

BRIEF REPORT

 Use of Common and Unique Techniques in the Early Treatment Phase for
Cognitive-Behavioral, Interpersonal/Emotional, and Supportive Listening
Interventions for Generalized Anxiety Disorder

 Brittany R. King, James F. Boswell,
and Carly M. Schwartzman
University at Albany, State University of New York

 Kyler Lehrbach
Rhode Island Hospital, Providence, Rhode Island

 Louis G. Castonguay and Michelle G. Newman
The Pennsylvania State University

Psychotherapy research often compares specific treatments to control conditions to establish efficacy of the specified treatment. Research has typically evaluated common factor elements (e.g., credibility, expectancy) in treatments only after the first or second session, largely as a manipulation check and under the assumption that such factors are static. This study observed therapist common factor and model-specific interventions in three treatment approaches from a randomized control trial for generalized anxiety disorder across the entire early phase of treatment (i.e., first five sessions). The parent randomized control trial compared two treatment conditions, using an additive design where patients were randomized to receive either interpersonal/emotional processing interventions or supportive listening after receiving a session of cognitive-behavioral therapy. The first five video-recorded sessions of $N = 40$ randomly sampled participants were observationally coded with a multidimensional intervention measure, with subscales reflecting diverse theoretical orientations and common factors. Multilevel modeling was used to examine intervention use and investigate differences between treatment conditions and segments. Among the results, common factor interventions were rated as significantly more typical in cognitive-behavioral therapy compared with supportive listening. The pattern of intervention use of other subscales was generally consistent with the orientation of the respective protocols. In the early phase of treatment, supportive listening conditions do not appear to function as common factor controls in the manner that many might assume. Common factors are potentially enhanced in bona fide treatments that include a more detailed, specific rationale and clear and cohesive techniques and goals.



Clinical Impact Statement

Question: In a randomized control trial for generalized anxiety disorder, what is the relative typicality of common factor and model-specific interventions in the early phase of treatment in more active and directive treatments compared with a supportive control condition? **Findings:** Common factor interventions were rated as significantly more typical in a more directive CBT treatment compared with a supportive listening control treatment. **Meaning:** Results provide evidence that supportive listening treatments may not function as common factor “controls” in the way that many might assume. **Next Steps:** Future research should continue to explore therapist intervention use in the early phase of treatment and across the entire course of therapy, with additional clinical populations, in naturalistic settings, and with a larger sample of therapists.

Keywords: common factors, early treatment, generalized anxiety disorder, psychotherapy process, interventions

Supplemental materials: <http://dx.doi.org/10.1037/pst0000277.supp>

This article was published Online First January 16, 2020.

Brittany R. King, James F. Boswell, and Carly M. Schwartzman, Department of Psychology, University at Albany, State University of New York;  Kyler Lehrbach, Department of Emergency Medicine, Rhode Island Hospital, Providence, Rhode Island; Louis G. Castonguay and  Michelle G.

Newman, Department of Psychology, The Pennsylvania State University.

Correspondence concerning this article should be addressed to Brittany R. King, Department of Psychology, University at Albany, State University of New York, Social Science 399, 1400 Washington Avenue, Albany, NY 12222. E-mail: bking3@albany.edu

Psychotherapy researchers often compare bona fide psychotherapy and control conditions in randomized control trials (RCTs) to establish efficacy of specified treatments. A variety of control conditions are used in such trials, including wait-list, treatment-as-usual, or “supportive therapy.” Active control conditions of this type ostensibly control for common factors (vs. model-unique interventions), such as support, attention, outcome expectancies, credibility, and the working alliance. In this way, researchers conclude that differences in outcomes can be attributed to the active ingredients or interventions specific to the unique therapy tested (Rosenthal & Frank, 1956).

As such, it is asserted that only when psychotherapy researchers determine that experimental and control conditions are equivalent with regard to common factors can strong conclusions about the active treatment be made (Boot, Simons, Stothart, & Stutts, 2013). Yet, perhaps with the exception of the working alliance, assessing and establishing such equivalence is rarely done beyond the first or second treatment session (Boot et al., 2013). Some RCTs fail to measure and compare common factors in treatment arms at all, whereas others report collecting patient ratings of a small subset of these constructs (e.g., expectancies, credibility) only once after the first or second session. In the context of controlled trials, we might expect comparable common factor elements this early in treatment (e.g., very soon after the provision of a treatment rationale), but these factors may begin to fluctuate even shortly after the first or second session, thus, highlighting the need for investigating these elements across treatment conditions, beyond the first couple of sessions (Constantino, Coyne, Boswell, Iles, & Višlā, 2018; Constantino, Višlā, Coyne, & Boswell, 2018). The early phase of treatment (often considered to be the first five sessions; Lutz et al., 2014) appears to be a particularly sensitive and important period in the psychotherapy change process (Flückiger, Grosse Holtforth, Znoj, Caspar, & Wampold, 2013), as highlighted by a recent special section in *Psychotherapy* dedicated to the topic of early treatment (Bedics, 2019). Perhaps, not surprisingly, several articles in this special section referenced the importance of common factors (King & Boswell, 2019). Therefore, it is important to examine these factors in conditions across a broader sample of the early treatment phase.

Direct and independent observations of therapists’ common technical (as well as model-unique) factors across sessions remain rare. In the context of the National Institute of Mental Health Treatment of Depression Collaborative Research Program (Elkin et al., 1989), Barnicot, Wampold, and Priebe (2014) reported that clinicians in the clinical management condition were rated on average to be less genuine and less empathic than the clinicians in the psychotherapy conditions. They argued that active psychotherapy conditions include a more detailed, specific rationale, as well as clear and cohesive techniques and goals. Active psychotherapy conditions are better equipped to engage and harness facilitative common factors (Wampold & Budge, 2012). Importantly, Barnicot et al. (2014) utilized patient self-report data collected after the second treatment session alone.

Based on this result and existing theory, we do not believe it is a “straw man” premise to question whether or not such “supportive listing” (SL) control conditions are indeed common factor controls/treatments. This does not necessarily imply that supportive control conditions are explicitly intended to *optimize* common factors. In fact, it may well be that common facilitative behaviors

(aimed, for example at strengthening the alliance) are more prevalent in active treatment conditions, yet this has received less attention across multiple early treatment sessions. There is much to learn about how therapist facilitative common factor behaviors may or may not differ when administering experimental and control conditions. This is also important because, unlike double-blind trials for pharmacological treatments, therapists are typically aware of which therapy they are delivering (that is, therapists are not blind to treatment condition), which may impact in-session behaviors and patient perceptions (Wampold, Frost, & Yulish, 2016).

Using archived audio-video recordings sampled from an RCT for generalized anxiety disorder (GAD; Newman et al., 2011), the aim of this study was to directly observe therapist interventions, both model-specific and common factors, in three treatment approaches during the entire early phase of treatment. Using an additive design, the parent RCT was a comparison between two conditions: CBT plus interpersonal/emotional processing (CBT plus I/EP) and CBT plus SL. The first five sessions of randomly sampled cases were coded using an assessment tool that included model-unique and common factor subscales. We investigated differences, in terms of level and pattern of interventions, between the conditions (CBT plus I/EP vs. CBT plus SL) and segments (CBT, I/EP, and SL).

Method

Participants

We randomly sampled $n = 40$ cases from the larger $N = 70$ completer sample in the Newman et al. (2011) RCT (20 CBT plus I/EP cases and 20 CBT plus SL cases). Although the current study was focused on psychotherapy process, the patient cases were sampled to be representative of a range of treatment outcomes (i.e., different degrees of responders and nonresponders). For additional details regarding the trial and recruitment, please see Newman et al. (2011). All participants were adults who met criteria for a principal GAD diagnosis. Demographic and baseline characteristics of the current study sample are shown in Table 1. Chi-square and t test indicated that the characteristics of this subsample did not differ significantly from the overall sample.

Measures

Multitheoretical List of Therapeutic Interventions. The Multitheoretical List of Therapeutic Interventions (MULTI) was used to observationally rate therapist intervention use (McCarthy & Barber, 2009). The MULTI is comprised of 60 items rated on a 5-point scale (1 = *not at all typical of the session*; 5 = *very typical of the session*). Items assess specific behaviors of the therapist and are written in a jargon-neutral manner to avoid drawing upon theoretical biases of the raters. It consists of eight subscales: Cognitive Therapy, Behavioral Therapy, Interpersonal Therapy, Dialectical-Behavioral Therapy, Psychodynamic, Person-Centered, Process-Experiential, and Common Factors. Of particular interest in this study is the Common Factors subscale, which includes the following seven items: “Worked to give the client hope or encouragement,” “Conveyed belief in the effectiveness of the methods he/she was using to help the client,” “Was warm, sympathetic, and accepting,” “Made the

Table 1
Demographic and Baseline Characteristics of the Sample

Characteristic	<i>M(SD)</i> or %(<i>n</i>)
Age, <i>M(SD)</i>	39.0 (12.7)
Gender	
Female	75% (<i>n</i> = 30)
Male	25% (<i>n</i> = 10)
Ethnicity	
White	90% (<i>n</i> = 36)
Asian	5% (<i>n</i> = 2)
Hispanic	2.5% (<i>n</i> = 1)
Education	
High school	22.5% (<i>n</i> = 9)
College	47.5% (<i>n</i> = 15)
Master's	22.5% (<i>n</i> = 9)
PhD	7.5% (<i>n</i> = 3)
Marital status	
Married	42.5% (<i>n</i> = 17)
Single	35% (<i>n</i> = 14)
Divorced	20% (<i>n</i> = 8)
HARS, <i>M(SD)</i>	23.0 (6.1)
PSWQ, <i>M(SD)</i>	67.8 (8.0)
GAD duration, years, <i>M(SD)</i>	12.5 (16.0)

Note. HARS = Hamilton Anxiety Rating Scale; PSWQ = Penn State Worry Questionnaire; GAD = generalized anxiety disorder.

session a place where the client could get better to solve his/her problems," "Worked with the client together as a team," "Listened carefully to what the client was saying," and "Focused on improving the client's ability to solve his/her own problems." The MULTI has demonstrated adequate-to-good internal consistency, interrater reliability, and validity (McCarthy & Barber, 2009).

A group of two doctoral students and two advanced undergraduate research assistants were extensively trained over a period of four months to criterion reliability (average subscale intraclass coefficient [ICC]_[2,2] = .80) with training videos (Shrout & Fleiss, 1979). For each participant, a random pair of judges was selected to rate each session. The trial's additive design meant that each session was roughly two hours in length. Regardless of condition, the first 50 min involved CBT. The second hour/session is where participants differed by either receiving 50 min of I/EP or 50 min of SL. Both raters coded each segment (CBT, I/EP, or SL) for each patient. Sessions were rated independently, and then each rater met to achieve consensus. The consensus ratings were used in the present analyses. The use of consensus or average coder ratings can be found in published research. We elected to use consensus ratings, in part, because they yield scores that are consistent with the original measure ratings/metric and averaging potentially reduces variance. Coders met periodically with the second author to discuss rating questions to prevent drift. Using initial independent ratings of MULTI items, overall interrater reliability for the coded sample was ICC_[2,2] = .804, which is considered good-to-excellent (Shrout, 1995). Item-level ICCs ranged between .61 and .94. Although segment type was not directly disclosed to raters, it is likely inaccurate to state that the raters were "blind." The basic differences between the segments were broadly apparent to the raters when viewing the sessions. The raters were not any more or less blind than raters in a typical psychotherapy RCT.

Control variables. Although not a primary focus of the current study, we wanted to increase confidence that the randomly

sampled participants did not differ from the larger sample on certain baseline variables. We similarly wanted to check if differences existed between conditions. Consequently, we examined baseline levels of the (a) Penn State Worry Questionnaire (Meyer, Miller, Metzger, & Borkovec, 1990), (b) Hamilton Anxiety Rating Scale (Hamilton, 1959), and (c) duration of GAD.

Procedure

The Newman et al. (2011) trial adopted an additive design where participants in both conditions always began with 50 minutes of CBT and then received 50 minutes of either I/EP or SL. Both treatments involved a total of 14 approximately 2-hr sessions. We operationalized the early phase of treatment as comprising the first five sessions, which is consistent with existing literature (Lutz et al., 2014).

Treatment

Both treatments in the RCT were conducted by three doctoral-level therapists with at least two years of experience with both treatments. Therapists were crossed such that each therapist participated in both conditions. The CBT sessions followed a manualized CBT protocol previously developed and used in past trials at Pennsylvania State University (Borkovec, Newman, Pincus, & Lytle, 2002). The I/EP protocol was developed by the trial investigators and sought to process client emotions and needs within an interpersonal context. SL sessions functioned as a control condition and followed a manual developed in previous research trials (Borkovec et al., 2002). During SL sessions, clients were to explore experiences and events with minimal guidance from the therapist, and therapists were prohibited from using any directive methods and instead facilitated discussions with reflections, supportive statements, and empathic listening. As described in the original trial report, a subset of sessions were independently rated for adherence and competence, and all were considered satisfactory (Newman et al., 2011).

Data Analysis

MULTI subscale scores were calculated by averaging the respective orientation-specific and common factor items. The levels and slopes (trajectories over the first five sessions) for the MULTI subscales were examined with multilevel models (Raudenbush & Bryk, 2002). We elected to test two-level models (rather than three-level models due to the small number of therapists), with clients at level 2 and session-to-session intervention ratings within clients at level 1. The following variables were entered as predictors of intervention use: treatment condition (testing the difference between the experimental and control conditions), segment (testing the difference between the individual segments of CBT, I/EP, and SL), time (i.e., session/weeks 1–5), as well as condition by time (differences in intervention trajectory as a function of experimental vs. control condition), and segment by time interactions (differences in intervention trajectory as a function of CBT, I/EP, or SL segment). Although we were primarily interested in common factors between the segments (e.g., SL vs. CBT), segments were technically nested within treatment condition, and it seemed most prudent to include the condition effect in the models. Models were

run in SAS 9.4 (SAS Institute Inc, 2013). Due to the small sample size, model estimation was conducted with restricted maximum likelihood estimation and a Kenward–Roger adjustment to the degrees of freedom (Kenward & Roger, 1997).

Results

Preliminary Analyses

Preliminary analyses indicated that the coded subsample did not differ from the larger RCT sample on demographic variables, initial anxiety severity, or duration of GAD ($ps > .05$). Table 2 reports bivariate correlations among the eight MULTI subscales across all sessions and segments. Due to the high positive correlations among the Cognitive Therapy, Behavioral Therapy, and Dialectical-Behavioral Therapy subscales ($r = .86$ to $.95$), these were combined (averaged) to create a Directive Interventions subscale. Similarly, Psychodynamic and Person-Centered ($r = .76$) subscales were combined to create an Exploratory Interventions subscale. Thus, five subscale variables were used as dependent variables in the analyses. Previous studies have similarly aggregated the MULTI subscales (Castonguay et al., 2017). Some segments ($n = 44$ out of 400; 11%) were determined to be uncodeable due to administrative error or poor quality. We created a dummy variable to compare patients and sessions with and without any missing segments. Analysis of variance and χ^2 tests were conducted to compare these two groups on baseline characteristics, including initial severity; none of these tests were statistically significant ($ps > .34$). We determined that it was appropriate to consider the data to be missing at random. We used multiple imputation procedures in SAS to estimate session segment-level MULTI ratings (Rubin, 1996), and analyses were conducted with the imputed data.

Intervention Descriptives

The subscale means (and standard deviations) across all sessions and segments are as follows: Common Factor: 3.71 (.29); Directive: 2.26 (.32); Exploratory: 2.19 (.27); Process-Experiential: 1.92 (.37); Interpersonal: 1.50 (.34). Common Factor use was most typical overall, followed by directive and exploratory interven-

tions. Figure 1 displays the observed subscale means by treatment segment (CBT, I/EP, and SL). Although purely descriptive, the patterns indicate that level of intervention use varied within each segment type, yet these were generally consistent with the segment protocol.

Intervention Models

Multilevel models were used to test (a) condition, (b) segment, (c) time, (d) Condition \times Time, and (e) Segment \times Time as fixed effect predictors of intervention use. Five separate models were tested for each subscale. Given the primary focus of the current analysis, we report results for the full common factors model in the following text and in Table 3. Summaries of the other model results are also provided (tables with full model results can be found in the online supplemental materials).

In the model predicting common factor use (Table 3), significant main effects for condition, $F(1, 75.7) = 8.61, p < .01$, and segment, $F(2, 354) = 5.05, p < .01$, were observed. Common factors were observed to be *less* typical in the CBT-SL control condition than in the CBT-I/EP experimental condition. In addition, common factors were significantly *more* typical of CBT segments compared with SL segments. No statistically significant difference emerged between the CBT and I/EP segments.

Results from the model predicting directive interventions use indicated significant main effects of condition, $F(1, 87.8) = 18.73, p < .0001$, and segment, $F(2, 354) = 124.37, p < .001$, such that directive interventions were less typical in the CBT-SL condition compared with CBT-I/EP, and significantly more typical in the CBT segments compared with both I/EP and SL (see Table S1 in the online supplemental materials for full model results). In the model predicting exploratory intervention use, significant main effects for condition, $F(1, 71.9) = 22.27, p < .0001$, and segment, $F(2, 354) = 114.59, p < .001$, were also found, such that exploratory interventions were less typical in the CBT-SL condition compared with CBT-I/EP, and also significantly less typical in the CBT segments compared to both I/EP and SL segments (see Table S2 in the online supplemental materials for full results).

In the model predicting process-experiential intervention use, significant main effects for condition, $F(1, 90.7) = 49.44, p < .0001$, and segment, $F(2, 354) = 34.37, p < .001$, were observed, such that process-experiential interventions were less typical in the CBT-SL condition than in CBT-I/EP. Process-experiential interventions were significantly less typical in the CBT segment when specifically compared with the I/EP segment, yet process-experiential interventions were significantly more typical in the CBT segment compared with the SL segment (see Table S3 in the online supplemental materials for full results). In the model predicting interpersonal therapy use, significant main effects for condition, $F(1, 89.8) = 31.52, p < .0001$, and segment, $F(2, 354) = 80.31, p < .001$, were observed, such that interpersonal interventions were less typical in the CBT-SL condition than in CBT-I/EP and significantly less typical in the CBT segment compared with both I/EP and SL segments (see Table S4 in the online supplemental materials for full results). For all models, no significant main effects of time or interactions with time were observed.

Table 2

Bivariate Correlations of Eight MULTI Subscales Across All Sessions and Segments

Subscale	DBT	BT	CF	CT	IPT	PC	PD
DBT							
BT	.95*						
CF	.53*	.57*					
CT	.86*	.92*	.61*				
IPT	-.09	-.07	.06	-.07			
PC	-.06	-.07	.24*	.03	.61*		
PD	-.08	-.1	.01	-.05	.78*	.76*	
PE	.53*	.56*	.46*	.58*	.50*	.61*	.54*

Note. MULTI = Multitheoretical List of Therapeutic Interventions; DBT = dialectical-behavioral therapy; BT = behavior therapy; CF = common factors; CT = cognitive therapy; IPT = interpersonal therapy; PC = person-centered; PD = psychodynamic; PE = process-experiential. * $p < .01$.

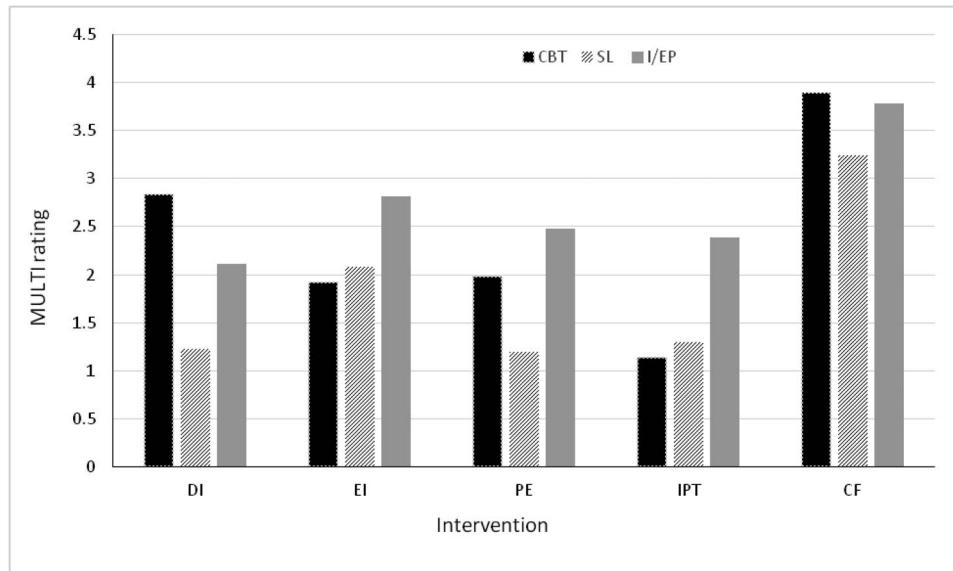


Figure 1. Means of the five Multitheoretical List of Therapeutic Interventions (MULTI) intervention subscales by segment. Values on the y axis represent mean ratings of intervention use on the 1 to 5 rating scale, with higher values indicating more intervention use. The x axis includes the five MULTI subscales. The bars represent segment type. CBT = cognitive-behavioral therapy; SL = supportive listening; I/EP = interpersonal/emotional processing; DI = directive interventions; EI = exploratory interventions; PE = process-experiential; IPT = interpersonal; CF = common factors.

Discussion

The use of control conditions is an important element in comparative psychotherapy research. Even when not directly assessed or compared, the use of supportive-listening type control conditions is assumed to account for common therapeutic factors associated with most active treatments. When directly assessed, research has typically focused on a single early session. This approach assumes that such factors remain relatively stable. This assumption may not be tenable, and accumulating research highlights the need for more time points to yield more dependable estimates of process variables (Dennhag, Gibbons, Barber, Gallop, & Crits-Christoph, 2012).

The current study assessed therapist use of common factor and model-unique interventions in experimental and control treatments for GAD, via direct and independent observation of trained observers. Results indicated that therapists' common factor interventions were rated as significantly higher (i.e., more typical) in the CBT segment compared with the supportive listening segment, across the early phase of treatment. Although orientation-specific interventions were also evaluated and compared, the ratings were in the expected direction given the nature of the study (e.g., directive interventions were rated highest in CBT conditions) and might be viewed as a further manipulation check, as well as further support for the validity of the MULTI.

Table 3
Mixed Model for Common Factor Intervention Use

Fixed effect	Estimate	SE	df	t	p	95% confidence interval	
						LL	UL
Condition segment	-.40	.11	144	-3.48	.001	-.63	-.17
CBT	—	—	—	—	—	—	—
SL	.28	.09	354	3.14	.002	.11	.46
I/EP	.04	.09	354	.48	.630	-.13	.22
Time	.00	.00	354	-.62	.534	-.01	.00
Time × Condition	.00	.01	354	-.69	.489	-.01	.01
Time × Segment							
CBT	—	—	—	—	—	—	—
SL	.02	.01	354	3.33	.001	.00	.03
I/EP	.01	.01	354	2.55	.011	.00	.02

Note. CBT = cognitive-behavioral therapy; SL = supportive listening; I/EP = interpersonal/emotional processing. Condition = CBT + I/EP vs. CBT + SL. Time represents Sessions 1 to 5.

Given the difference in common factor ratings between segments, these results suggest that it may be misleading to label supportive listening conditions as common factor control treatments. If simply based on relative typicality, CBT represented more of a common factor treatment condition than the ostensibly common factor control condition. Although speculative, there are several reasons for why this may occur. First, it may be the case that there were fewer opportunities for therapists to engage in common factor facilitating interventions in the supportive listening segment, compared with the more active and directive CBT approach. As noted, Barnicot et al. (2014) compared patient perceptions of common factors between CBT, interpersonal, and a clinical management condition for depression. Empathy and genuineness were rated significantly lower by patients in the clinical management group compared with the CBT and interpersonal groups. Indeed, the authors suggest that therapists' ability to utilize common factors may be hampered when providing a treatment that lacks a coherent rationale or specific techniques, which may subsequently hinder the development of a positive working therapeutic relationship. Common factors are potentially enhanced in more active treatments.

Second, when used in psychotherapy research, unlike double-blind pharmacological trials, therapists are likely to know if or when they are providing an active or control treatment, which may activate biases of different approaches and affect in-session behavior (Borkovec & Sibrava, 2005; Wampold et al., 2016). Ultimately, it appears difficult to create a control condition that consistently matches the treatment being studied on common factor elements (Borkovec & Nau, 1972). Still, it is argued that conclusions drawn from RCTs depend as much on control conditions as it does the active treatments (Mohr et al., 2009).

There are several limitations of the current study. First, although large in comparison to similar observational studies, this study involved a relatively small sample. The sample was also relatively homogenous in terms of client gender and race, and the RCT included only individuals with GAD, thus making it difficult to generalize to other individuals and clinical populations. In addition, the fixed order of the treatment segments (i.e., CBT always occurred before I/EP and SL segments) increases internal validity and yet could conceivably have unintended consequences. Therapist behavior might change as a function of time—what is delivered in the first hour of an interaction versus the second. In such a case, the relatively less typical common factor use is explained by the supportive listening condition always occurring later in the interaction. It is difficult to completely rule out this possibility. However, it is important to note that common factor use was rated significantly higher in the CBT-interpersonal/emotional processing condition compared with the CBT-supportive listening condition, and at the segment level, the difference between CBT and the interpersonal/emotional processing segment was not statistically significant. If the result was better accounted for by an ordering effect, a similar difference would be expected between these two segments. The fixed ordering of segments also makes it difficult to generalize findings to more naturalistic settings. Finally, the literature identifies many common factor elements and, although the MULTI items capture prevalent common factors, this is a single measure that adopts a particular perspective. It is possible that other common factor conceptualizations and measures (i.e., addi-

tion or removal of additional common factors) could yield different findings. Certainly the present results require replication.

Despite these limitations, the present results imply that researchers should more carefully evaluate common factor intervention use across conditions in psychotherapy trials. Supportive listening conditions do not appear to function as one-to-one common factor controls, in the manner that many might assume. Broadly, replication of this study is needed, and, resources permitting, future studies should also focus on covering the entire course of treatment to provide additional context. Technology may allow researchers to scale up their coding methodology, resulting in less resource-intensive coding of treatment activities. For example, there have been advances in automated coding of psychotherapy sessions and transcripts, such as the coding of therapist empathy (Xiao, Imel, Georgiou, Atkins, & Narayanan, 2015), using natural language processing and machine learning methods. In addition, future studies should assess common factor use in other clinical populations, in naturalistic settings, with different conceptualizations of common factors, and with a larger sample of therapists.

References

- Barnicot, K., Wampold, B., & Priebe, S. (2014). The effect of core clinician interpersonal behaviours on depression. *Journal of Affective Disorders, 167*, 112–117. <http://dx.doi.org/10.1016/j.jad.2014.05.064>
- Bedics, J. (2019). Initiating the therapeutic encounter: Beginning psychotherapy [special section]. *Psychotherapy, 56*, 5–6. <http://dx.doi.org/10.1037/pst0000223>
- Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science, 8*, 445–454. <http://dx.doi.org/10.1177/1745691613491271>
- Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3*, 257–260. [http://dx.doi.org/10.1016/0005-7916\(72\)90045-6](http://dx.doi.org/10.1016/0005-7916(72)90045-6)
- Borkovec, T. D., Newman, M. G., Pincus, A. L., & Lytle, R. (2002). A component analysis of cognitive-behavioral therapy for generalized anxiety disorder and the role of interpersonal problems. *Journal of Consulting and Clinical Psychology, 70*, 288–298. <http://dx.doi.org/10.1037/0022-006X.70.2.288>
- Borkovec, T. D., & Sibrava, N. J. (2005). Problems with the use of placebo conditions in psychotherapy research, suggested alternatives, and some strategies for the pursuit of the placebo phenomenon. *Journal of Clinical Psychology, 61*, 805–818. <http://dx.doi.org/10.1002/jclp.20127>
- Castonguay, L. G., Janis, R. A., Youn, S. J., Xiao, H., McAleavey, A. A., Boswell, J. F., . . . O'Leary, M. (2017). Clinicians' prediction and recall of therapeutic interventions: A practice research network study. *Counselling Psychology Quarterly, 30*, 308–322. <http://dx.doi.org/10.1080/09515070.2017.1334628>
- Constantino, M. J., Coyne, A., Boswell, J. F., Iles, B. R., & Vísňá, A. (2018). The credibility-outcome association: A meta-analysis. *Psychotherapy, 55*, 486–495. <http://dx.doi.org/10.1037/pst0000168>
- Constantino, M. J., Vísňá, A., Coyne, A., & Boswell, J. F. (2018). The expectation-outcome association: A meta-analysis. *Psychotherapy, 55*, 473–485. <http://dx.doi.org/10.1037/pst0000169>
- Dennhag, I., Gibbons, M. B., Barber, J. P., Gallop, R., & Crits-Christoph, P. (2012). How many treatment sessions and patients are needed to create a stable score of adherence and competence in the treatment of cocaine dependence? *Psychotherapy Research, 22*, 475–488. <http://dx.doi.org/10.1080/10503307.2012.674790>
- Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., . . . Parloff, M. B. (1989). National institute of mental health treatment of depression collaborative research program: General effec-

- tiveness of treatments. *Archives of General Psychiatry*, 46, 971–982. <http://dx.doi.org/10.1001/archpsyc.1989.01810110013002>
- Flückiger, C., Grosse Holtforth, M., Znoj, H. J., Caspar, F., & Wampold, B. E. (2013). Is the relation between early post-session reports and treatment outcome an epiphenomenon of intake distress and early response? A multi-predictor analysis in outpatient psychotherapy. *Psychotherapy Research*, 23, 1–13. <http://dx.doi.org/10.1080/10503307.2012.693773>
- Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, 32, 50–55. <http://dx.doi.org/10.1111/j.2044-8341.1959.tb00467.x>
- Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53, 983–997. <http://dx.doi.org/10.2307/2533558>
- King, B. R., & Boswell, J. F. (2019). Therapeutic strategies and techniques in early cognitive-behavioral therapy. *Psychotherapy*, 56, 35–40. <http://dx.doi.org/10.1037/pst0000202>
- Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., . . . Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of Consulting and Clinical Psychology*, 82, 287–297. <http://dx.doi.org/10.1037/a0035535>
- McCarthy, K. S., & Barber, J. P. (2009). The Multitheoretical List of Therapeutic Interventions (MULTI): Initial report. *Psychotherapy Research*, 19, 96–113. <http://dx.doi.org/10.1080/10503300802524343>
- Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 28, 487–495. [http://dx.doi.org/10.1016/0005-7967\(90\)90135-6](http://dx.doi.org/10.1016/0005-7967(90)90135-6)
- Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., . . . Kaplan, R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics*, 78, 275–284. <http://dx.doi.org/10.1159/000228248>
- Newman, M. G., Castonguay, L. G., Borkovec, T. D., Fisher, A. J., Boswell, J. F., Szkodny, L. E., & Nordberg, S. S. (2011). A randomized controlled trial of cognitive-behavioral therapy for generalized anxiety disorder with integrated techniques from emotion-focused and interpersonal therapies. *Journal of Consulting and Clinical Psychology*, 79, 171–181. <http://dx.doi.org/10.1037/a0022489>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
- Rosenthal, D., & Frank, J. D. (1956). Psychotherapy and the placebo effect. *Psychological Bulletin*, 53, 294–302. <http://dx.doi.org/10.1037/h0044068>
- Rubin, D. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 91, 473–489. <http://dx.doi.org/10.1080/01621459.1996.10476908>
- SAS Institute Inc. (2013). *What's new in SAS® 9.4*. Cary, NC: Author.
- Shrout, P. E. (1995). Measuring the degree of consensus in personality judgments. In P. E. Shrout & S. T. Fiske (Eds.), *Personality research, methods, and theory: A festschrift honoring Donald W. Fiske* (pp. 79–92). Hillsdale, NJ: Erlbaum, Inc.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428. <http://dx.doi.org/10.1037/0033-2909.86.2.420>
- Wampold, B. E., & Budge, S. L. (2012). The 2011 Leona Tyler Award address: The relationship and its relationship to the common and specific factors of psychotherapy. *The Counseling Psychologist*, 40, 601–623. <http://dx.doi.org/10.1177/0011000011432709>
- Wampold, B. E., Frost, N. D., & Yulish, N. E. (2016). Placebo effects in psychotherapy: A flawed concept and a contorted history. *Psychology of Consciousness: Theory, Research, and Practice*, 3, 108–120. <http://dx.doi.org/10.1037/cns0000045>
- Xiao, B., Imel, Z. E., Georgiou, P. G., Atkins, D. C., & Narayanan, S. S. (2015). “Rate my therapist”: Automated detection of empathy in drug and alcohol counseling via speech and language processing. *PLoS ONE*, 10, e0143055. <http://dx.doi.org/10.1371/journal.pone.0143055>

Received June 17, 2019

Revision received October 30, 2019

Accepted November 8, 2019 ■