist Effects, L. G. Castonguay and C. E. Hill (Eds.)

relevant assessments and in using the resulting information. We hasten to say, however, that when addressing these issues, at least as many questions will be raised as tentative answers offered. Furthermore, we appreciate that the solutions offered should be considered tentative until stronger empirical evidence supports them.

WHAT CAN AND SHOULD BE DONE

The ultimate goal of psychotherapy researchers and practitioners should be to foster treatment success and avoid harm (Castonguay, Boswell, Constantino, Goldfried, & Hill, 2010). Within that context, once it is clear which treatment variables explain between-therapist effects (see Chapters 2 and 3, this volume, for a discussion of candidate variables for explaining therapist differences), the field should move toward integrating these variables within different facets of mental health care. This could involve adapting training programs, supervision routines, and standard referral practices to maximize therapist effectiveness and minimize client deterioration. This is especially true for those outcome variables for which therapists makes a clear and consistent contribution.

With the proper measurement and estimation of therapist contributions to treatment outcomes, provider effectiveness data have the potential to improve the quality of care. Multiple studies involving large samples have demonstrated that individual therapist effectiveness is relatively stable across time and clients (e.g., Kraus et al., 2016; Wampold & Brown, 2005; see also Chapter 2, this volume). Therapist effects, therefore, appear to be predictable. On the basis of real-world data, simulations suggest that treatment effect sizes would increase significantly by matching or referring clients to the most effective therapists. For example, Imel, Sheng, Baldwin, and Atkins (2015) conducted a Monte Carlo simulation to examine the impact of removing therapists with the worst outcomes, defined as performing in the bottom 5% of the sample. Extrapolating over 10 years, they found significantly higher response rates when the lowest performing therapists were removed and replaced with a random sample of therapists from the better performing population, translating into thousands of additional treatment responders over time. Although the data on effective therapists have so far been based almost exclusively on symptom measures (as opposed to other aspects of functioning, such as personality reorganization, that many insight-oriented clinicians perceive as an important focus of change), we view these empirical results as highly meaningful clinically, and as one very compelling answer to the "so what?" question. We also believe that this and other findings about therapist effects (see Chapters 1, 2, and 3) provide an empirical basis to our own question: What can be done with data gathered through routine outcome monitoring? In our view, one way that such an assessment can (and should) be used is to help identify the therapists to whom we most want to refer clients (not to dictate, but to provide relevant information). Such an assessment can also identify the therapists likely to be more effective with particular clients and the therapists who are unlikely to best serve particular clients (at least in the current phase of their career and/or without additional training and supervision).

The identification of stable outcome differences among providers and services is already being harnessed in other areas of health care. In medicine, treatment providers and hospital systems are often ranked on the basis of established criteria (http://www.leapfroggroup.org/), and this information is disseminated in various ways to stakeholders and decision makers (Scanlon, Lindrooth, & Christianson, 2008). Behavioral health care, however, has received virtually no attention in this area. It seems important to consider the feasibility and merit of using similar methods, as the predominant existing methods of referring clients to therapists are based on questionable assumptions, to say the least (e.g., the therapist's availability, reputation, or connection with a referring provider).

It should be recognized that strategies for making effectiveness data available for health care decisions are complex, with many options, but few good ones, at least partially because of the diversity of stakeholders (e.g., clients, therapists, hospitals, payers, trainers, state licensing boards). For example, although online provider ratings have become more common, they are typically based on self-selecting clients' general impressions that generate a certain number of endorsed "gold stars" with highly questionable validity (Boswell, Constantino, Kraus, Bugatti, & Oswald, 2016).

Each stakeholder group has its own unique needs and interests in the use of therapist effectiveness data, and these needs and concerns may not always align. Moreover, each stakeholder is confronted with several unresolved issues regarding use of such data. Although the discussion is far from exhaustive, we explore some of the needs of and problems faced by four critical stakeholder perspectives: payers (e.g., insurance companies), clients, therapists, and trainers. A consideration of these distinct needs and problems will help clarify not only what can (and should) be done with data monitoring but also what should not be done.

Payers

Payers are motivated to improve the efficiency and effectiveness of treatment. In the rapidly changing health care landscape, payers have been under increasing pressure from self-insured employers to tier their networks (Scanlon et al., 2008) and to display outcome "report cards" (Chernew, Gowrisankaran, & Scanlon, 2008). In a *tiered network*, health insurance plan providers or hospitals are separated into different levels on the basis of established cost and quality metrics. In such a system, individuals can choose to seek services from a lower tier provider or hospital, but they will be responsible for a higher copayment relative to seeking services from a Tier 1 provider or hospital. As alluded to previously, quality data from hospital systems are already being used to inform health care decision making at the client, employer, and insurance levels. One could argue that therapist effectiveness data could be similarly useful and valuable. To be of optimal value from the payer perspective, however, collection and access to these data would need to influence payment and payment models. A major movement is to pay out more to the providers who demonstrate consistently positive outcomes (Greene & Nash, 2008). Different labels are used for this approach, including pay-for-performance and performance-based incentives (Bremer, Scholle, Keyser, Knox Houtsinger, & Pincus, 2008).

This approach, as we discuss next, raises serious fears for providers. From a public health perspective, one could argue that a pay-for-performance approach on the basis of demonstrated change could disincentivize therapists from seeing more difficult clients. Clinical experience and research reveal subgroups of clients who demonstrate relatively flat change trajectories even after a significant dose of treatment. It may be the case that therapy offers a stabilizing function that reduces the risk of inpatient hospitalization or self-harming behaviors, rather than demonstrable improvements in symptoms and functioning on the basis of standard measurement tools. This raises an understandable concern regarding performance-based payment models. However, models could be constructed that adjust for client characteristics that are associated with attenuated response. In these models, the trajectory of clients with poorer prognosis would not be compared with the expected trajectory of motivated and less impaired clients, because these would not represent a relevant benchmark. Decreased frequency or absence of inpatient hospitalization could also function as a key performance indicator.

Nevertheless, a number of critical issues need to be addressed before performance-based payment models and incentives could be safely and fairly implemented in mental health practice, including the reliability and validity of outcomes measurement, and the selection of outcomes that are of greatest value. For example, should payers reimburse therapists for providing an established, evidence-based treatment with a high level of fidelity and/or for demonstrating significant functional improvements in individual clients on the basis of standardized self-report measures? There are no doubt inherent limitations to relying solely on self-report measures. To muddy the waters further, the results from Kraus, Castonguay, Boswell, Nordberg, and Hayes (2011; see also Kraus et al., 2016) indicate that a "better therapist" is relative to the client problems one is treating and an individual therapist should be considered *better matched* rather than generally "better" in an absolute sense. In addition, if a good therapist match exists, the referral system within the payer network would need access to the outcome data within whatever referral procedures the payer has in place (e.g., an online referral database with distance to travel and other selection criteria).

Finally, with regard to payers, one might wonder about the significance of an effect that appears to explain between 5% and 8% of the variance in client outcomes. The results of Imel et al. (2015) demonstrate that removing underperforming therapists translates into thousands of additional treatment responders, which would significantly improve public health. Other research (e.g., Kraus et al., 2016) demonstrates larger outcome effect sizes for therapists who are labeled *above average* on the basis of their performance track records. It is certainly the case that the importance of therapist differences should be weighed in the context of what we know about other treatment factors, such as client characteristics and interventions effects. In this respect, the size of the therapist effect is notable and should not be discounted. Furthermore, if a payer is in the hypothetical position of choosing between paying attention to therapist differences or mandating the use of empirically supported treatments (ESTs), it is considerably more cost-effective and logistically feasible to identify effective therapists on the basis of their observed performance on standardized measures than it would be to confirm the presence or absence of therapist EST certification and assess intervention-specific adherence and competence on an ongoing basis.

Clients

We have collected survey data demonstrating that clients struggle to find good-fitting therapists and highly value the idea of using therapist performance information when selecting therapists (Boswell et al., 2016). However, clients are more ambivalent about needing or wanting direct access to therapist effectiveness information. Although speculative, this may be because clients are less confident in their own ability to interpret or make use of therapist performance information. However, what is clear from the survey data is that an overwhelming majority of clients want assurances that those helping them find the right therapist are using such data, if the data are available.

If we can identify therapist effects in specific outcome domains, research highlights the outcome implications of steering clients to more effective therapists (e.g., Imel et al., 2015), and the preliminary evidence cited previously indicates that clients favor using such information. From the perspective of client treatment benefit and choice, we can and should use therapist effectiveness data to inform mental health care decisions. However, other research findings underscore the importance of client preferences and the relative values clients place on different treatment characteristics. For example, Swift and Callahan (2010) found that clients were willing to discount a significant percentage of benchmarked intervention empirical support for assurances that they could develop a positive therapeutic relationship with therapists. Although this study focused on attitudes toward ESTs, the implications logically extend to "empirically supported providers." The point here is that outcome data should not be viewed as the sole basis for referral or matching. Rather, the data should be viewed as one source of information to support these crucial decisions (as well as other important decisions that must be made once treatment has begun; see Castonguay, Barkham, Lutz, & McAleavey, 2013).

In addition, it is unclear which dissemination and implementation methods for provider effectiveness information are optimal. For example, just as there are multiple EST lists available in the literature, multiple versions of report cards could be established by a health plan, hospital system, or trade association (e.g., American Psychological Association). Furthermore, if there are relative effectiveness ratings by problem category (e.g., mood vs. substance use) or another domain classification system, how does a client know which one to prioritize, particularly when problem comorbidity is the rule?

Therapists

The term *accountability* has been used extensively in discussions about routine outcome monitoring (ROM) and therapist effectiveness, yet calls for increased accountability rarely originate from the therapist perspective. They are more likely to come from policy experts, payers, and clients. Understandably, therapists have concerns regarding outcome data, performance measurement, and report cards. Interestingly, in the same survey previously referenced (Boswell et al., 2016), therapists reported generally positive attitudes toward ROM. In fact, most therapists welcomed the idea of having a system that could help them "find" clients with whom they have a high likelihood of success. Relatedly, most therapists endorsed the belief that they are more or less effective with different types of clients.

Supporting the need for a system to assist therapists in identifying well-matched clients, research has shown that therapists tend to overestimate their own general effectiveness with clients, as well as their effectiveness relative to that of other therapists (e.g., Walfish, McAlister, O'Donnell, & Lambert, 2012). Traditional alternatives to ROM-based therapist effectiveness information include peer identification (or self-identification) as an *expert*, *master*, or a *specialist* in a particular area. It is not uncommon for such labels to be conferred on treatment developers. Even when a particular treatment has demonstrated efficacy in controlled research, it is taken for granted that the treatment developer, in particular, is an effective therapist. In the absence of systematically tracking the outcomes of this therapist's clients, this assertion is highly speculative. For example, it is unlikely that all of the inventors of surgical devices are outstanding surgeons, although it is most likely that they are surgeons.

In short, it has historically been common practice to make anecdotal claims about one's own or another therapist's effectiveness. This implies that therapist effectiveness information is valued at some level. It seems reasonable to expect that such assertions are based on the best available empirical evidence. The availability of such evidence is itself a critical issue. Depending on the clinical contexts that lead certain therapists to use routine outcome assessment, outcome databases may include a selective sample of providers and clients. For example, more seasoned therapists who see primarily self-paying clients may be less likely to routinely monitor their clients' progress and would not be represented in data repositories.

Regardless of one's professional stage or typical payment practices, we believe that therapists should be aware of their own relative areas of strength and weakness; however, if individual therapist effectiveness data are used in a way that is clearly detrimental to therapists' livelihoods, there will be no motivation for therapists to cooperate. For example, therapist effectiveness information could be published for the general public's consumption, mirroring one approach already being implemented in medicine (Henderson & Henderson, 2010). However, we caution against publishing individual therapist effectiveness data. Rather than focus on therapists' strengths, we believe this would lead to a greater focus on therapists' weaknesses. Although therapist effectiveness, as typically measured, is relatively stable across time, there is emerging research implying that therapists can achieve better outcomes when they engage in more deliberative practice activities. A highly innovative study by Chow et al. (2015) found that therapists' reported time spent engaging in deliberative practice related to their work with clients was significantly associated with their clients' outcomes. Interestingly, no significant associations were observed between specific activities (e.g., attending a training workshop) and outcome. However, the amount of reported cognitive effort exerted while reviewing therapist recordings, for example, was significantly correlated with outcome.

In another example, benchmarked outcome reports, which provide information regarding outcomes from a provider or setting relative to an existing standard or similar outcomes from a provider or setting, demonstrated that an adolescent inpatient substance abuse treatment facility was achieving suboptimal outcomes in the area of violence and anger (Adelman, McGee, Power, & Hanson, 2005). Consequently, the program sought additional training for addressing anger. When the targeted training was provided to program therapists, violence outcomes improved.

These findings highlight the importance of exercising caution when responding to observed differential effectiveness among therapists. If therapists can improve their outcomes in a particular domain through additional training and deliberative practice, then at least some who initially demonstrate relative ineffectiveness can become effective therapists. This is one of the reasons why the warning signals from data monitoring may be relevant for some therapists in the current phase of their career (in general or with regard to specific types of clinical problems). Arguably, ongoing postlicense training and supervision is typically unsystematic and underemphasized in the United States. Steering clients to better-matched therapists who have demonstrated effectiveness is only one side of the coin. Therapists who appear to struggle will require additional training and supervision resources.

We think publishing point-in-time therapist effectiveness data would ultimately deter therapist buy-in. Furthermore, we strongly hold that there should be equal public policy emphasis on helping low-performing therapists improve their skills.

Trainers

The studies by Chow et al. (2015) and Adelman et al. (2005) highlight the training implications of "harnessing" therapist effects and identifying effective therapists. In considering the training implications of therapist effects, we once again raise the question of "therapists are effective under what conditions?" For example, it may be the case that effective therapists possess a certain degree of basic interpersonal skills, such as empathic attunement. However, if therapists with very high empathic attunement do not have optimal skills to help clients control impulsive and dangerous behaviors (toward themselves and/or others), therapists may have poorer outcomes with certain types of clients (e.g., clients with problems with substance abuse or violence). We expect that among the population of all therapists, each individual may possess differential patterns of strengths and weaknesses across outcome domains (or clusters of domains) and across the various skills and attributes that are associated with good outcomes.

When coupled with accumulating research findings supporting the importance of deliberative practice and structured supervision for client outcomes, as well as changes in therapist behavior (e.g., Chow et al., 2015; Hill et al., 2015), we believe the utility of routinely evaluating client outcomes and treatment process to support professional development is no longer a matter of debate. Benchmarked outcome information can direct therapists, regardless of career stage, to engage in deliberative practice or seek additional consultation and training in targeted areas of relative ineffectiveness (Castonguay et al., 2010). For example, therapists who are relatively ineffective at improving sleep outcomes can be provided with specific training in evidence-based sleep interventions through training programs or when a workshop is offered at a professional conference.

The tracking of therapist outcomes has additional developmental implications. For example, after a short time in graduate training, therapists-in-training may have concrete data indicating that, compared with other therapists-in-training who treat similar clients, they are less effective with clients presenting with substance abuse issues. This information could motivate particular therapists-in-training to seek additional supervised training experiences with clients experiencing substance abuse and identify substance abuse treatment as an area in need of additional training in predoctoral internship applications. Without improvement, these therapists-in-training might steer away from treating clients with substance abuse altogether.

MAKING USE OF WHAT WE KNOW AND CAN ASSESS: CHALLENGES AND SOLUTIONS

We have focused on the challenges in identifying effective therapists and making use of therapist effectiveness information. In this section, we refocus on some of the big challenges to integrating attention to therapist effects in routine practice, training, and policy, as well as offer potential solutions on the basis of existing research, trends in the field, and our own experience.

Engaging Therapists

Therapists are skeptical of how these data might be used, and they will have little incentive to engage in routine outcome or process assessment if the potential (real and perceived) costs are too high. From a Skinnerian perspective, should highly effective therapists be rewarded (e.g., with increased reimbursement or an enhanced referral stream), or should ineffective therapists (e.g., the 4% with no effective outcome domain in Kraus et al., 2011) be punished (e.g., removed from a network's preferred provider list until they document improvement)? Consistent with Skinner's view on the differential impact of various kinds of contingencies, we believe that a focus on positive reinforcement (see Gates et al., 2005) and therapists' relative strengths will yield greater engagement. Even if therapists are temporarily restricted from treating clients with a certain type of problem (e.g., depression), they could still treat other types of problems (e.g., anxiety) while working on increasing their competency in treating depression. This underscores that therapists are not globally effective or ineffective but are effective or ineffective under certain conditions. This allows positive reinforcement of effective domains and provides markers for additional training in other domains (a frame that we would hope would foster therapist engagement in outcome monitoring and use). However, it will take time and a critical mass of "corrective experiences" related to positive impacts on therapists' practice and clients' outcomes before the collection and use of effectiveness data becomes standard practice.

Socialization in Training

When asked about the factors that shape their professional identity and approach to psychotherapy, therapists often identify experiences during graduate training (e.g., supervision; Goldfried, 2001). Therapy practice and supervision during training creates a schema for conducting therapy, including thinking about client problems, how clients change, and the roles of therapist and client. Simply put, if therapists-in-training are not exposed to integrating routine assessment and data collection in their work with clients or their own development, they will be less likely to integrate routine assessment and data collection as professionals or to be receptive to actuarial feedback. Although we are speculating here, therapists without such exposure may ultimately harbor negative or suspicious views toward such practices.

The earlier that therapist-level effectiveness data and feedback (outcome and process) are introduced in therapists' training, the more likely it is that such assessment will become a part of therapists' routine clinical practice. In line with the recommendations of Barkham, Lutz, Lambert, and Saxon (Chapter 1, this volume), a potential solution is the early introduction of data-driven self-reflection and supervision in graduate training, which is transtheoretical and transdiagnostic (Constantino, Boswell, Bernecker, & Castonguay, 2013). Early socialization to this level of feedback may not only enhance the impact of training but also foster positive attitudes toward measurement-based care. Training directors and supervisors are likely to see benefit in making such routine assessment a standard part of the training curriculum. At the very least, it provides an opportunity for ongoing program evaluation. This would include making review of information from routine monitoring (process and/or outcome) an explicit part of supervision. A supervisor's valuing of such information and feedback might serve as a model for enhancing treatment responsiveness through discussions of what to do, and what not to do, in response to the obtained feedback.

That said, the use of relative effectiveness information to direct deliberative practice experiences is not limited to therapists' graduate training.

Rather, we envision this information being directly linked to continuing education experiences as a professional. In fact, reflective practice/self-assessment is a competency benchmark according to the American Psychological Association (Kaslow et al., 2009). The validity of this approach for training and professional development is dependent on the use of performance benchmarks that control for client characteristics that are systematically related to outcome yet are not in therapists' control (i.e., risk-adjustment, baseline severity).

System and Stakeholder Integration

It is a complex task to weigh all of the potential provider selection options, which include such things as the supply of local therapists; the relative skills, strengths, and weaknesses of therapists; the unique needs and problems of each client; and the current capacity of each therapist (e.g., waiting lists), among other factors. Effective communication is vital, and ultimately, information must make its way to the relevant decision makers. We do not see how direct-to-consumer or even direct referral source access to raw therapist effectiveness data will be the most efficient way to improve population outcomes. Therefore, a solution will require some type of "expert system," either the use of a computerized expert system and/or a health care expert (e.g., a primary care physician) that could help clients interpret the data.

We envision the following use of therapist effectiveness data that meet the various anticipated needs of key stakeholders. Clients desiring a scientifically based referral to a well-matched therapist would take their first multidimensional outcome assessment, not just before the start of treatment but also before a referral is made. For a system-wide application across an entire health plan using the same outcome system, this could work on the payer's website where the payer lists and helps clients find specialists, where the system is connected to a primary care physician's office, and where the client–therapist match could be completed online before the client's visit to the physician's office. For more localized applications, the system could sit within a group practice, community mental health system, or hospital-based practice where clients are to be matched to therapists within the localized practice setting.

Results of this assessment would be scored by a central processing system and be augmented with a list of well-matched therapists. This list could be geographically boundless (in the case of telemedicine) or as restricted as the client requires (e.g., within a few blocks of a public transportation stop). It would also include all of the typical filters for things such as type of insurance accepted, age of the therapist, and therapist degree/orientation. The feedback could encourage the user to relax their criteria so that at least three options are delivered, allowing for choice. Payers could receive a signal that a therapist (or a short list of therapists) was "well matched," triggering a higher rate of reimbursement suggested for top-tier network providers. Therapists would be incentivized to participate because their strengths would be rewarded with higher pay and they would automatically receive referrals that are in their "wheelhouse." With funding from the Patient-Centered Outcomes Research Institute, we are conducting a mixed-method study, including a randomized controlled trial, to examine the feasibility and impact of this approach. To our knowledge, this will be the first randomized controlled trial for a referral process in any field of medicine and may shed light on how health care can be improved by creating real-world applications of therapist effect data. This study will be based in the largest primary care outpatient practice in New England, but the same methodology should be applicable to all of the settings described previously.

Policymakers

Finally, keeping decision makers informed so that enacted policies are empirically based and result in sufficient benefit to relevant stakeholders represents a core challenge. For example, if clients do not value therapist performance information or problem-domain–based matching, then they will not be motivated to access an "expert system." We simply do not know if relative to other therapist characteristics (e.g., experience), clients would prefer to be assigned or referred to therapists on the basis of therapist effectiveness. This should be investigated, as should the possibility that the quality of the therapeutic relationship might overcome initial apprehensions that clients may have about the way they were referred to therapists.

Another key example comes from the adherence–competence literature. A high level of adherence and competence in applying an evidence-based treatment does not guarantee a better client outcome (Webb, DeRubeis, & Barber, 2010), yet state licensing boards appear to be more focused on therapist fidelity than therapist effectiveness. These are far from mutually exclusive and may often be positively correlated; however, policymakers must be educated on the complexities of behavioral health treatment outcomes. This will require input from not only researchers but also therapists and clients. Consequently, one solution is to support the active participation of diverse stakeholders in this ongoing discussion. In addition, implementation interventions and policy initiatives can themselves be a focus of research (Boswell, Kraus, Miller, & Lambert, 2015).

CONCLUSION

In answer to the "so what?" question posed throughout this chapter, we believe that therapist effect data may have an important place in transforming and improving the health care system. We believe that the positive implications of therapist effects for routine practice, training, and policy will be maximized by taking a multidimensional and pantheoretical approach to measurement that involves diverse outcome and process domains. For example, future work should prioritize the identification and measurement of processes and outcomes that are more directly linked with psychodynamic and humanistic therapies, such as enhanced reflective functioning and self-actualization. Although data monitoring should never by viewed as the only basis for mental health care decisions, we believe it can provide information to improve referral practice, training programs, and supervision routines, as well as to maximize and enhance therapist effectiveness. Furthermore, we believe that decision making and policy initiatives that are informed by observed therapist effects should focus on therapists' relative strengths and disseminating easily interpretable information to relevant stakeholder(s). Conversely, observed relative ineffectiveness should trigger targeted training and consultation resources to providers and care systems.

Of course, all of this is easier said than done. We end this chapter after having raised more questions than answers. Considerable work will be required to realize the potential implications of therapist effects. For example, stakeholders must carefully examine the relative value of different outcome and process domains where the identification of therapist differences are most meaningful. Symptom-based outcomes may be more relevant in certain contexts or psychotherapy approaches, whereas other psychological constructs (e.g., self-actualization) may be more important in others. In addition, research is sorely needed on the client-level outcomes of performance-based incentives for therapists, or, conversely, the use of penalties for therapists for lack of demonstrated effectiveness. In our view, the best way to face these challenges and find actionable and retainable solutions that may improve health care is to pursue them through active partnerships between researchers, therapists, clients, and policymakers.

REFERENCES

Adelman, R., McGee, P., Power, R., & Hanson, C. (2005). Reducing adolescent clients' anger in a residential substance abuse treatment facility. *Joint Commission Journal on Quality and Patient Safety*, *31*, 325–327.

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. 258–297). Hoboken, NJ: John Wiley & Sons.

Boswell, J. F., Constantino, M. J., Kraus, D. R., Bugatti, M., & Oswald, J. (2016). The expanding relevance of routinely collected outcome data for mental health care decision making. *Administration and Policy in Mental Health*, *43*, 482–491.

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

Bremer, R. W., Scholle, S. H., Keyser, D., Knox Houtsinger, J. V., & Pincus, H. A. (2008). Pay for performance in behavioral health. *Psychiatric Services*, *59*, 1419–1429. http://dx.doi.org/10.1176/ps.2008.59.12.1419

Castonguay, L. G., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 85–133). Hoboken, NJ: John Wiley & Sons.

Castonguay, L. G., Boswell, J. F., Constantino, M. J., Goldfried, M. R., & Hill, C. E. (2010). Training implications of harmful effects of psychological treatments. *American Psychologist*, *65*, 34–49. http://dx.doi.org/10.1037/a0017330

Chernew, M., Gowrisankaran, G., & Scanlon, D. P. (2008). Learning and the value of information: Evidence from health plan report cards. *Journal of Econometrics*, *144*, 156–174. http://dx.doi.org/10.1016/j.jeconom.2008.01.001

Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy*, *52*, 337–345. http://dx.doi.org/10.1037/pst0000015

Constantino, M. J., Boswell, J. F., Bernecker, S. L., & Castonguay, L. G. (2013). Context-responsive integration as a framework for unified psychotherapy and clinical science: Conceptual and empirical considerations. *Journal of Unified Psychotherapy and Clinical Science*, *2*, 1–20.

Gates, L. B., Klein, S. W., Akabas, S. H., Myers, R., Schawager, M., & Kaelin-Kee, J. (2005). Outcomes-based funding for vocational services and employment of people with mental health conditions. *Psychiatric Services*, *56*, 1429–1435. http://dx.doi.org/10.1176/appi.ps.56.11.1429

Goldfried, M. R. (Ed.). (2001). *How therapists change: Personal and professional reflections*. Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10392-000

Greene, S. E., & Nash, D. B. (2008). Pay for performance: An overview of the literature. *American Journal of Medical Quality*, *24*, 140–163. http://dx.doi.org/10.1177/1062860608326517

Henderson, A., & Henderson, S. (2010). Provision of a surgeon's performance data for people considering elective surgery. *Cochrane Database of Systematic Reviews*, *11*, CD006327. Advance online publication.

Hill, C. E., Baumann, E., Shafran, N., Gupta, S., Morrison, A., Rojas, A. E., . . . Gelso, C. J. (2015). Is training effective? A study of counseling psychology doctoral trainees in a psychodynamic/interpersonal training clinic. *Journal of Counseling Psychology*, *62*, 184–201. http://dx.doi.org/10.1037/cou0000053

Imel, Z. E., Sheng, E., Baldwin, S. A., & Atkins, D. C. (2015). Removing very low-performing therapists: A simulation of performance-based retention in psychotherapy. *Psychotherapy*, *52*, 329–336. http://dx.doi.org/10.1037/pst0000023

Kaslow, N. J., Grus, C. L., Campbell, L. F., Fouad, N. A., Hatcher, R. L., & Rodolfa, E. R. (2009). Competency assessment toolkit for professional psychology. *Training and Education in Professional Psychology*, *3*, S27–S45. http://dx.doi.org/10.1037/a0015833

Kraus, D. R., Bentley, J. H., Alexander, P. C., Boswell, J. F., Constantino, M. J., Baxter, E. E., & Castonguay, L. G. (2016). Predicting therapist effectiveness from their own practice-based evidence. *Journal of Consulting and Clinical Psychology*, *84*, 473–483. http://dx.doi.org/10.1037/ccp0000083

Kraus, D. R., Castonguay, L., Boswell, J. F., Nordberg, S. S., & Hayes, J. A. (2011). Therapist effectiveness: Implications for accountability and patient care. *Psychotherapy Research*, *21*, 267–276. http://dx.doi.org/10.1080/10503307.2011.563249

Scanlon, D. P., Lindrooth, R. C., & Christianson, J. B. (2008). Steering patients to safer hospitals? The effect of a tiered hospital network on hospital admissions. *Health Services Research*, *43*, 1849–1868. http://dx.doi.org/10.1111/j.1475-6773.2008.00889.x

Swift, J. K., & Callahan, J. L. (2010). A comparison of client preferences for intervention empirical support versus common therapy variables. *Journal of Clinical Psychology*, 66, 1217–1231. http://dx.doi.org/10.1002/jclp.20720

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

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 and Clinical Psychology*, *73*, 914–923. http://dx.doi.org/10.1037/0022-006X.73.5.914

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *78*, 200–211. http://dx.doi.org/10.1037/a0018912