Professional Psychology: Research and Practice
2007, Vol. 38, No. 5, 493–500

0735-7028/07/$12.00 DOI: 10.1037/0735-7028.38.5.493

Assessing Intervention Competence and Its Relation to Therapy Outcome: A Selected Review Derived From the Outcome Literature

Jacques P. Barber, Brian A. Sharpless, Susan Klostermann, and Kevin S. McCarthy
University of Pennsylvania School of Medicine

The assessment of intervention competence possesses an obvious relevance for practitioners and clinical scientists alike. It is often assessed as part of the evaluation of treatment integrity in clinical research in general, and in randomized clinical trials (RCTs) in particular. The authors first attempt to add clarity to the concept and better differentiate intervention competence from closely related constructs. Next, the authors review and evaluate the main measures of therapist competence used in RCTs, relying on this conceptual foundation to provide suggestions for future measures. The empirical literature on the relation between therapist competence and clinical outcome is then reviewed. The relation, while positive, is weaker than expected, and factors having a potential bearing on this are discussed. The authors then recommend that new measures be created and that the assessment of limited-domain competence be supplemented by explorations of global competence. Due to the potential ramifications for the field, the authors also recommend that caution be exercised in the task of operationally defining competence.

Keywords: therapist competence, competency, therapy outcome, treatment integrity, process of change

Competencies of various kinds have permeated our lives and have garnered popular and professional attention. This is likely due to factors such as the increased degree of specialization in various disciplines in conjunction with societal demands for accountability. In professional psychology in particular, we have witnessed a burgeoning interest in competency-based education and practice termed the *competencies-based movement* (Kaslow, 2004). Kaslow (2004) described eight domains of psychologist competencies (e.g., ethical, assessment) warranting attention. This paper will focus on one of them, namely intervention competencies, and what we have learned about competence assessment from randomized clinical trials (RCTs). Despite the narrow focus on RCTs, we will address four questions holding practical implications for professional psychologists. First, what is intervention competence, and can conceptual clarity be added to this concept? Second, how is intervention competence typically measured in RCTs, and can improvements be made in its measurement? Third, what is the relation between competence and therapy outcome? And finally, what are the implications of this current state of knowledge for our field? Before addressing these issues, however, some contextual discussion of intervention competency is warranted.

Intervention competence has a number of ethical and practical implications relevant to professional psychologists. For instance, the American Psychological Association (APA) recently released its *Policy Statement on Evidence-Based Practice in Psychology* (APA, 2005), which directly mentions competence, or "clinical expertise" (p. 1). More fundamentally, the current ethics code of the APA (2002) provides four relevant points for intervention competence: (a) Psychologists should only provide services within areas of competence; (b) in areas that lack clear guidelines, psychologists should take steps to ensure the competence of their work and protect their patients; (c) psychologists should increasingly develop and maintain their competence; and (d) psychologists should base their work upon established scientific and pro-

JACQUES P. BARBER received his PhD in clinical psychology from the University of Pennsylvania. He is a professor of psychology in psychiatry at the University of Pennsylvania's Center for Psychotherapy Research. His research interests include the study of the efficacy of different forms of psychotherapy for patients with various disorders. In addition, he is interested in the mechanisms of change involved in those interventions and focuses his research on examining the impact of both the therapeutic relationship and the specific techniques used in therapy on patients' outcome.

BRIAN A. SHARPLESS received his PhD in clinical psychology and MA in philosophy from Pennsylvania State University. He is currently a project director at the University of Pennsylvania's Center for Psychotherapy Research. His research interests include anxiety, psychotherapy research, the philosophical foundations of psychology, and psychoanalytic psychotherapy.

SUSAN KLOSTERMANN received her BA in psychology from Swarthmore College. She was a research assistant at the University of Pennsylvania's Center for Psychotherapy Research and is now working at Western Psychiatric Institute and Clinic, Pittsburgh, Pennsylvania. Her research interests include understanding mood and anxiety disorders in children and adolescents.

KEVIN S. MCCARTHY received his MA in clinical psychology from the University of Pennsylvania. He is currently a PhD candidate in the Department of Psychology at the University of Pennsylvania and works at the Center for Psychotherapy Research. His research interests are perceptions of interventions in therapy, psychotherapy process, and interpersonal relationships.

THIS ARTICLE WAS WRITTEN with support from the National Institute of Mental Health, Grants MH 061410 and MH 070664.

CORRESPONDENCE CONCERNING THIS ARTICLE should be addressed to Jacques P. Barber, Suite 648, Center for Psychotherapy Research, Department of Psychiatry, University of Pennsylvania School of Medicine, 3535 Market Street, Philadelphia, PA 19104-3309. E-mail: barberj@mail.med.upenn.edu

fessional knowledge. The implications of this formulation are legion. For instance, as APA does not support practice outside of one's areas of competence, it would probably be very difficult to find psychologists who could document adequate competence in more than one or two specific therapies (of course, the definitions of *specific therapies* and *adequate competence* are a matter of debate). Moreover, there would be few trainers available who could document appropriate levels of competence.

Although it could be argued that the APA guidelines essentially beg the question of the nature of competence, this opacity may, in reality, be representative of the state of the field. The tacit sense of competence that is promulgated may serve the purpose of avoiding premature intellectual foreclosure on the concept of competence, as operationalizing competence could have significant ramifications for the field as a whole as well as third-party payment.

What Is Intervention Competence?

Competent delivery of a treatment involves a series of steps, such as the therapist knowing the treatment, having treatment capability (viz., that the therapist is actually capable of delivering the treatment), having a willingness to deliver the treatment as intended, and being able to carry it out with a particular patient (i.e., the patient's willingness and readiness). We will not address these issues in this paper. However, because our discussion of the nature of competence is predicated on the belief that competence can be conceptually (and practically) distinguished from other related constructs such as treatment integrity and adherence, these will be discussed.

Treatment Integrity

Treatment integrity typically refers to the extent to which a therapeutic intervention was implemented as it was intended (and competently delivered) as well as its ability to be differentiated from other approaches (e.g., Perepletchikova & Kazdin, 2005). It is of primary importance in clinical trials, as interventions must be successfully implemented before any meaningful determinations of outcome can be made. Manualizing therapy along critical dimensions of the intended treatment is a necessary prerequisite for ensuring that the treatment is implemented as intended. Another prerequisite involves the training of therapists and their adherence to a manual.

Adherence

The meaning of adherence within the context of a manualized treatment is the extent to which a therapist utilizes procedures specified by the treatment manual while minimizing recourse to extramanual components. This is often specified in terms of intervention frequency but can also be thought of as the degree to which a treatment has been actually delivered (e.g., Barber, Liese, & Abrams, 2003). Adherence is a prerequisite for and the sine qua non of competence, as it is not possible to competently deliver a specified treatment sans adherence to that treatment (Perepletchikova & Kazdin, 2005). Therefore, competence requires conforming to some treatment guidelines, however they are defined. Adherence is conceptualized as including treatment behaviors that are (a) unique and essential, (b) essential but not unique, (c) neither unique nor essential (but not proscribed), and (d) proscribed (Waltz, Addis, Koerner, & Jacobson, 1993). High levels of the first two components (even the first three) would typify good adherence.

Competence

Definitions of intervention competence are not self-evident, and the range of definitions available in the literature instantiates this point. We have found a modified version of Epstein and Hundert's (2002) definition to be helpful. Namely, competence can be thought of as the judicious application of communication, knowledge, technical skills, clinical reasoning, emotions, values, and contextual understanding for the benefit of the individual and community being served. Further, we suggest that the crucial distinction between adherence and competence is the type of knowledge that each demonstrates. Adherence demonstrates that one knows "how" to intervene and can actually do so. In a sense, adherence is context-independent knowledge. Competence, on the other hand, is always context dependent and requires a knowledge of "when" and "when not" to intervene. It is predicated on a (possibly tacit) sense of appropriateness, responsiveness, good judgment, and clinical acumen. Nested within the contextual nature of competence is also an appreciation of and comfort with issues of diversity at the surface (e.g., diversity of gender, ethnicity, or pathology) as well as deeper senses of the term (e.g., diversity of values or knowledge).

Two meanings of "competence." Even if competence can be disentangled from adherence at a conceptual level, competence within a psychotherapy context possesses at least two meanings. We term them *global competency* and *limited-domain competency*. Global competency is the broad idea that a therapist possesses clinical acumen and that competence pervades the therapist's interventions. It is the sense that a therapist appropriately and independently manages a number of clinical problems and can adequately help patients realize their treatment goals. This is the type of competence presumably desired by those who determine internship and entry-level practice readiness. Although this is an important meaning of competence worthy of exploration in its own right, most (if not all) of the empirical investigations of competence have assessed it within a limited-domain context using experienced therapists. Limited-domain competence is a subset of global competence that is solely expressed within the context of a specific psychotherapy intervention or treatment modality. Limited-domain competency may be the desired competency assessed for a specialty certification. Needless to say, global and limited-domain competencies do not necessarily or perfectly overlap.

Within the limited-domain context, skillful application of the unique (treatment-specific) and general (Binder, 2004) techniques of psychotherapy have been the focus of attention, with the former receiving the majority of attention in competency measures. But regardless of technique specificity, competent application requires that techniques be wisely applied while the idiosyncratic context of the patient is simultaneously considered (e.g., Stiles, Honos-Webb, & Surko, 1998). Thus, competent therapists are flexible in their judicious implementation and nonimplementation of therapeutic techniques (Binder, 2004; Borkovec & Sharpless, 2004). Functioning in this competent manner requires a great deal of judgment and clinical acumen. This is especially the case given the paucity of

data available to suggest which specific moment-to-moment interventions are associated with improvement for a specific patient within a particular therapeutic interaction.

Levels of competence. To further complicate matters, there clearly exists a continuum of competence, yet thresholds on this continuum for specific professional functions are not clearly forthcoming and unambiguous. For instance, identifying the appropriate threshold of competence for a psychology intern would presumably be different from the acceptable threshold of competence required for a seasoned psychologist serving as a protocol therapist in an RCT. The concept of competence may have a different meaning in RCTs when compared to general practice. Competence displayed in an RCT may be an expression of *freedom within restraint* due to the necessity for strong experimental controls and the limited-domain context. The example can be made even more complex by introducing Beutler's (1999) discussion of manualizing flexibility, or the attempt to train therapists to flexibly and appropriately use manualized techniques. In a general practice setting, psychologists have more freedom to be flexibly oriented to the particularity of an individual's clinical problem. Therefore, assessment of competence with these different purposes in mind may be separate yet overlapping tasks. In fact, the current task of an APA committee (Assessment of Competency Benchmarks Work Group [ACBWG], 2007) is to formulate "benchmarks" for four levels of professional development (practicum, internship, entry-level practice, and advanced training readiness). Our work and review focuses primarily on the fourth developmental level (*readiness for advanced training*), for which no benchmarks have yet been formulated. However, our findings may be applicable not only to the other levels discussed by ACBWG but also to an expert level not addressed by this APA committee.

In summary, defining therapists' competence is no easy task. In fact, we have serious doubts that the field, using the measures reviewed, is ready to meet the challenges raised by different bodies asking for evaluations of therapists' competency. Be this as it may, there are continued pressures to come up with such requirements.

How Has Competence Been Measured?

By their actions, many researchers seem to assume that the way to assess competent delivery of a treatment is to examine whether specific techniques are competently delivered. Researchers have focused less on general skills and global competency perhaps due to the need to assess competency for specific interventions (e.g., cognitive behavioral therapy for depression); the belief that these specific interventions are responsible for the change observed during treatments; and the difficulties involved in trying to conceptualize, define, and measure global competencies. Table 1 displays a number of the measures of competence that have been developed and for which psychometric data are available. This list is not exhaustive, but these particular measures were chosen because they each emphasize therapists' in-session performance (as opposed to outcome-based criteria or standard milestones such as American Board of Professional Psychology certification), represent a variety of therapeutic approaches, and can be used by independent observers. In addition to reviewing these measures, we will provide recommendations for future test construction and evaluate whether these measures are suitable for many of the field's purposes.

Comparative Analysis of Limited-Domain Competency Measures

The measures described in Table 1 differ widely in structure. For instance, the number of items vary from 1 (Multicenter Collaborative Study for the Treatment of Panic Disorder Global Competence Item [MCSTPD-GCI], Huppert et al., 2001) to 82 (Adherence/Competence Scale for Supportive-Expressive Psychotherapy for Cocaine Dependence [ACS-SEC], Barber, Krakauer, Calvo, Badgio, & Faude, 1997). This likely reflects the different approaches to defining what a competent therapist would (or could) do in session and is evidence of the measures' heterogeneity. Several (e.g., Yale Adherence/Competence Scale [YACS], Carroll et al., 2000; Contextual Assessment of Therapist Adherence/Competence [CATAC], Hogue et al., 2007; Penn Adherence/Competence Scale for Supportive-Expressive Psychotherapy [PACS-SET], Barber & Crits-Christoph, 1996) also possess a modular structure allowing for the assessment of competence in specific (e.g., supportive techniques, 12-step facilitation) and more general domains. Whereas overall ratings are useful for certain situations, this modular approach combines flexibility and specificity and may be worth pursuing in subsequent measures. According to the definitions presented above, all measures assessed competence to an extent. However, some (e.g., Short-Term Anxiety Provoking Psychotherapy Rating Form [STCRF], Svartberg, 1989; Cognitive Therapy Scale [CTS], Young & Beck, 1980) seem to assess competence that is linked with elements of adherence. Others (e.g., PACS-SET, ACS-SEC, CATAC) largely rate adherence and competence separately, and one (Cognitive Therapy Adherence/Competence Scale [CTACS], Barber et al., 2003) measures adherence and the more specific competence domains of *appropriateness* and *quality* of the interventions. Further, all of the reviewed measures assess limited-domain competency. Interestingly, while the measures tend to focus on specific aspects of limited-domain competence, many include a single-item estimate of overall session competence (e.g., CTACS, PACS-SET, CATAC). However, none of them present the correlations between the two.

Psychotherapy researchers have argued that experts are needed in order to decide that an intervention was delivered competently (e.g., Barber & Crits-Christoph, 1996). Therefore, measures such as these require an expert in the specific modality that is assessed in order to render judgments of competence (for an exception, see Paivio et al., 2004). We are not aware of empirical data supporting this point. Further, the specific criteria on how experts were defined (and even how expertise was evaluated) were not always made explicit, and we recommend including this information in future studies

Data for the reliability of the measures were also reviewed. Internal consistency coefficients for each measure were generally high (see Table 1), indicating that if a therapist was judged as highly competent on one intervention, he or she was likely to be judged highly on related interventions. Interrater reliability estimates for these measures tended to be low, especially when fewer judges' ratings were averaged in the estimate (e.g., Vallis, Shaw, & Dobson, 1986). Poor interrater reliability might be due to four problems: (a) different judges' understandings of competence, (b) differences in how much attention judges pay to different aspects of the treatment delivery (e.g., implicitly weighting certain inter-

Table 1
Reliability and Validity of Limited-Domain Competence Measures

Competence measure (author, year; no. of items)	Therapy modality	Reliability		Validity
		Internal[a]	Interrater[b]	
Unimodal				
CTS (Young & Beck, 1980; 11)	CBT	.95	.74	Scale discriminates between acceptable vs. nonacceptable sessions (Vallis et al., 1986) and predicts personality disorder improvement (Hoffart et al., 2005[c,d]) and depression improvement (Kingdon et al., 1996[c,d]; Shaw et al., 1999[c]; Trepka et al., 2004[d]) but not patient-rated global symptom improvement (Shaw et al., 1999[c]).
CTACS (Barber et al., 2003; 25)	CBT	.93	.80	CBT sessions scored higher than dynamic or drug-counseling sessions (Barber et al., 2003).
MACT-RS (Davidson et al., 2004; 11)	CBT		.66[e]	Scale predicted global symptom improvement (Davidson et al., 2004[c,d]).
MCSTPD-GCI (Huppert et al., 2001; 1)	CBT			Moderate CBT competence, but not high or low, predicted outcome on multiple measures (Huppert et al., 2001).
PACS-SET (Barber & Crits-Christoph, 1996; 45)	Dynamic	.95	.73	SE sessions scored higher than CBT sessions; expert therapists scored higher than less experienced therapists (Barber & Crits-Christoph, 1996); scale predicted depression improvement (Barber, Crits-Christoph, & Luborsky, 1996).
ACS-SEC (Barber et al., 1997; 82)	Dynamic	.97	.43	SE sessions scored higher than CBT or IDC sessions (Barber et al., 1997); scale predicted less drug improvement (Barber et al., in press).
STCRF (Svartberg, 1989; 11)	Dynamic	.71	.70	Correlated with therapist-rated competence, STAPP sessions scored higher than nondirective sessions (Svartberg, 1989); scale predicted less global improvement (Svartberg & Stiles, 1994[c]).
ACS-IDC for Cocaine Dependence (Barber, Mercer, et al., 1996; 43)	IDC	.84[g]	.93	IDC sessions scored higher than SE or CBT sessions (Barber, Mercer, et al., 1996); competence did not predict outcome (Barber et al., 2006[c]).
TFS (Paivio et al., 2004; 1)	EFT		.87[h]	Scale did not predict symptomatic improvement (Paivio et al., 2004[c,i]).
TSRF (Chevron & Rounsaville, 1983; 14)	IPT		.88[f]	Scale correlated with four IPT competence measures; scale predicted patient-rated improvement (O'Malley et al., 1988[c]).
Multimodal				
CATAC (Hogue et al., 2007; 9)	CBT, FT		.38[g]	
YACS (Carroll et al., 2000; 55)	CBT, CM, IDC		.87[g]	

Note. CTS = Cognitive Therapy Scale; CBT = cognitive behavioral therapy; CTACS = Cognitive Therapy Adherence/Competence Scale; MACT-RS = Manual-Assisted Cognitive Behavior Therapy Rating Scale; MCSTPD-GCI = Multicenter Collaborative Study for the Treatment of Panic Disorder Global Competence Item; PACS-SET = Penn Adherence/Competence Scale for Supportive-Expressive Psychotherapy; SE = supportive-expressive psychotherapy; ACS-SEC = Adherence-Competence Scale for Supportive-Expressive Psychotherapy for Cocaine Dependence; IDC = individual drug counseling; STCRF = Short-Term Anxiety Provoking Psychotherapy Rating Form; STAPP = short-term anxiety provoking psychotherapy; ACS-IDC = Adherence/Competence Scale for Independent Drug Counseling; TFS = Therapist Facilitating Scale; EFT = emotion-focused therapy; TSRF = Therapist Strategy Rating Form; IPT = interpersonal therapy; CATAC = Contextual Assessment of Therapist Adherence/Competence; FT = family therapy; YACS = Yale Adherence/Competence Scale; CM = clinical management.
[a] All values are Cronbach's α for total scale score unless otherwise noted. [b] All values are intraclass-correlation coefficients for total scale score and have been recalculated for two judges unless otherwise noted. [c] Therapists other than PhD-level psychologists were included. [d] Outcome determined post-termination. [e] Number of judges required to obtain this statistic not reported. [f] Pearson product moment coefficient. [g] Median of subscales. [h] A *k* statistic was used for early sessions. [i] Competence was judged by MA-level psychologists.

ventions as more important to competence), (c) difficulties in operationalizing competence, and (d) the use of uniformly competent therapists in RCTs (viz., therapists are selected for their competence, leading to very small variance in competence ratings and unstable statistical estimates of reliability). To increase the likelihood of obtaining reliable ratings, we recommend (a) aggregating multiple scores of multiple judges; (b) training judges in the use of competence measures to better standardize their definitions of competence and their interpretations of specific items; and (c) training them across sites, as systemic differences can arise between groups of raters who receive different training. The CTS, the most widely used measure of competence in RCTs, was evaluated in this manner, and it was found to be unreliable (Jacobson, 1998). However, when trained together, CTS raters can be reliable.

Some evidence for the validity of these measures is presented as well. High internal consistency estimates give some support for criterion validity and suggest that theoretically similar interventions are being delivered together (e.g., therapists are competently administering nearly all cognitive therapy interventions together in a session). Evidence for criterion validity is also found in the ability of many of the measures to discriminate among treatments based on the limited-domain competence levels of the therapists in each treatment (see e.g., Barber & Crits-Christoph, 1996, in which dynamic therapists were rated higher on a dynamic psychotherapy

competence scale than were cognitive therapists). Several factors might limit the validity of these measures and to our knowledge have not been explored in the literature. First, while the measures were used with many different patient populations (e.g., personality disorder, depression, cocaine dependence), no explicit test for differences in the application of the measures has been conducted. Barber and colleagues (Barber & Crits-Christoph, 1996; Barber et al., 1997) in fact developed separate measures for the same treatment for different disorders. A second limitation is the scant attention paid to issues of patient and therapist diversity. It would be important for new measures to explicitly assess competence in tailoring interventions to a patient's particular cultural background just as other measures now assess therapists' awareness and acceptance of multicultural differences. A third limitation is that these scales were developed for a specific treatment manual, and often for a specific disorder. Thus, it is unclear how these scales could be used for evaluating therapists' competence in using similar but not identical treatment (e.g., can the CTS be used to evaluate cognitive therapy for anxiety disorders or a personality disorder).

Finally, there are also a number of interesting innovations that appear promising and certainly warrant further exploration as competence assessment becomes increasingly sophisticated. As one example, the Manual-Assisted Cognitive Behavior Therapy Rating Scale (MACT-RS) and PACS-SET contain one or more items assessing patient difficulty. As this would, a priori, appear to hold relevance for assessing competence and answering questions the field may have (e.g., is it less difficult to be competent with "easy" patients), we recommend that this area should receive more empirical attention.

In summary, these competence measures provide the initial steps toward the creation of measures that the field needs to document intervention competency. However, limits of definitional clarity (adherence combined with competence) and limits of reliability are of concern, as is their limited focus on specific diagnoses. As these scales were developed for very particular purposes and were used in very controlled conditions, reliability would likely be lowered in less structured situations. Therefore, we advocate the development of new, easier to use scales that could be applicable to a wide range of practitioners.

What Is the Relation Between Competence and Therapy Outcome?

As clinicians and clinical scientists we are interested in the empirical question of whether or not competence matters. The answer to this seemingly simple question is, in the end, quite complex. Our intuition is that competent therapists should have better treatment outcomes than noncompetent therapists. Of course, this issue harkens back to the introductory reflections on the definition of competence and the difficulties contained therein. Operating under the assumption that the measures described above capture a "good enough" sense of what it means to be a competent therapist within their respective limited domains, does competence yield a better outcome?

The final column of Table 1 describes the relation between competence measures reviewed in this article and patient outcome. The majority of studies reported a small but positive correlation between competence and outcome in several limited domains. However, a few studies demonstrated no relation of competence to outcome, and many reported a positive correlation with some outcome measures but not with others. Moreover, two studies observed a negative association (i.e., the more competent the therapist, the worse the patients' outcome).

The inconsistency in the results from these studies may be due to several factors that we will broadly group into statistical–methodological and conceptual domains. Regarding the former, the relation of competence to outcome may be a small effect that is difficult to detect. In order to maximize detection, future studies would likely benefit from including large samples of patients and therapists. The effect of competence on outcome may have also been small due to a restricted range of competence. Because they were based on RCTs, the studies tended to rely on experienced or well-trained therapists, a factor likely leading to low variability in competence and/or outcome. Studies examining a wider range of competence (e.g., trainees as in Milne, Baker, Blackburn, James, & Reichelt, 1999) and outcome may help clarify the relation between the two. Further, competence may be a proximal but not distal predictor of outcome or have different relations to outcome at different stages of therapy. Longitudinal studies of competence and outcome are needed. Finally, there may be inconsistencies in the studies reviewed because none shared the same method of modeling the relation of competence to outcome. For instance, some lacked statistical measurement of the relation (Milne et al., 1999), and others looked at zero-order effects of competence on outcome (Huppert et al., 2001; O'Malley et al., 1988; Paivio et al., 2004). Baseline symptom level and symptom change up until the measurement of competence are often predictors of later outcome (cf. Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2000; Barber, Crits-Christoph, & Luborsky, 1996) and may need to be accounted for to clarify the relation of competence to subsequent symptom change.

Inconsistencies in the relation between competence and outcome may also be due to conceptual issues. As stated in the introduction, adherence and competence are conceptually distinct, but the degree to which they have been disjoined has not been consistent across measures. Adherence may have a very complex relation to outcome (Barber et al., 2006), and failing to fully separate out adherence may obfuscate any relation between competence and outcome. Further, competence may not have a direct relation to outcome but may instead moderate the effects of other process variables. For instance, therapeutic alliance or adherence may both be predictors of outcome in the presence of competence. Not enough studies have investigated the interactive effects of competence, alliance, and adherence. Similarly, curvilinear associations, like that found by Huppert et al. (2001), might be another way in which competence affects outcome. The highest quality delivery of certain interventions may not be what is related to outcome as we expect; rather, some more moderate level of competence could be more predictive of improvement. Finally, all of the measures reviewed assessed only limited-domain competence. It is indeed conceivable that a good outcome may result from competence demonstrated within a single domain, but it is also possible that demonstrating global competence (or even competence across multiple domains) may be more consistently predictive of outcome. This hypothesis, however, awaits further empirical support.

Discussion

We have briefly reviewed some of the conceptual issues related to intervention competency, first contrasting competence with related concepts and then proposing a distinction between global and limited-domain competence. We next reviewed and evaluated a number of measures designed to assess competence and provided recommendations for developing future measures. All measures reviewed were found to focus on the assessment of limited-domain competencies. Most (if not all) demonstrated some degree of reliability, especially if more than one judge was used. In the majority of cases, however, judges used to rate competence were themselves experts in the treatment they assessed. The necessity for using multiple expert clinicians as judges (as well as the question of how to define them) has implications for assessing therapists on a large scale, as does the very specific focus of these competence measures.

The extant data on competence and therapy outcome was also evaluated. We concluded that the evidence is consistent with a positive association between competence and patient outcome. However, it seems that the overall strength of the relation is weaker than would be expected, and as we are relying on published reports, the actual relation may be even less robust. We have discussed possible conceptual and methodological reasons for this weak relation.

Taken together, these findings possess implications for the field. The current state of competence assessment derived from RCTs is not yet ready to meet many of clinical psychology's needs for objective competence criteria (e.g., licensure, internship). However, we do have the beginnings of such tools. As psychologists (and the public) desire competence assessment for many purposes, it is inevitable that these different goals necessitate a more sophisticated competence assessment as well as a plurality of measures. Several of the field's purposes may perhaps best be served by global competence assessments pending further conceptual and empirical work on this construct (see also Kaslow & Keilin's [2006] discussion of *superordinate competencies*). However, as no measures described here were developed for this purpose, it is not surprising that none yet meet this requirement. In those cases where limited-domain assessment is the desired goal (i.e., in specialty certification), the available across-sites reliability data are not heartening. However, with more judge training, reliability could be improved. Thus, additional data are needed to test the possibility of widespread reliability and utilization.

Areas for Future Research and Conceptual Clarification

A number of unanswered questions related to competence assessment remain. First, the reviewed measures were developed for use in RCTs. Thus, they were utilized with manual-based treatments for patients with specific disorders. We are not aware of efforts to assess general practitioner competence in more general settings using these or other instruments. However, as discussed earlier, benchmarks are currently being formulated (ACBWG, 2007).

Second, we are not yet aware of the potential impact of diversity issues on competence assessment. Is it easier for therapists to be competent with patients who are similar? Do different groups differ in level of competence? Do diversity variables moderate outcome? These are important questions that have not yet been explored.

Third, given the emphasis on accountability, the formulation of thresholds for competence is paramount (see also ACBWG, 2007; Kaslow, 2004). However, situating these thresholds in a useful manner requires (a) reliable instruments possessing sufficient depth, breadth, and construct validity as well as (b) a clearly articulated sense of what level of competence is acceptable for what purpose. Neither is yet available and formalized in a readily usable manner, except perhaps for the CTS's redline of 39.

Fourth, questions remain over the best way to assess competence. Waltz et al. (1993) and Stiles et al. (1998) emphasize considering the overall therapeutic context (stage of therapy, etc.). We would agree, and taking this idea further, it may be the case that an ability to adequately and judiciously handle situations that occur at the extremes of therapy (e.g., coping with patient hostility or handling boundary violations) should also be evaluated, as they may be more indicative of therapeutic competence than behaviors evidenced during "business as usual" therapy. This ability to improvise and flexibly respond to unanticipated situations has been discussed in the literature (e.g., Binder, 2004). To understand a broader context, though, raters may need to listen to the entire therapy (or at least numerous continuous sessions) in order to adequately sample interchanges that are both representative and relatively uncommon. This would be very costly and difficult to do.

Fifth, the issue of using outcome as an alternative indicator of competence deserves exploration. This is present in other fields. In seeking out a heart surgeon, for instance, it is reasonable to investigate surgeons' reputation and credentials as well as their success rate. It stands to reason that a more competent surgeon would have a higher success (i.e., survival) rate. Similarly, we as consumers and researchers might want to assess psychologists' competence in this manner. Although there may be merits to this approach, there may be disadvantages as well. For instance, even more so than with surgical patients, outcome in psychotherapy is due to many factors besides the therapist's competence; and these factors are not always amenable to therapist control. Further, a competent surgeon who takes tough cases may demonstrate a lower success rate. Including information on outcome, however, may be a useful adjunct for assessing competence in a multi-trait, multi-method, multi-informant manner (see Kaslow, 2004).

Finally, in order to assess global competence, researchers and theorists will necessarily have to confront the issue of competence within and between paradigms of psychotherapy. This will lead to interesting questions, as it is possible (and indeed likely) that exemplars of competent behavior viewed from an intraparadigmatic perspective may be viewed as incompetent through an extraparadigmatic lens. There is some evidence that *master therapists* representing different therapeutic modalities more closely resemble other master therapists than they do less masterful therapists from their own respective modalities (Goldfried, Raue, & Castonguay, 1998). However, this research did not utilize the kinds of tools that we have reviewed in this paper and did not directly address cross-paradigmatic competence.

Dangers of Competence Assessment

In closing, assessing competence is a complex task, and there are a number of potential risks that should be borne in mind. Though operational definitions are necessary preconditions of controlled research, by their nature they are inherently incomplete. We are not assessing competence in itself but rather an operationalization of that construct. Reification is an ever-present threat, and this could very well lead to premature intellectual foreclosure of a key psychological construct with significant professional ramifications. Specifically, one of the potential dangers would be that third-party payers would utilize measures of competence in order to choose "competent" therapists for reimbursement. Again, this in itself is not problematic if we as a field are confident in our definitions of competence. However, due to the intricacy of the concept and the remaining questions proposed above, we are not yet at that juncture.

References

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57,* 1060–1073.

American Psychological Association. (2005, August). *Policy statement on evidence-based practice in psychology.* Washington, DC: Author.

Assessment of Competency Benchmarks Work Group. (2007). *A developmental model for defining and measuring competence in professional psychology.* Unpublished manuscript.

Barber, J. P., Connolly, M. B., Crits-Christoph, P., Gladis, L., & Siqueland, L. (2000). Alliance predicts patients' outcome beyond in-treatment change in symptoms. *Journal of Consulting and Clinical Psychology, 68,* 1027–1032.

Barber, J. P., & Crits-Christoph, P. (1996). Development of an adherence/competence scale for dynamic therapy: Preliminary findings. *Psychotherapy Research, 6,* 81–94.

Barber, J. P., Crits-Christoph, P., & Luborsky, L. (1996). Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *Journal of Consulting and Clinical Psychology, 64,* 619–622.

Barber, J. P., Gallop, R., Crits-Christoph, P., Barrett, M. S., Klostermann, S., McCarthy, K. S., & Sharpless, B. A. (in press). The role of alliance and techniques in predicting outcome of supportive-expressive dynamic therapy for cocaine dependence. *Psychoanalytic Psychology.*

Barber, J. P., Gallop, R., Crits-Christoph, P., Frank, A., Thase, M. E., Weiss, R. D., et al. (2006). The role of therapist adherence, therapist competence, and the alliance in predicting outcome of individual drug counseling: Results from the NIDA Collaborative Cocaine Treatment Study. *Psychotherapy Research, 16,* 229–240.

Barber, J. P., Krakauer, I., Calvo, N., Badgio, P. C., & Faude, J. (1997). Measuring adherence and competence of dynamic therapists in the treatment of cocaine dependence. *Journal of Psychotherapy Practice and Research, 6,* 12–14.

Barber, J. P., Liese, B., & Abrams, M. J. (2003). Development of the Cognitive Therapy Adherence and Competence Scale. *Psychotherapy Research, 13,* 205–221.

Barber, J. P., Mercer, D., Krakauer, I., & Calvo, N. (1996). Development of an adherence/competence rating scale for individual drug counseling. *Drug and Alcohol Dependence, 43,* 125–132.

Beutler, L. E. (1999). Manualizing flexibility: The training of eclectic therapists. *Journal of Clinical Psychology, 55*(4), 399–404.

Binder, J. L. (2004). *Key competencies in brief dynamic psychotherapy: Clinical practice beyond the manual.* New York: Guilford Press.

Borkovec, T. D., & Sharpless, B. (2004). Generalized anxiety disorder: Bringing cognitive behavior therapy into the valued present. In S. Hayes, V. Follette, & M. Linehan (Eds.), *Mindfulness and acceptance: Expanding the cognitive-behavior tradition* (209–242). New York: Guilford Press.

Carroll, K. M., Nich, C., Sifry, R. L., Nuro, K. F., Frankforter, T. L., Ball, S. A., et al. (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence, 57,* 225–238.

Chevron, E. S., & Rounsaville, B. J. (1983). Evaluating the clinical skills of psychotherapists: A comparison of techniques. *Archives of General Psychiatry, 40,* 1129–1132.

Davidson, K., Scott, J., Schmidt, U., Tata, P., Thornton, S., & Tyrer, P. (2004). Therapist competence and clinical outcome in the prevention of parasuicide by manual assisted cognitive behaviour therapy trial: The POPMACT study. *Psychological Medicine, 34*(5), 855–863.

Epstein, R. M., & Hundert, E. M. (2002). Defining and assessing professional competence. *Journal of the American Medical Association, 287,* 226–235.

Goldfried, M. R., Raue, P. J., & Castonguay, L. G. (1998). The therapeutic focus in significant sessions of master therapists: A combination of cognitive-behavioral and psychodynamic-interpersonal interventions. *Journal of Consulting and Clinical Psychology, 66,* 803–810.

Hoffart, A., Sexton, H., Nordahl, H. M., & Stiles, T. C. (2005). Connection between patient and therapist and therapist's competence in schema-focused therapy of personality problems. *Psychotherapy Research, 15*(4), 409–419.

Hogue, A., Chinchilla, P., Fried, A., Dauber, S., Liddle, H. A., Henderson, C., et al. (2007). *A contextual approach to evaluating the fidelity of behavioral interventions for adolescent substance abuse.* Unpublished manuscript.

Huppert, J. D., Bufka, L. F., Barlow, D. H., Gorman, J. M., Shear, M. K., & Woods, S. W. (2001). Therapists, therapist variables, and cognitive-behavioral therapy outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69*(5), 747–755.

Jacobson, N. S. (1998, July). *How important are cognitive interventions in cognitive therapy?* Paper presented at the World Congress of Behavioral and Cognitive Therapies, Acapulco, Mexico.

Kaslow, N. J. (2004). Competencies in professional psychology. *American Psychologist, 59,* 774–781.

Kaslow, N. J., & Keilin, W. G. (2006). Internship training in clinical psychology: Looking into our crystal ball. *Clinical Psychology: Science and Practice, 13,* 242–248.

Kingdon, D., Tyrer, P., Seivewright, N., Ferguson, B., & Murphy, S. (1996). The Nottingham Study of Neurotic Disorder: Influence of cognitive therapists on outcome. *British Journal of Psychiatry, 169,* 93–97.

Milne, D. L., Baker, C., Blackburn, I., James, I., & Reichelt, K. (1999). Effectiveness of cognitive therapy training. *Journal of Behavior Therapy and Experimental Psychiatry, 30,* 81–92.

O'Malley, S. S., Foley, S. H., Rounsaville, B. J., Watkins, J. T., Sotsky, S. M., Imber, S. D., & Elkin, I. (1988). Therapist competence and patient outcome in interpersonal psychotherapy of depression. *Journal of Consulting and Clinical Psychology, 56,* 496–501.

Paivio, S. C., Holowaty, K. A. M., & Hall, I. E. (2004). The influence of therapist adherence and competence on client reprocessing of child abuse memories. *Psychotherapy, 41,* 56–68.

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12,* 365–383.

Shaw, B. F., Elkin, I., Yamaguchi, J., Olmstead, M., Vallis, T. M., Dobson, K. S., et al. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 67,* 837–846.

Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice, 5*(4), 439–458.

Svartberg, M. (1989). Manualization and competence monitoring of short-

term anxiety-provoking psychotherapy. *Psychotherapy: Theory, Research, and Practice, 26,* 564–571.

Svartberg, M., & Stiles, T. C. (1994). Therapeutic alliance, therapist competence, and client change in short-term anxiety-provoking psychotherapy. *Psychotherapy Research, 4*(1), 20–33.

Trepka, C., Rees, A., Shapiro, D. A., Hardy, G. E., & Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research, 28*(2), 143–157.

Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986). The cognitive therapy scale: Psychometric properties. *Journal of Consulting and Clinical Psychology, 54*(3), 381–385.

Waltz, J., Addis, M. E., Koerner, K., & Jacobson, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology, 61*(4), 620–630.

Young, J., & Beck, A. T. (1980). *Cognitive Therapy Scale: Rating manual.* Unpublished manuscript, University of Pennsylvania.

Received December 1, 2006
Revision received June 28, 2007
Accepted June 29, 2007 ■