

A Further Investigation of the Good-Enough Level Model Across Outcome Domains and Termination Status

D. Martin Kivlighan III
University of Iowa

Yi-Jiun Lin
University of Maryland, College Park

Karen P. Egan
University of Virginia

Theodore Pickett Jr.
University of Maryland, College Park

Simon B. Goldberg
University of Wisconsin—Madison

Accurate estimations of progress in psychotherapy are necessary for therapists to identify clients at risk of deterioration and potentially reduce premature terminations. This need has resulted in a large body of literature examining the rate and trajectory of change in psychotherapy; however, few studies have tested these dose-response relationships outside of global measures of mental health. Moreover, there is a paucity of research examining the relationship between progress in treatment, treatment length, and premature termination. In this study, we conducted multivariate multilevel analyses to test the good-enough level model across the three domains of the phase model of psychotherapy: psychological symptoms, life functioning, and well-being. In addition, we tested changes in well-being, psychological symptoms, and life functioning, treatment length, and an interaction between treatment progress and treatment length as predictors of premature termination. Data for this study consisted of 438 clients who were treated by 57 therapists within a brief therapy model. Results failed to support the good-enough level model for changes in well-being, psychological symptoms, and life functioning, such that the rate of change across all three scales did not significantly vary as a function of treatment length. However, exploratory analyses revealed a significant interaction effect between changes in well-being, treatment length, and premature termination, indicating that clients who experienced high rates of change in well-being early in treatment were more likely to prematurely terminate treatment. Clinical implications and future research are discussed.

Clinical Impact Statement

This study examined the good enough model (GEL) across multiple outcomes and termination status.

Question: We tested the relationship between treatment progress across multiple outcomes, treatment length, and the probability of dropout. **Findings:** This study failed to support the GEL model for changes in well-being, symptom reduction, or life functioning. Early and rapid changes in well-being were associated with higher probability of dropout. **Meaning:** This study suggests the importance of monitoring clients' changes in outcomes across the phase model of psychotherapy in an effort to reduce early and premature dropout. **Next Steps:** Findings should be replicated and additional ways in which progress may related to dropout should be explored.

Keywords: GEL model, treatment progress, premature termination, psychotherapy, phase model

Monitoring the trajectory and rate of change in psychotherapy significantly impacts treatment processes and outcomes (Barkham, Mellor-Clark, & Stiles, 2015; Miller, Hubble, Chow, & Seidel, 2015). For example, in a special issue of *Psychotherapy*, Miller et

al. (2015) cited evidence suggesting that routine outcome monitoring (ROM) may increase the efficacy of mental health treatments, significantly decrease dropout rates, and shorten treatment lengths. Moreover, ROM may also serve to enhance therapist

This article was published Online First November 26, 2018.

D. Martin Kivlighan III, Department of Psychological and Quantitative Foundations, University of Iowa; Yi-Jiun Lin, Counseling Center, University of Maryland, College Park; Karen P. Egan, Department of Athletics, University of Virginia; Theodore Pickett Jr., Counseling Center, University of Maryland, College Park; Simon B. Goldberg,

Department of Counseling Psychology, University of Wisconsin—Madison.

Correspondence concerning this article should be addressed to D. Martin Kivlighan III, Department of Psychological and Quantitative Foundations, University of Iowa, N334 Lindquist Center, IA City, IA 52242. E-mail: martin-kivlighan@uiowa.edu

development over time and quality improvement at the system level (Fortney et al., 2017). Although much is unknown about how therapists use ROM to inform their practice and enhance the effectiveness of the services delivered, feedback systems operate under the assumption that therapists are better able to identify clients at risk of deterioration with feedback about client progress, or lack thereof. Ultimately, the utility of ROM largely rests on accurate estimations of the trajectory of change over the course of therapy, and identifying clients who are deteriorating compared with these normed trajectories (Shimokawa, Lambert, & Smart, 2010).

Researchers have commonly relied on dose-response studies to identify expected trajectories of change in psychotherapy. In a seminal article testing the dose-effect relationship, Howard, Kopta, Krause, and Orlinsky (1986) found that approximately 50% of clients improved by Session 8, 75% of clients improved by Session 26, and 85% of clients improved by Session 52. In addition, results indicated that the rate of improvement followed a negatively accelerating growth curve, wherein patients experienced greater rates of change early in treatment, which subsequently plateaued in later sessions (Howard et al., 1986). More recently, the good-enough level model (GEL; Barkham et al., 1996) has received increased attention (Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009; Falkenström, Josefsson, Berggren, & Holmqvist, 2016; Goldberg, Hoyt, Nissen-Lie, Nielsen, & Wampold, 2018; Owen, Adelson, Budge, Kopta, & Reese, 2016; Stulz, Lutz, Kopta, Minami, & Saunders, 2013). The GEL model proposes that the negatively accelerating rate of change in treatment (Howard et al., 1986) is an artifact of aggregating across patients who attended various lengths of treatment (Barkham et al., 1996). As the name implies, clients discontinue treatment once they have sufficiently improved to a *good enough level* (Owen et al., 2016). For example, Baldwin et al. (2009) found that the rate of change for 4,676 clients significantly varied as a function of individual treatment length and concluded that “on average patients appear to remain in treatment until they have achieved sufficient change—their GEL—and then terminate treatment” (p. 208). Collectively, findings from these studies suggest that, on average, clients’ progress in treatment follows a negatively accelerating curve, but that the rate of change significantly varies according to treatment length. From a ROM standpoint, having accurate predictions of the expected rate of change is vital for identifying clients who are “off track” or at risk for deterioration. Unfortunately, the majority of dose-response studies have examined the rate and trajectory of change using global measures of general distress, resulting in a lack of information regarding the rates of change for other domains, such as symptom reduction, functioning, or well-being. Whereas the trajectory of change may be similar across diverse domains of change, few studies have tested this assumption.

In fact, we only know of one study that has tested the GEL model across outcome types (Owen et al., 2016). Owen and colleagues (2016) examined variations in the GEL model across outcome type and found evidence of the GEL model for psychological symptoms, well-being, and life functioning. Moreover, they found that the rate of change differed according to outcome type, such that the rate of change in life functioning was less than the rate of change in symptom reduction and well-being (Owen et al., 2016). ROM systems need to model accurate trajectories of change, which may differ across outcome types, in order for

therapists to identify clients at risk of deterioration. Owen et al. (2016) provided valuable information for understanding how clients’ progress may differ according to outcome type; however, Owen et al. (2016) did not account for the dependency between the scales of the Behavioral Health Measure (i.e., psychological symptoms, well-being, and life functioning). When outcome measures contain multiple scales, these data are statistically dependent, and analyses examining multiple domains of change within an outcome measure should account for this level of dependency (Snijders & Bosker, 2012). Multivariate models represent one approach to model multiple dependent variables simultaneously to examine changes in one domain, (e.g., psychological symptoms) while controlling for changes across other domains (e.g., well-being and life functioning). Given the lack of attention to this level of statistical dependency in previous analyses of multiple dependent variables, additional studies are needed.

An additional limitation of the GEL model literature is related to the inconsistent application of inclusion criteria. More specifically, there appear to be meaningful discrepancies or inconsistencies in the inclusion criteria applied across studies of the GEL model. For example, some studies included samples of clients whose baseline scores fell within a clinical distress range (Owen et al., 2016), whereas other studies utilized samples of all clients regardless of pretreatment clinical distress (Falkenström et al., 2016). In addition, some studies included samples of clients who attended a minimum of two sessions (Owen et al., 2016; Reese, Toland, & Hopkins, 2011), and others included samples of clients who attended a minimum of three sessions (Baldwin et al., 2009; Falkenström et al., 2016). Despite this inconsistent application of inclusion criteria, no dose-response study, to our knowledge, has reported comprehensive sensitivity analyses of their main findings. Given these limitations in the GEL model literature, a primary aim of this study was to test for variations in the GEL model according to outcome type while controlling for the dependency between scales and perform sensitivity analyses of our main findings.

Treatment Progress and Premature Termination

In addition to alerting therapists about clients who may be at risk of deterioration, routinely monitoring clients’ progress may prove beneficial in alerting therapists to clients who may be at risk of prematurely terminating treatment. Unfortunately, few studies have examined the relationship between client progress, treatment length, and termination status. The GEL model posits that clients terminate treatment once they have reached a sufficient GEL. However, Stulz et al. (2013) found considerable variability in clients’ individual slopes, especially for those who terminated treatment early, and suggested that “some clients may have terminated early because they felt better, whereas others may have terminated early because therapy was not working well or for other reasons” (p. 599). In other words, clients may decide to prematurely terminate treatment based either on their progress or lack thereof in therapy. Recently, Swift and Greenberg (2015) delineated several strategies for engaging clients in psychotherapy and reducing rates of premature termination. Swift and Greenberg argued clients “may be at an increased risk of prematurely terminating if they expect immediate and global changes and they mistake initial improvements as complete recovery from all of

their symptoms and problems” (p. 174, 2015). As such, they encourage therapists to educate clients on the phase model of psychotherapy, which theorizes that changes in psychotherapy occur in three successive stages: remoralization, remediation, and rehabilitation (Howard, Lueger, Maling, & Martinovich, 1993). Clients experience (a) improvement in their well-being in early sessions, followed by (b) psychological symptom reduction, and then (c) increased functioning in later sessions. Swift and Greenberg (2015) suggested that clients may prematurely discontinue treatment following early gains in well-being, but prior to full recovery, or reductions in symptoms and increased functioning in later sessions. This suggests that clients may decide to continue treatment, or not, based on their progress across multiple domains of change.

A small body of research has found that clients who experience lower rates of change are more likely to prematurely terminate therapy (Kegel & Flückiger, 2015; Mohl, Martinez, Ticknor, Huang, & Cordell, 1991). For example, Kegel and Flückiger (2015) found that clients’ treatment progress was related to their termination status, such that clients who experienced less change in self-esteem over the course of treatment were more likely to prematurely terminate psychotherapy. However, no study has tested if clients’ progress in well-being, psychological symptoms, or life functioning is related to their termination status as proposed by Swift and Greenberg (2015). In addition, it is unclear if the relationship between treatment progress and termination status differs according to treatment length. Whereas clients may terminate treatment once they reach their GEL, it may also be that clients prematurely terminate treatment when they fail to reach their GEL by a certain point in therapy. Moreover, clients may prematurely terminate therapy based on the progression or lack thereof in one domain of change (e.g., well-being) without consideration of additional change domains (e.g., psychological symptoms and life functioning). Therefore, a second aim of this study was to examine the relationship between treatment progress, treatment length, and premature termination.

Present Study

The purpose of this study was twofold. First, we tested the GEL model across outcome domains while controlling for the statistical dependency between scales. More specifically, we applied a multivariate multilevel model to test the GEL model for psychological symptoms, well-being, and life functioning simultaneously. In addition, we tested the impact of differing inclusion criterion on our main findings by applying various inclusion criteria to our sample. We conducted our analyses with four samples: all clients who attended at least two sessions, clinically distressed clients who attended at least two sessions, all clients who attended at least three sessions, and clinically distressed clients who attended at least three sessions. For simplicity and consistency in the presentation of our main findings, our primary analyses were conducted with the sample of all clients (regardless of pretreatment clinical distress) who attended at least two sessions. Sensitivity analyses were conducted with the remaining three samples and compared with the findings from this primary sample. Based on Owen et al. (2016), we hypothesized that the GEL model would be sup-

ported for psychological symptoms, well-being, and life functioning, after controlling for the dependency of data between scales. Similarly, the GEL model has been supported with samples derived from the application of various inclusion criteria, and, as such, we hypothesize that the GEL model will be supported regardless of the inclusion criteria applied to the data.

Second, we examined the relationship between treatment progress, treatment length, and termination status. Informed by preliminary evidence of the association between the rate of change in psychotherapy and premature termination (Kegel & Flückiger, 2015), the current study modeled clients’ changes in well-being, psychological symptoms, and life functioning; treatment length; and an interaction between changes in these three outcomes and treatment length as predictors of therapist-reported termination status (i.e., mutual termination or premature termination). Given the limited research on the impact of clients’ progress in accordance to the phase model and termination status over the course of treatment, these analyses were exploratory in nature, and therefore, no a priori hypotheses were made.

Method

Participants

Participants were 786 clients who received individual therapy from a university counseling center at a large mid-Atlantic university in the United States from 2011 to 2014. Our sample was reduced by excluding clients who received more than one treatment episode between 2011 and 2014, attended less than two therapy sessions (Owen et al., 2016; Reese et al., 2011), or experienced a significant gap in treatment (i.e., 90 days or greater; Baldwin et al., 2009; Falkenström et al., 2016; Goldberg et al., 2018; Reese et al., 2011; Stulz et al., 2013). Applying these criteria, we identified 438 clients for the current study. The majority of clients were female (62.3%; $n = 273$), and 37.7% ($n = 165$) were male. Client age ranged from 18 to 48 years, and the average age was 22.82 ($SD = 4.05$) years. In regards to race, clients self-identified the following way: White 63.7% ($n = 279$), Asian/Native Hawaiian/other Pacific Islander 14.38% ($n = 63$), Black/African American 11.19% ($n = 49$), Biracial 5.48% ($n = 24$), American Indian/Alaska Native 0.23% ($n = 1$), and 5.02% did not identify their race ($n = 22$). The majority of the sample were seniors (43.2%; $n = 189$), followed by graduate students (20.3%; $n = 89$), juniors (18.3%; $n = 80$), sophomores (11.9%; $n = 52$), freshman (4.3%; $n = 19$), and lastly “other” (0.19%; $n = 8$). The mean number of therapy sessions attended was 5.54 ($SD = 2.68$), with an overall range of two to 20 sessions. Lastly, 57 therapists were included in the current sample. Therapists saw 7.68 clients on average ($SD = 6.84$), with a range of one to 25 clients per therapist. Therapists consisted of licensed psychologists, doctoral trainees, and predoctoral interns, and endorsed a variety of theoretical orientations.

Measures

Behavioral Health Measure–20. The Behavioral Health Measure–20 (BHM-20) is a 21-item self-report measure of global mental health (Kopta & Lowry, 2002). In addition to assessing

global mental health, the BHM-20 assesses three domains of mental health: well-being, psychological symptoms, and life functioning (Kopta & Lowry, 2002). The Well-Being scale consists of three items that assess levels of motivation, distress, and life satisfaction. The Psychological Symptoms scale consists of 13 items that assess a range of disorder-specific symptoms, such as depression, anxiety, and substance use. Lastly, the Life Functioning scale consists of four items that assess interpersonal and work/school levels of functioning. Items are assessed on a 4-point Likert-type scale, with higher scores indicating greater levels of mental health. Previous studies have found moderate-to-strong psychometric properties of the BHM-20 (Kopta & Lowry, 2002). In the current study, Cronbach's α s were .80, .88, and .70 for well-being, psychological symptoms, and life functioning, respectively. Clients completed the BHM-20 prior to each therapy session via a computer-based system.

Premature termination. Premature termination is ill-defined in the literature; however, Swift, Greenberg, Whipple, and Kominiaik (2012) suggested that premature termination occurs when a client discontinues treatment unilaterally, without their therapist's agreement and/or without discussing the decision with their therapist. In accord with this definition, premature termination was determined by therapist report of client reason for termination. At the participating site, therapists recorded the reason for client termination at the end of each treatment episode. Clients were included in the analysis of termination status, if they mutually or prematurely terminated treatment, per therapist report. This variable was subsequently dummy coded (0 = *mutual termination*, 1 = *premature termination*).

Data Analysis

Hierarchical linear modeling (Raudenbush, Bryk, Cheong & Congdon, du Toiy, 2011) was used to estimate multivariate multilevel models of the GEL model for clients' well-being, psychological symptoms, and life functioning. Due to the nested nature of the data and multivariate structure of the Behavioral Health Measure (BHM) scales, we created a four-level model, where well-being, psychological symptoms, and life functioning scales were nested within the BHM assessment, sessions were nested within clients, and clients were nested within therapists. Informed by prior research (Baldwin et al., 2009; Falkenström et al., 2016; Owen et al., 2016; Reese et al., 2011), growth models were estimated with a linear, log-linear, quadratic, and cubic growth terms; however, our models would not converge with a cubic growth term, and therefore we opted to compare fit statistics from the linear, log-linear, and quadratic growth models. To identify the growth model of best fit, we utilized Akaike information criterion (AIC) and Bayesian information criteria (BIC). Multivariate models included growth terms for well-being, psychological symptoms, and life functioning simultaneously to control for changes in all three BHM scales within the same model. Following these analyses, we tested the GEL model by entering an interaction term between number of sessions and the log-linear growth in life functioning, psychological symptoms, and well-being. Similar to the tests of model fit comparison, we used a multivariate approach to test the GEL model with all three outcome types

within the same model. The multivariate GEL model was as follows:

Level-1 model:

$$\begin{aligned} BHM_{ijkl} = & \pi_{1jkl} \times (Life\ Functioning_{ijkl}) \\ & + \pi_{2jkl} \times (Psychological\ Symptoms_{ijkl}) \\ & + \pi_{3jkl} \times (Well-being_{ijkl}) + e_{ijkl} \end{aligned}$$

where BHM_{ijkl} is the average BHM score for scale (i ; Life Functioning, Psychological Symptoms, or Well-Being), at session (j), for patient (k), seen by therapist (l). Intercept π_{1jkl} is the within-BHM effect of life functioning on BHM after controlling for the within-BHM effect of psychological symptoms and well-being on BHM, π_{2jkl} is the within-BHM effect of psychological symptoms on BHM after controlling for the within-BHM effect of life functioning and well-being, and π_{3jkl} is the within-BHM effect of well-being on BHM after controlling for the within-BHM effect of life functioning and psychological symptoms. Lastly, e_{ijkl} is the random error.

Level-2 model:

$$\begin{aligned} \pi_{1jkl} = & \beta_{10kl} + \beta_{11kl} \times (Session_{jk}) + r_{1jkl} \\ \pi_{2jkl} = & \beta_{20kl} + \beta_{21kl} \times (Session_{jk}) + r_{2jkl} \\ \pi_{3jkl} = & \beta_{30kl} + \beta_{31kl} \times (Session_{jk}) + r_{3jkl} \end{aligned}$$

where the fixed intercepts (β_{10kl} , β_{20kl} , and β_{30kl}) represent the overall mean for life functioning, psychological symptoms, and well-being, respectively, at baseline for all clients across all therapists. The fixed slopes (β_{11kl} , β_{21kl} , β_{31kl}) represent the overall mean log-linear change in life functioning, psychological symptoms, and well-being, respectively, for each session across all clients and all therapists. Last, r_{1jkl} , r_{2jkl} , and r_{3jkl} represent the error of prediction for life functioning, psychological symptoms, and well-being at session (j), of a given patient (k), seen by a given therapist (l).

Level-3 model:

$$\begin{aligned} \beta_{10kl} = & \gamma_{100l} + \gamma_{101l} \times (\#Sessions_{kl}) + u_{10kl} \\ \beta_{11kl} = & \gamma_{110l} + \gamma_{111l} \times (\#Sessions_{kl}) + u_{11kl} \\ \beta_{20kl} = & \gamma_{200l} + \gamma_{201l} \times (\#Sessions_{kl}) + u_{20kl} \\ \beta_{21kl} = & \gamma_{210l} + \gamma_{211l} \times (\#Sessions_{kl}) + u_{21kl} \\ \beta_{30kl} = & \gamma_{300l} + \gamma_{301l} \times (\#Sessions_{kl}) + u_{30kl} \\ \beta_{31kl} = & \gamma_{310l} + \gamma_{311l} \times (\#Sessions_{kl}) + u_{31kl} \end{aligned}$$

where the fixed intercepts (γ_{100l} , γ_{200l} , and γ_{300l}) represent the overall mean for life functioning, psychological symptoms, and well-being, respectively, at baseline across all therapists. The fixed intercepts (γ_{110l} , γ_{210l} , and γ_{310l}) represent the overall mean change in life functioning, psychological symptoms, and well-being, respectively, across all therapists. The fixed slopes (γ_{101l} , γ_{201l} , and γ_{301l}) represent the effect of the number of sessions on overall mean in life functioning, psychological symptoms, and well-being, respectively, at baseline across all therapists. The fixed slopes (γ_{111l} , γ_{211l} , and γ_{311l}) represent the effect of the number of sessions on the overall mean change in life functioning, psychological symptoms, and well-being, respectively, at baseline across all therapists. Last, the error terms (u_{10kl} , u_{11kl} , u_{20kl} , u_{21kl} , u_{30kl} , and u_{31kl}) represent the error of prediction for life functioning, psychological symptoms, and well-being, respectively, of a given patient (k), seen by a given therapist (l).

Level-4 model:

$$\begin{aligned}\gamma_{100i} &= \delta_{1000} + v_{100i} \\ \gamma_{101i} &= \delta_{1010} + v_{101i} \\ \gamma_{110i} &= \delta_{1100} + v_{110i} \\ \gamma_{111i} &= \delta_{1110} + v_{111i} \\ \gamma_{200i} &= \delta_{2000} + v_{200i} \\ \gamma_{201i} &= \delta_{2010} + v_{201i} \\ \gamma_{210i} &= \delta_{2100} + v_{210i} \\ \gamma_{211i} &= \delta_{2110} + v_{211i} \\ \gamma_{300i} &= \delta_{3000} + v_{300i} \\ \gamma_{301i} &= \delta_{3010} + v_{301i} \\ \gamma_{310i} &= \delta_{3100} + v_{310i} \\ \gamma_{311i} &= \delta_{3110} + v_{311i}\end{aligned}$$

where v_{100i} , v_{200i} , and v_{300i} represent therapist variability around the fixed intercept for life functioning, psychological symptoms, and well-being, respectively. v_{101i} , v_{201i} , and v_{301i} represent therapist variability for the effect of the number of sessions on the overall baseline mean for life functioning, psychological symptoms, and well-being, respectively. v_{110i} , v_{210i} , and v_{310i} represent therapist variability around the slope for life functioning, psychological symptoms, and well-being, respectively. Lastly, v_{111i} , v_{211i} , and v_{311i} represent therapist variability for the effect of number of sessions on the slope of life functioning, psychological symptoms, and well-being, respectively.

Next, we tested treatment progress and treatment length as predictors of termination status with our primary sample. To do so, we first ran separate log-linear growth models for well-being, psychological symptoms, and life functioning and retained the unique client-level intercept and slope residuals from these models. The values reflected clients' overall levels on the BHM subscales (i.e., intercept residuals), as well as their change trajectories (i.e., slope residuals). Lastly, we constructed a two-level (i.e., clients nested within therapists) logistics model with the retained well-being, psychological symptoms, and life functioning intercept and slope coefficients, treatment length, and an interaction between treatment length and the well-being, life functioning, and psychological symptoms slopes as predictors of the probability of premature termination. Random effects for all parameters were included in the model. The model testing progress and treatment length as predictors of termination status was as follows:

Level-1 model:

$$\begin{aligned}\text{Prob}(\text{Termination}_{ij} = 1 | \beta_j) &= \phi_{ij} \\ \log[\phi_{ij}/(1 - \phi_{ij})] &= \eta_{ij} \\ \eta_{ij} &= \beta_{0j} + \beta_{1j} \times (\# \text{Sessions}_{ij}) + \beta_{2j} \times (\text{LF Intercept}_{ij}) \\ &+ \beta_{3j} \times (\text{LF Slope}_{ij}) + \beta_{4j} \times (\text{Symptom Intercept}_{ij}) \\ &+ \beta_{5j} \times (\text{Symptom Slope}_{ij}) + \beta_{6j} \times (\text{WB Intercept}_{ij}) \\ &+ \beta_{7j} \times (\text{WB Slope}_{ij}) + \beta_{8j} \times (\text{LF Slope} \times \# \text{Sessions}_{ij}) \\ &+ \beta_{9j} \times (\text{Symptom Slope} \times \# \text{Sessions}_{ij}) \\ &+ \beta_{10j} \times (\text{WB Slope} \times \# \text{Sessions}_{ij})\end{aligned}$$

where η_{ij} represents the probability of premature termination for a given patient (i) seen by a given therapist (j). The fixed intercept (β_{0j}) reflects the overall mean probability of premature termination across all therapists. β_{1j} represents the relationship between treatment length and the probability of premature termination across all therapists; β_{2j} , β_{4j} , and β_{6j} represent the relationship between baseline scores for life functioning, psychological symptoms, and

well-being, respectively, and the probability of premature termination across all therapists; β_{3j} , β_{5j} , and β_{7j} represent the relationship between the log-linear slope in life functioning, psychological symptoms, and well-being, respectively, and the probability of premature termination across all therapists; and β_{8j} , β_{9j} , and β_{10j} represent the relationship between treatment length, the log-linear slope in life functioning, psychological symptoms, and well-being, respectively, and the probability of premature termination across all therapists.

Level-2 model:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \\ \beta_{2j} &= \gamma_{20} + u_{2j} \\ \beta_{3j} &= \gamma_{30} + u_{3j} \\ \beta_{4j} &= \gamma_{40} + u_{4j} \\ \beta_{5j} &= \gamma_{50} + u_{5j} \\ \beta_{6j} &= \gamma_{60} + u_{6j} \\ \beta_{7j} &= \gamma_{70} + u_{7j} \\ \beta_{8j} &= \gamma_{80} + u_{8j} \\ \beta_{9j} &= \gamma_{90} + u_{9j} \\ \beta_{10j} &= \gamma_{100} + u_{10j}\end{aligned}$$

where u_{0j} reflects therapist variability around the fixed intercept; u_{1j} reflects therapist variability around the slope for treatment length; u_{2j} , u_{4j} , and u_{6j} reflect therapist variability around the slope for life functioning, psychological symptoms, and well-being at baseline, respectively; u_{3j} , u_{5j} , and u_{7j} reflect therapist variability around the slope for log-linear changes in life functioning, psychological symptoms, and well-being, respectively; and u_{8j} , u_{9j} , and u_{10j} reflect therapist variability around the slope for the relationship between treatment length, log-linear changes in life functioning, psychological symptoms, and well-being, respectively, and the probability of premature termination.

Results

The average pretreatment score was 1.92 ($SD = .76$) for well-being, 2.93 ($SD = .58$) for psychological symptoms, and 2.21 ($SD = .70$) for life functioning. The average posttreatment score was 2.33 ($SD = .80$) for well-being, 3.20 ($SD = .55$) for psychological symptoms, and 2.46 ($SD = .79$) for life functioning. Well-being and psychological symptoms were correlated at .75 ($p < .001$), well-being and life functioning were correlated at .75 ($p < .001$), and psychological symptoms and life functioning were correlated at .65 ($p < .001$). To examine standardized pre-post changes in life functioning, well-being, and psychological symptoms, we calculated Cohen's d . According to Cohen (1988), the pre-post effect size for well-being ($d = 0.54$) and psychological symptoms ($d = 0.47$) was medium, and the pre-post effect size for life functioning ($d = 0.36$) was small. In regards to termination status, 26% of the clients in the sample prematurely terminated treatment.

Model fit comparisons indicated that a log-linear growth term was the best fitting model for well-being, psychological symptoms, and life functioning. Further sensitivity analyses indicated that a log-linear growth term was the best fitting model for all samples except the sample of all clients who attended a minimum of three sessions, wherein a linear growth term was

Table 1
Multivariate Multilevel Analysis of the Good-Enough Level Model for Well-Being, Psychological Symptoms, and Life Functioning

Predictor	Estimate	SE	95% confidence interval	T ratio	df	p
Well-being						
Intercept, γ_{100}	1.9	.04	[1.81, 1.99]	42.51	672	<.001
#Sessions, γ_{101}	-0.01	.02	[-0.04, 0.02]	-0.7	672	.482
Session, γ_{110}	0.25	.03	[0.19, 0.30]	8.1	672	<.001
#Sessions, γ_{111}	-0.01	.01	[-0.03, 0.002]	-1.71	672	.088
Psychological Symptoms						
Intercept, γ_{200}	3.09	.04	[3.01, 3.18]	75.51	672	<.001
#Sessions, γ_{201}	-0.01	.01	[-0.03, 0.02]	-0.42	672	.676
Session, γ_{210}	-0.01	.03	[-0.07, 0.04]	-0.45	672	.656
#Sessions, γ_{211}	0.01	.01	[-0.01, 0.02]	1.05	672	.294
Life Functioning						
Intercept, γ_{300}	2.20	.04	[2.11, 2.28]	52.23	672	<.001
#Sessions, γ_{301}	-0.01	.01	[-0.04, 0.02]	-0.61	672	.544
Session, γ_{310}	0.15	.03	[0.09, 0.21]	4.81	672	<.001
#Sessions, γ_{311}	-0.01	.01	[-0.03, 0.008]	-1.06	672	.289

Note. #Sessions = treatment length mean-centered; Session = log-linear growth term.

the best fitting model. Specifically, a log-linear growth pattern was the best fit for samples with all clients attending two or more sessions (AIC = 35,504.30; BIC = 35,728.83), clinically distressed clients attending two or more sessions (AIC = 9864.77; BIC = 10,019.44), and clinically distressed clients attending three or more sessions (AIC = 9190.59; BIC = 9334.43); however, a linear growth pattern was the best fit for the sample of all clients attending three or more sessions (AIC = 33,004.67; BIC = 3320.65).

Tests of the GEL Model

To address Study Aim 1, we tested the GEL model across outcome type and inclusion criteria. Our hypothesis—that the GEL model would be supported for psychological symptoms, well-being, and life functioning, after controlling for the dependency of data between scales—was not supported (Table 1 and Figure 1). Specifically, log-linear changes in well-being, psychological symptoms, or life functioning did not significantly vary as a

function of treatment length (all $ps > .05$). Level-4 random effects were nonsignificant for all parameters. Findings from the analyses of the GEL model did not significantly differ according to the inclusion criteria applied to our data set. Specifically, we did not find support for the GEL model for changes in well-being, life functioning, and psychological symptoms across all four samples.¹ Given the lack of support for the GEL model across outcome types, we ran a post-hoc analysis of the GEL model with the BHM global distress score (i.e., total BHM score). Results from this analysis provided support for the GEL model for clients' global psychological distress. Specifically, the rate of change in clients' global distress significantly varied as a function of the number of sessions attended ($\gamma_{110} = -0.01$, $SE = 0.01$, 95% confidence interval [CI] [-0.02, -0.0004], $t(56) = -2.33$, $p = .024$). We also ran a post-hoc analysis of the GEL model for well-being, life functioning, and psychological symptoms without controlling for dependency between scales to compare our findings to Owen et al. (2016). Results from these analyses provided support for the GEL model for well-being ($\gamma_{110} = -0.02$, $SE = 0.01$, 95% CI [-0.03, -0.002], $t(56) = -2.27$, $p = .027$), and psychological symptoms ($\gamma_{110} = -0.01$, $SE = 0.004$, 95% CI [-0.02, -0.0002], $t(56) = -2.00$, $p = .05$), but not life functioning ($\gamma_{110} = -0.01$, $SE = 0.01$, 95% CI [-0.02, 0.001], $t(56) = -1.85$, $p = .069$).

Tests of Treatment Progress, Treatment Length, and Premature Termination

To address Study Aim 2, we tested clients' log-linear changes in well-being, psychological symptoms, and life functioning; treatment length; and an interaction between changes in all three outcomes and treatment length as predictors of the probability of

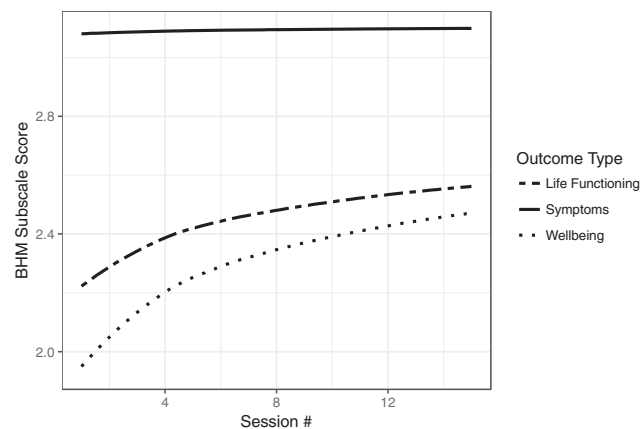


Figure 1. Multivariate analysis of log-linear changes in well-being, psychological symptoms, and life functioning.

¹ The GEL model including all random effects would not converge with the samples of clinically distressed clients who attended a minimum of two or three sessions. Therefore, these analyses were run with models including only random effects for changes in well-being, psychological symptoms, and life functioning.

Table 2
Multilevel Analysis of Treatment Length, Rates of Change, and Interaction Terms Between Treatment Length and Rates of Change Predicting Premature Termination

Predictor	Coefficient	Odds Ratio	95% confidence interval	<i>p</i>
Termination, γ_{00}	-2.51	0.08	[.05, 0.14]	<.001
#Sessions, γ_{10}	-0.16	0.85	[.72, 1.01]	.068
WB Pre, γ_{20}	0.42	1.53	[.44, 5.29]	.498
WB Slope, γ_{30}	7.97	2,888.04	[.61, 13,671,625.77]	.064
SYM Pre, γ_{40}	-1.18	0.31	[.08, 1.24]	.096
SYM Slope, γ_{50}	-0.29	0.75	[.00, 42,255.20]	.958
LF Pre, γ_{60}	0.27	1.31	[.54, 3.14]	.544
LF Slope, γ_{70}	-4.50	0.01	[.00, 14.51]	.214
WB Slope \times #Sessions, γ_{80}	-1.57	0.21	[.06, 0.70]	.012
SYM Slope \times #Sessions, γ_{90}	-0.17	0.84	[.19, 3.72]	.819
LF Slope \times #Sessions, γ_{100}	0.86	2.37	[.74, 7.61]	.145

Note. WB = well-being; SYM = psychological symptoms; LF = life functioning; #Sessions = treatment length mean-centered; pre = baseline score.

premature termination. This analysis tested whether the link between premature termination and the rate of change in outcome varied depending on the length of treatment. For these analyses, we reduced our sample to only include clients for whom their therapist reported that they prematurely or mutually terminated treatment, resulting in a sample of 369 clients treated by 53 therapists. There was a significant interaction effect between log-linear changes in well-being, treatment length, and the probability of premature termination (Table 2). The odds ratio for this interaction term was 0.21 (95% CI [.06, .70]), meaning that for every one-point change in well-being, the odds of premature termination in early sessions increase by 21%. This interaction effect is illustrated in Figure 2. Level-2 random effects (i.e., therapist effects) were nonsignificant for all parameters.

Discussion

The first aim of this study was to test the GEL model across outcome type and inclusion criteria. Our results failed to provide

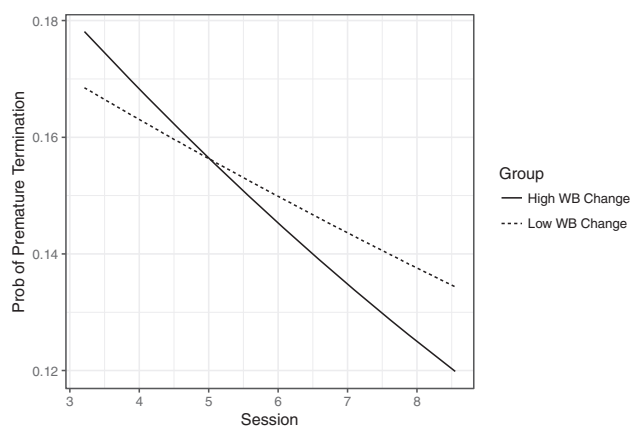


Figure 2. Association between log-linear changes in well-being, treatment length, and the probability of premature termination. “High” well-being change and “Low” well-being change represent 1 *SD* above and below the mean change in well-being, respectively.

support for the GEL model when well-being, psychological symptoms, and life functioning were modeled simultaneously. Specifically, the rate of change in clients’ well-being, psychological symptoms, and life functioning did not significantly vary because a function of the number of sessions attended when the dependency between scales was accounted for. These findings are in contrast to those by Owen et al. (2016), who found that the rate of change in clients’ well-being, psychological symptoms, and life functioning varied as a function of treatment length. There are a variety of possibilities for these mixed findings, namely, differences across studies and samples, for example, differences in the average rate of change, inclusion criteria, average treatment length, sample size, and number of sites included existed between this study and Owen et al. (2016). Another possible explanation is that when dependency between scales is accounted for, the rate of change in well-being, psychological symptoms, and life functioning does not vary as a function of treatment length. To further examine our results in comparison with those by Owen et al. (2016), we ran a post-hoc analysis that did not control for the dependency of scales. Results from this exploratory analysis provided partial support for the GEL model. Specifically, we found evidence for the GEL model for well-being and psychological symptoms, but not life functioning. Together, it appears that when dependency between scales is not accounted for, analyses of the rate of change across outcome domains may produce inaccurate trajectories. Hierarchical modeling is an important consideration in analyses of multiple dependent measures and multiple comparisons in order to control for statistical dependency (Snijders & Bosker, 2012).

As a second exploratory analysis, we conducted a post-hoc analysis of the GEL model for changes in clients’ global psychological distress and found support for the GEL model. Specifically, the rate of change in clients’ global distress significantly varied as a function of the number of sessions attended, such that changes in global distress were greater for clients’ who attended fewer sessions. This study provides preliminary evidence suggesting that the rate of change significantly varies as a function of treatment length for measures of global psychological distress, but not for measures of well-being, psychological symptoms, and life functioning after

controlling for the dependency between these scales. However, it should be noted that this was an unplanned and novel finding that will require independent replication in future research.

In addition, findings from tests of the GEL model did not vary according to the inclusion criteria applied. One concern about the GEL model literature was related to the inconsistent application of inclusion criteria across studies. As such, we attempted to test the effect of applying differing inclusion criteria to our sample on the results from our test of the GEL model. These sensitivity analyses suggested that regardless of the inclusion criteria applied, findings remained consistent. As it is, the inclusion criteria applied (e.g., inclusion of all clients vs. only clinically distressed clients) do not appear to impact findings from tests of the GEL model.

As a second aim of this study, we tested the association between treatment progress, treatment length, and termination status. Results revealed a significant three-way interaction between changes in well-being, treatment length, and termination status. This interaction effect suggests that clients' who experienced rapid rates of change in well-being early in treatment were more likely to prematurely terminate treatment in early sessions. Clients may make decisions about termination related to changes in their well-being and timing, perhaps without considering progress in psychological symptoms or life functioning. This finding provides preliminary support for [Swift and Greenberg's \(2015\)](#) position that clients may prematurely terminate treatment early following increases in well-being but prior to experiencing full remission, such as symptom reduction or changes in life functioning. On the other hand, this finding is in line with the GEL model, such that clients who experienced rapid changes in well-being, prematurely terminated treatment early. However, this finding was not supported for life functioning or psychological symptoms. It should also be noted that this finding is in contrast to the findings of [Kegel and Flückiger \(2015\)](#), who found that the lack of treatment progress was positively related to the probability that a client would prematurely terminate treatment. However, [Kegel and Flückiger \(2015\)](#) measured treatment progress as self-esteem, and the average treatment length between studies greatly differed. Although clinical wisdom and assumptions underlying ROM systems suggest that clients may be at a higher risk of drop out when they experience a lack of progress in treatment, our findings suggest that rapid improvement in clients' well-being early in treatment was associated with premature termination.

There are several implications given our findings. First, this study provides an example and application of a multivariate approach, which underscores the importance of controlling for multiple dependent variables and comparisons. It is not uncommon for psychotherapy researchers to assess multiple domains of change in studies, or analyze multiple subscales of larger outcome measures, which often results in statistical dependency. Analyses should account for dependency of multiple dependent variables and multiple comparisons and apply appropriate models to reduce bias and ensure accuracy in findings.

Clinically, our findings regarding early and rapid changes in clients' well-being and premature termination suggest that therapists should monitor clients' well-being over the course of treatment in an effort to reduce dropout. We echo [Swift and Greenberg's \(2015\)](#) suggestion for therapists to educate their clients on the phase model of psychotherapy and inform them that symptom reduction and changes in life functioning may subsequently follow

earlier changes in well-being. Clinicians may be wise to assess for early and rapid increases in clients' well-being and engage discussions of treatment goals to assess if further change, such as symptom reduction and functioning, is indicated. It may be that early and rapid changes in clients' well-being reflect their experiences of expectancy for change, remoralization, and hope ([Frank & Frank, 1993](#); [Wampold & Imel, 2015](#)) and, although important therapeutic processes, might not represent full remission of symptoms or change in functioning. However, it may also be that clients determine their good-enough level based on their experiences of well-being and may successfully terminate treatment following early and rapid changes in well-being, despite termination being recorded as premature.

As with every study, our findings should be interpreted with an understanding of several limitations. First, compared with previous research examining dose-response relationships ([Baldwin et al., 2009](#); [Howard et al., 1986](#); [Owen et al., 2016](#); [Reese et al., 2011](#)), the current sample size is relatively small ($n = 438$). It may be that our inability to detect a significant interaction between treatment length and the rate of change when simultaneously modeling the subscales of the BHM was related to power. However, we should note that the number of therapists ($n = 57$) at our highest level (i.e., Level 4) is sizable. Second, our sample consisted of a majority of White college-student population receiving brief therapy from a university counseling center. As such, our findings might not be generalizable beyond a university counseling center serving predominately White college students within a brief therapy model. Third, client data, such as diagnosis and presenting concerns, were unavailable. Related, limited therapist data were available, including how therapists utilized the ROM data in practice. Although it was not customary for therapists to review outcome data with their clients at the participating site, some may have done so as part of their own clinical practice. It may be that the inability to blind therapists to their clients' progress in treatment impacted our findings, although research suggests that therapists have a tendency to overestimate their effectiveness ([Tracey, Wampold, Lichtenberg, & Goodyear, 2014](#)) and that signals to therapists about clients who are at risk of deterioration might be ineffective in impacting subsequent treatment response ([Amble, Gude, Ulvenes, Stubdal, & Wampold, 2016](#)). Regardless, a limitation of this study is that termination status and client outcome were inherently dependent, as therapists could not be blind to clients' report of their progress in therapy (even if they could be blinded to the measurement of outcome on the BHM). Of course, this limitation exists for the majority of studies examining the relationship between treatment progress and termination status. Lastly, the BHM-20 was the only outcome measure used to measure well-being, symptoms, and functioning, limiting our assessment of additional outcomes of psychotherapy, such as insight, and potentially more in-depth measures of well-being, psychological symptoms, and life functioning.

Conclusions

Recent attention to the rate and trajectory of change in psychotherapy has increased; however, it is our hope that the current study advances dose-response research in several meaningful ways. As a primary contribution, this is the first study to test, simultaneously, the GEL model across outcome types, to control

for the dependency of multiple outcome domains. Moreover, this study advances the literature on premature termination and suggests that rapid and early improvements in clients' well-being may be a precursor to premature termination. Given the results of this study, it may prove helpful for clinicians to monitor clients' changes in well-being and provide education of the phase model in an effort to reduce premature termination. Future studies are needed to continue to understand the varied ways in which clients' change over the course of treatment and across a variety of domains, as well as how clients' progress over the course of treatment may be related to premature termination.

References

- Amble, I., Gude, T., Ulvenes, P., Stubdal, S., & Wampold, B. E. (2016). How and when feedback works in psychotherapy: Is it the signal? *Psychotherapy Research, 26*, 545–555. <http://dx.doi.org/10.1080/10503307.2015.1053552>
- Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: Contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77*, 203–211. <http://dx.doi.org/10.1037/a0015235>
- Barkham, M., Mellor-Clark, J., & Stiles, W. B. (2015). A CORE approach to progress monitoring and feedback: Enhancing evidence and improving practice. *Psychotherapy, 52*, 402–411. <http://dx.doi.org/10.1037/pst0000030>
- Barkham, M., Rees, A., Shapiro, D. A., Stiles, W. B., Agnew, R. M., Halstead, J., . . . Harrington, V. M. (1996). Outcomes of time-limited psychotherapy in applied settings: Replicating the second Sheffield psychotherapy project. *Journal of Consulting and Clinical Psychology, 64*, 1079–1085. <http://dx.doi.org/10.1037/0022-006X.64.5.1079>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum, Inc.
- Falkenström, F., Josefsson, A., Berggren, T., & Holmqvist, R. (2016). How much therapy is enough? Comparing dose-effect and good-enough models in two different settings. *Psychotherapy, 53*, 130–139. <http://dx.doi.org/10.1037/pst0000039>
- Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). A tipping point for measurement-based care. *Psychiatric Services, 68*, 179–188. <http://dx.doi.org/10.1176/appi.ps.201500439>
- Frank, J. D., & Frank, J. B. (1993). *Persuasion and healing: A comparative study of psychotherapy*. Baltimore, MD: Johns Hopkins University Press.
- Goldberg, S. B., Hoyt, W. T., Nissen-Lie, H. A., Nielsen, S. L., & Wampold, B. E. (2018). Unpacking the therapist effect: Impact of treatment length differs for high-and low-performing therapists. *Psychotherapy Research, 28*, 532–544.
- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41*, 159–164. <http://dx.doi.org/10.1037/0003-066X.41.2.159>
- Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology, 61*, 678–685. <http://dx.doi.org/10.1037/0022-006X.61.4.678>
- Kegel, A. F., & Flückiger, C. (2015). Predicting psychotherapy dropouts: A multilevel approach. *Clinical Psychology and Psychotherapy, 22*, 377–386. <http://dx.doi.org/10.1002/cpp.1899>
- Kopta, S. M., & Lowry, J. L. (2002). Psychometric evaluation of the Behavioral Health Questionnaire-20: A brief instrument for assessing global mental health and the three phases of psychotherapy outcome. *Psychotherapy Research, 12*, 413–426. <http://dx.doi.org/10.1093/ptr/12.4.413>
- Miller, S. D., Hubble, M. A., Chow, D., & Seidel, J. (2015). Beyond measures and monitoring: Realizing the potential of feedback-informed treatment. *Psychotherapy, 52*, 449–457. <http://dx.doi.org/10.1037/pst0000031>
- Mohl, P. C., Martinez, D., Ticknor, C., Huang, M., & Cordell, L. (1991). Early dropouts from psychotherapy. *Journal of Nervous and Mental Disease, 179*, 478–481. <http://dx.doi.org/10.1097/00005053-199108000-00005>
- Owen, J. J., Adelson, J., Budge, S., Kopta, S. M., & Reese, R. J. (2016). Good-enough level and dose-effect models: Variation among outcomes and therapists. *Psychotherapy Research, 26*, 22–30.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & du Toit, M. (2011). HLM7: Hierarchical linear and nonlinear modeling [Computer software manual]. Lincolnwood, IL: Scientific Software International.
- Reese, R. J., Toland, M. D., & Hopkins, N. B. (2011). Replicating and extending the good-enough level model of change: Considering session frequency. *Psychotherapy Research, 21*, 608–619. <http://dx.doi.org/10.1080/10503307.2011.598580>
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology, 78*, 298–311. <http://dx.doi.org/10.1037/a0019247>
- Snijders, T., & Bosker, R. (2012). *Multilevel modeling: An introduction to basic and advanced multilevel modeling*. London, United Kingdom: Sage.
- Stulz, N., Lutz, W., Kopta, S. M., Minami, T., & Saunders, S. (2013). The shape of the dose-effect relationship in outpatient psychotherapy: Does treatment duration matter? *Journal of Counseling Psychology, 60*, 593–600.
- Swift, J. K., & Greenberg, R. P. (2015). *Premature termination in psychotherapy: Strategies for engaging clients and improving outcomes*. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/14469-000>
- Swift, J. K., Greenberg, R. P., Whipple, J. L., & Kominiak, N. (2012). Practice recommendations for reducing premature termination in therapy. *Professional Psychology, Research and Practice, 43*, 379–387. <http://dx.doi.org/10.1037/a0028291>
- Tracey, T. J., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist, 69*, 218–229. <http://dx.doi.org/10.1037/a0035099>
- Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. New York, NY: Routledge.

Received September 16, 2016

Revision received July 16, 2018

Accepted August 8, 2018 ■