Special Issue: Integration of Practice, Advocacy, and Research in Counseling Psychology

The Counseling Psychologist
2021, Vol. 49(7) 1013–1037

DOI: 10.1177/00110000211029271
journals.sagepub.com/home/tcp

Center Effects: Counseling Center Variables as Predictors of Psychotherapy Outcomes

Dever M. Carney[1], Louis G. Castonguay[1], Rebecca A. Janis[1], Brett E. Scofield[1], Jeffrey A. Hayes[1], and Benjamin D. Locke[1]

Abstract

Treatment context may have a unique impact on psychotherapy outcomes, above and beyond client, therapist, and therapy process variables. University counseling centers represent one such treatment context facing increasing treatment demands. This study examined the role of counseling centers and center variables in explaining differences in psychotherapy outcomes. The Center for Collegiate Mental Health, a large practice–research network, contained data from 116 counseling centers, 2,362 therapists, and 58,423 clients. Multilevel modeling tested if some counseling centers systematically achieved better outcomes than others (a "center effect"). Outcome was operationalized as clients' magnitude and rate of change in distress across treatment. Results showed a relatively small "center effect" for both outcomes. Analyses sought to explain that center effect through administrative policies and characteristics. As a group, these variables partially explained the center effect. None explained a large portion of total outcome variance. Potential future implications for policy and advocacy efforts are discussed.

[1]Pennsylvania State University, University Park, PA, USA

Corresponding Author:
Dever M. Carney, Department of Psychology, Pennsylvania State University, 346 Bruce V. Moore Building, University Park, PA 16802, USA.
Email: dfc5168@psu.edu

Keywords
collegiate mental health, psychotherapy outcome, center effects, organizational effects, contextual factors

Significance of the Scholarship to the Public
This study examined the link between college counseling centers, their policies, and student psychotherapy outcomes. Findings indicate that the specific center where someone receives treatment does not explain much of how they do in treatment and that client variables are more predictive of how people change across treatment.

Researchers have spent decades exploring what influences the outcome of therapy, including factors relating to the client (e.g., personality), the therapist (e.g., interpersonal style), type of treatment, and process factors (e.g., therapeutic alliance). Historically, it has been difficult to statistically account for the fact that relationships between these factors are inherently hierarchical and nonindependent. Individual sessions (Level 1) are "nested" within clients (Level 2), who are nested within therapists (Level 3), who are then nested within organizations (Level 4). With the development of multilevel modeling (MLM; Raudenbush & Bryk, 2002), analyses can now parse the proportion of variance in psychotherapy outcomes attributable to each level. Effects specific to the client-, therapist-, and organizational-level can be calculated in terms of intraclass correlation coefficients (ICCs). Although a substantial number of psychotherapy studies have utilized MLM to parse out contributions from client- and therapist-levels, less attention has been given to contextual or organizational factors. An "organizational effect" (or center effect) has been suggested, such that some organizations may be better than others at facilitating therapeutic change, above and beyond the therapists they employ and the clients they see (Falkenström et al., 2018). This may be due to organizational-level factors such as type of therapy setting (e.g., hospital), clinical model in place (e.g., treat or refer), specific organizational policies (e.g., session limits), or staff climate and culture. Researchers have argued that these factors likely account for some of the unexplained variance in therapy outcomes (e.g., Wampold & Imel, 2015) and could thus inform administrative best practices.

Falkenström et al. (2018) reviewed the literature on organizational effects in psychotherapy. Across the 19 studies they examined, every paper found some evidence of an organizational effect, demonstrating that the treatment setting is likely an important source of variance in therapy outcomes. The majority of the studies used MLM, and evidence for organizational differences was observed first through statistically significant differences between clinical agencies. For example, one study found that 12% of client change was attributable to variation across sites (Glisson & Green, 2011). Of all the studies reviewed, none were explicitly designed to test questions around organizational effects. The authors therefore encouraged researchers to directly examine the explanatory power of organizational variables. They also noted that large practice–research networks (PRNs) may be especially well suited to investigating these effects, due to the size and complexity of the data being gathered.

One area that would benefit from this line of organizational research is mental health among college students. Not only is the field inherently divided into many traditionally independent organizations (i.e., campuses) well-suited to MLM, but college and university counseling centers (UCCs) across the United States have been experiencing growing demand in a number of areas (Xiao et al., 2017). According to a large-scale analysis of hundreds of UCCs from 2009–2015, therapy utilization increased on average by 30–40% compared to only a 5% increase in student enrollment (Center for Collegiate Mental Health [CCMH], 2016). Furthermore, a growing proportion of clients presented with threat-to-self indicators, who then utilized about 25% more resources (CCMH, 2016). These issues pressured centers to evaluate and/or change policies to maximize their effectiveness. Using multisite data to examine how organizational-level features impact outcomes has the potential to guide policymakers towards best administrative practices.

Contextual Factors Related to Outcomes

UCCs vary greatly from one to another. They differ, for example, in the type of academic institution (e.g., private or public) they are attached to, and the ethnic composition of their clientele. Although such defining variables could potentially have a relationship with treatment outcome, the UCC is not able to modify them. In the current study, we investigate center variables that not only may predict outcome, but may also be the targets of actionable change to the provision of services. One such policy is the frequency of treatment. A recent study indicated that clients who attended more frequent sessions demonstrated more rapid improvement (Reese et al., 2011). Erekson et al. (2015) demonstrated that more frequent therapy sessions were associated with clients achieving gains more quickly, but not necessarily more change overall. These studies used client-level data, but the results may have center-level implications for how to set treatment frequency policies. A more direct examination of these implications using center-level data could provide additional insight.

UCC policies around session limits may also be related to outcome. Research has shown that the majority of clients require 14 sessions or more to achieve clinically significant change, suggesting that lower session limits in UCCs may not be optimal (Wolgast et al., 2004). On the other hand, findings also show that briefer treatment can be effective, especially for clients who start out with lower levels of distress (Bohart & Wade, 2013). A recent study found that UCCs providing treatment with explicitly defined session limits (especially in the context of a smaller staff) had better outcomes than UCCs with ambiguous or no limits (Coleman et al., 2019). The mixed findings described used client-level data, and merit further investigation at the center level to better inform UCC decision-making on session limits.

The provision of supplementary services to students may also impact their therapy outcomes. For example, a UCC may also offer in-house psychiatric care, academic counseling, career counseling, and/or neuropsychological testing. Based on theories behind integrated care that suggest a more holistic approach can improve outcomes (Hammer et al., 2019), UCCs that address students' needs in multiple domains or collaborate with other providers may foster more positive outcomes. In addition, UCCs make decisions about the number of clients that receive psychotherapy simultaneously. Although this may be contingent on the size of their institution and staff, decisions about number of clients seen simultaneously may affect how clients attend, engage with, and benefit from treatment. One qualitative study interviewed students about what was "unhelpful or hindering to counseling" and students identified "feeling like part of an assembly line" as a barrier to improvement (Paulson et al., 2001). Larger UCCs may engender that type of "assembly line" experience, although to our knowledge no studies have examined the link between number of clients and outcomes.

Another actionable contextual factor that could be linked with outcome is the organization's accreditations. In college counseling, UCCs with training programs can submit applications to become accredited with the American Psychological Association (APA; n.d.) and the International Accreditation of Counseling Services (IACS). The APA accredits UCC training programs for predoctoral psychology interns and postdoctoral fellowships based on their ability to use best educational practices to prepare graduates to provide exceptional, evidence-based services. The IACS accredits the UCC itself based on their services, function within the university, and administrative practices. Because these statuses are recognition of high educational and professional standards in the field, earning one of these accreditation statuses is intended to be a proxy for the highest quality of care. Thus, a center successfully maintaining accreditation may predict more positive outcomes for their clients. To our knowledge, this has yet to be investigated empirically.

Methodological Considerations

In addition to the limited number of studies on organizational factors, most research has failed to delineate the specific part of the outcome variance that is uniquely due to the center or organization. One of the challenges in parsing out outcome variance is acquiring the necessary sample sizes (Kim et al., 2006). This step is crucial, as generating multilevel models that lack sufficient power at any of the levels can lead to biased parameter estimates and inflated rates of Type I errors (Schiefele et al., 2017). So far, very few projects have been able to muster the sample sizes to meet recommended guidelines (for guidance on a priori power analysis see de Jong et al., 2010). This is especially true at the highest level of analysis (organizations), as there are often logistical barriers to recruiting multiple sites (Maas & Hox, 2005). Of the studies that link center-level variables with client outcomes, only a few (e.g., Erekson et al., 2015) used MLM to account for hierarchical data structures, and the majority of these types of studies were conducted within one center.

Current Study

The goals of the current study were to (a) fill gaps in the literature on organizational effects by employing MLM to estimate unique center contributions to client outcomes, and (b) explore actionable center characteristics and policies that could potentially be modified to influence such outcomes. The study set out to accomplish these goals by utilizing a sizeable, representative, and heterogeneous sample of clients and organizations derived from the CCMH, an extensive PRN infrastructure of UCCs (McAleavey et al., 2015). Two research questions were explored: (a) How much of the variance in psychotherapy outcomes was attributable to the center? and (b) Did specific counseling center characteristics or policies help explain this proportion of variance in outcomes? To investigate these questions, the relationship between six center-level factors and therapeutic outcome (measured both in terms of magnitude and rate of symptom change) were examined: session limits, session frequency, APA and IACS accreditation status, counseling center size (number of clients seen annually), and the integration of various services into the treatment setting.

Methods

Participants

Clients. The clients were graduate or undergraduate students presenting for treatment at UCCs across the United States during the 2017–2018 and 2018–2019 academic years. A general data reduction process (see Table 1) excluded

Table 1. Steps for Sample Reduction

Procedure	Client, *n*	Therapist, *n*	Center, *n*
Posttreatment analyses			
Starting sample	334,996	5,454	169
Remove unattended sessions	319,925	5,372	169
Remove unused appointment types	302,994	4,968	169
Remove clients without primary therapist	216,182	4,314	163
Remove clients with <2 appts	149,913	4,201	163
Remove clients with <2 CCAPS	94,017	3,809	163
Remove clients without valid first/last CCAPS	70,616	3,538	159
Remove clients without 1+ CCAPS above low cut	67,329	3,525	159
Remove therapists with <5 clients	65,411	2,690	147
Remove centers without clients in both samples	59,674	2,470	134
Remove centers with <50 clients	59,194	2,402	118
Remove centers without data on predictors	58,423	2,362	116
Rate of change analyses			
Starting sample	334,996	5,454	169
Remove unattended sessions	319,925	5,372	169
Remove unused appointment types	302,994	4,968	169
Remove clients without primary therapist	216,182	4,314	163
Remove clients with <2 appts	149,913	4,201	163
Remove clients with <2 CCAPS	94,017	3,809	163
Remove clients without 2+ linked CCAPS	84,762	3,624	160
Remove clients without CCAPS at baseline	77,438	3,540	160
Remove clients without 1+ CCAPS above low cut	73,848	3,524	160
Remove therapists with <5 clients	72,084	2,721	151
Remove centers without clients in both samples	59,674	2,470	134
Remove centers with <50 clients	59,194	2,402	118
Remove centers without data on predictors	58,423	2,362	116

Note. CCAPS = College Counseling Assessment of Psychological Symptoms.

clients who could not be assigned a primary therapist (see the therapist section next), who attended less than two group or individual therapy sessions, and who had completed less than two symptom questionnaires (to ensure the outcomes could be computed). Data were also removed for clients who did not demonstrate, at baseline, some elevation on the measure of symptoms described below. At the outset, the dataset contained 334,996 clients. The final sample contained 58,423 clients meeting inclusion criteria, with an average age of 21.92 years (range: 18.00–65.39 years), 65.51% female, 32.06% male, and about 2% transgender or self-identifying. Participants were 65.33% White, 9.81% Hispanic/Latino/a, 8.85% African American/Black, 8.64% Asian/Asian American, 5.19% biracial, 1.50% self-identifying, 0.45%

American Indian/Alaskan Native, and 0.22% Native Hawaiian/Pacific Islander. They represented a range of religious, country of origin, and sexual orientation identities.

Therapists. From across all contributing UCCs, therapists in the initial dataset represent treatment providers who were designated in a center's local electronic medical record system. To be linked to a client, counselors were required to be that client's "primary therapist," operationalized as the provider for at least 50% of that client's appointments. Data from therapists with less than five clients meeting the above criteria were removed, ensuring that MLM with therapist as a grouping factor could be accomplished. The original dataset contained 5,454 therapists, and the final sample contained 2,362 (see Table 1). Some therapists completed an optional demographic survey, but over half did not respond. Because of the low response rate, demographic summaries are thought to be unrepresentative and are not reported.

Centers. CCMH members are granted access to proprietary instruments, reporting tools, trainings, national benchmarking data, and datasets for research. Centers may not join due to lack of time, resources, technological capabilities, prioritization of research, or information about CCMH. Centers in this study were CCMH members who contributed data to the national data repository. Centers were removed (see Table 1) if they had fewer than 50 clients with valid data, to ensure representative and reliable data at an organizational level. Centers were also excluded if they did not contribute data on the predictor variables. The original sample contained 169 centers, and the final sample had 116 centers. Centers saw an average of 1,280 clients each year (range: 164–5,371) and employed an average of 27 therapists (range: 2–104). Across centers, the average number of sessions per client was 5.42 (range: 2.80–9.64) with an average of 18.12 days between appointments (range: 9.95–25.19). Thirty-nine percent (n = 45) had session limits, 40% (n = 46) were APA accredited, 53% (n = 61) were IACS accredited, and 78% (n = 91) offered some type of integrated services.

Measures

College Counseling Assessment of Psychological Symptoms (CCAPS). Client symptoms were assessed via the CCAPS, a multidimensional measure specifically designed to assess collegiate mental health concerns (Locke et al., 2011). It demonstrates strong psychometric properties, and a validated short form with 34 of the original 62 items that was created for use as a repeated measurement (Locke et al., 2012). The current study includes data from CCAPS-34 administrations, as well as CCAPS-62s scored as CCAPS-34s, in order to maximize the available data. The CCAPS-34 has seven subscales: Depression, Generalized Anxiety, Social Anxiety, Academic Distress, Eating Concerns, Alcohol Use, and Hostility. Complementing these is the Distress Index, which calculates a general distress score using select items from other subscales. Clients are asked to rate themselves on items relating to the past two weeks. Reports are made on a 5-point Likert-type scale, ranging from 0 (*not at all like me*) to 4 (*extremely like me*). Total subscale scores are the average of all items that load onto a particular subscale. Higher subscale scores indicate more distress. CCAPS low cut-off points were calculated by statistically differentiating between samples of treatment-seeking and nontreatment-seeking students, and scores below the cut-off points (different on each subscale) reflect little to no distress (McAleavey et al., 2012). Data were removed from clients who did not endorse, at baseline, distress above the CCAPS low cut-off point for at least one of the subscales, to ensure that statistically there was room for change. As a check for reliability, internal consistency of the CCAPS within centers in the current sample was calculated. Although a few centers had low internal consistencies on some subscales, the majority of centers displayed acceptable levels. The overall Cronbach's alphas (taking into account all clients from all centers) and range of values within centers are as follows: Depression (overall = .85, range: .78–.89), Generalized Anxiety (.79, .67–.83), Social Anxiety (.80, .70–.86), Academic Distress (.80, .73–.87), Eating Concerns (.88, .78–.94), Hostility (.82, .74–.86), Alcohol Use (.82, .67–.91), and the Distress Index (.88, .85–.91).

Standardized Data Set (SDS). The SDS is a standardized questionnaire created from a collection of UCC intake materials. It contains items for clients to answer at intake regarding their demographics and mental health history, as well as items answered annually by local UCC administrators that assess organizational policies and characteristics.

Procedure

Data were collected through the CCMH PRN. Each UCC secured and maintained approval from their local institutional review board. Participating UCCs collected client-, therapist-, and center-level data in a naturalistic setting via the CCAPS and SDS. Appointment data were also collected via the center's electronic medical record system, and included information such as date, treating therapist, attendance (e.g., attended or no show), and appointment type (e.g., group or individual). All data were then de-identified and contributed to the national CCMH repository.

Statistical Analyses

Analyses aimed to address the research questions sequentially: (a) examination of the effect of center on outcome without any predictors, and (b), addition of predictors to identify which variables account for the variance explained by the center. These steps were conducted in two separate analyses: one with the posttreatment CCAPS score as the measure of outcome (controlling for initial CCAPS score) and the next with rate of symptom change as the outcome.

Center-Level Predictor Variables. Six center-level variables were used to predict outcomes. Four were obtained from SDS items with dichotomous yes or no responses, including IACS and APA accreditations, session limits, and integrated services. Session limits were also categorized as existing or not, regardless of the specific policy. Centers with integrated services provided at least one of the following in their counseling center: career, disability, learning, health, and testing services, drug and alcohol treatment, and employee assistance.

The other two center-level predictors were continuous and calculated from electronic medical record data: center size (operationalized as the number of students served annually, averaged between the two years of data) and session frequency (average number of days between attended appointments, including both group and individual therapy). Higher values of this variable indicate less frequent sessions. For the purposes of this study—interested in routine psychotherapy practices—appointments such as screens, psychiatric, and couple therapy were excluded. An additional center-level variable was not a predictor of interest but was added to help disentangle client- and center-level effects: center-level mean baseline on a given CCAPS subscale (i.e., scores averaged across all of a center's clients).

Outcome as Post-Treatment CCAPS Scores. MLM was used to control for the fact that the data are nonindependent. A series of models were built with three levels: clients within therapists (Level 1), therapists within UCCs (Level 2), and UCCs (Level 3), allowing the examination of unique contributions to outcomes from each source. Analyses were conducted using maximum likelihood estimation with the *nlme* package (Pinheiro et al., 2013) and the *r2glmm* package (Jaeger, 2017) in the R programming language (version 3.5.2; R Development Core Team, 2014).

For these analyses, CCAPS had to be administered within a two-week window around clients' first and last sessions, respectively, to ensure scores truly captured the beginning and end of treatment. First, a client's final score on a given CCAPS subscale was modeled in a multilevel regression with therapist and center included as grouping variables (Model 1). This model also included client baseline CCAPS scores as a control. The client baseline score was centered around the counseling center's mean baseline distress to disentangle the client-level and center-level effects (Raudenbush & Bryk, 2002), thus the effect of client-level initial distress is interpreted in comparison to other clients at the same center. Model 1 also included the center's mean initial score on a CCAPS subscale, which was grand mean centered. This accounts for the fact that clients and centers with lower average initial distress can experience less change statistically due to a lower upper bound.

Model 1 generated ICCs, which determined how variance in client outcome was allocated across client, therapist, and center levels. A generic equation structurally equivalent across subscales was as follows:

$$\text{Client-Level}: [\text{Client last CCAPS}]_{\text{ijc}} = \beta_{0\text{jc}} + \beta_{1\text{jc}}\left(\text{Client first CCAPS}_{\text{ijc}}\right) + e_{\text{ijc}}$$

$$\text{Therapist-Level}: \beta_{0\text{jc}} = \pi_{00\text{c}} + u_{0\text{jc}}$$

$$\text{Center-Level}: \pi_{00\text{c}} = \gamma_{000} + \gamma_{001}\left(\text{Center first CCAPS}_{\text{c}}\right) + u_{00\text{c}}$$

Here, $[\text{Client last CCAPS}]_{\text{ijc}}$ represents the final CCAPS subscale score for client *i* seen by therapist *j* at center *c*. Random intercepts were included at the therapist and center levels ($u_{0\text{jc}}$ and u_{00c}), allowing therapists and centers to have a unique deviation from the average final CCAPS score, in addition to a term of residual variance (e_{ijc}). The variation associated with the center-level divided through the total variance is the "center effect."

Next, the second model examined which UCC policies and characteristics might explain outcome variance attributable to the center (Model 2). This model added the six center-level predictors: session limits, IACS and APA accreditations, center size, session frequency, and integrated services. To aid interpretation of findings, continuous predictors (center size, session frequency, average initial CCAPS) were grand mean centered, while dichotomous predictors (session limits, accreditations, integrated services) were left uncentered. Center size was also rescaled such that increments in one unit represented an additional 100 clients (instead of a single client). Other predictors and outcomes were left in original units to allow coefficients to be more readily interpretable. The effects of these center-level predictors were evaluated after controlling for client and center baseline distress. The resulting intercept represented the predicted final CCAPS score for a client with average initial distress, at an average-sized center with an average session frequency, with no session limits, accreditations, or integrated services. Alpha cutoffs for significance were set to a more stringent $p < 0.01$ to account for the large sample size. Variance attributable to each level of grouping (or the ICC) was calculated as a ratio of specific-level error variance divided by the total error variance from Model 2 (Raudenbush & Bryk, 2002).

Likelihood-ratio tests (LRT) were used to compare model fit and to assess whether adding center-level predictors improved the overall model fit (Bolker et al., 2009). The LRT tested the null hypothesis that there was no significant difference in fit between two nested models and is modeled as a chi-square distribution where the degrees of freedom are equal to the difference in parameters between the two models. We report a generalized R^2 as an approximation of the percent variance explained by each predictor (Edwards et al., 2008). The ICCs (the variance attributable to grouping factors) can be compared across models, as can R^2 values (variance explained by the predictors), but not to each other. Finally, proportional reduction in error (PRE) at the center level was calculated to evaluate the proportion of variance explained at the center level by center-level predictors (Raudenbush & Bryk, 2002). This was calculated for a center's average baseline CCAPS by comparing a model with average center baseline to a model without, as well as for all center-level predictors by comparing a model with all predictors (including center baseline) to a model without any center predictors.

Outcome as Rate of Change Across Treatment. To examine the rate at which a client's distress changed across treatment, we generated longitudinal MLMs with four levels: sessions (Level 1), clients (Level 2), therapists (Level 3), and centers (Level 4). Rate of change analyses followed the iterative model building approach described above (i.e., Models 1 and 2). Here, CCAPS administrations were required to occur within three days of a session to be linked with that appointment. Clients' CCAPS scores were centered around their baseline, anchoring each client's intercept at zero. As such, only fixed and random slopes (but not intercepts) were estimated in the model, similar to methodology described by Lutz et al. (2007). Session numbers were log-transformed to model change in line with the dose-effect model (Hansen et al., 2006). The random effect of session number (slope) allowed slopes to vary by client, therapist, and center. Equations structurally equivalent across subscales are as follows:

$$\text{Session-Level}: [\text{Client CCAPS}]_{\text{tijc}} = \beta_{0\text{ijc}} + \beta_{1\text{ijc}}\left(\text{log session number}_{\text{tijc}}\right) + e_{\text{tijc}}$$

$$\text{Client-Level}: \beta_{0\text{ijc}} = 0 + 1\left(\text{Client first CCAPS}_{\text{ijc}}\right)$$

$$\beta_{1\text{ijc}} = \pi_{1\text{ijc}} + u_{1\text{ijc}}$$

$$\text{Therapist-Level}: \pi_{1ijc} = \gamma_{100c} + u_{10jc}$$

$$\text{Center-Level}: \gamma_{100c} = \theta_{1000} + \theta_{1001}\left(\text{Center first CCAPS}_c\right) + u_{100c}$$

Here, $[\text{Client CCAPS}]_{tijc}$ represents the CCAPS subscale score at session *t* for client *i* seen by therapist *j* at center *c*. Random intercepts for the effect of session number were included at the client, therapist, and center levels (u_{1ijc}, $u_{10jc,}$ and u_{100c}), allowing clients, therapists, and centers to have unique deviation from the average slope, in addition to a term of residual variance at the session level (e_{tijc}).

Similar to posttreatment analyses, continuous predictors were grand mean centered to aid interpretation, while dichotomous predictors were left uncentered. As with the previous model, predictors were left in the original units, except for center size. The resulting slope coefficient represents the predicted change on the CCAPS per session, for a client with average initial distress, at an average-sized UCC with an average session frequency, with no accreditations, limits, or integrated services. Center effects were calculated as the center slope variance divided by the sum of center, therapist, and client slope variance (Lutz et al., 2007), representing the amount of variance in client rate of change that is attributable to centers. Similar to post-treatment analyses, we report ICCs and LRT values, but due to computational limitations were unable to calculate generalized R^2 effect sizes for the four-level longitudinal models. Finally, PRE in slopes was calculated similarly to the post-treatment models, comparing models with and without center baseline, as well as models with and without any center predictors.

Power

Calculating power a priori for MLM, especially two- and three-level models, is complex: dependent not only on overall sample size, but also on the sample, ICC, and covariates at each level. Schiefele et al. (2017) developed recommendations for variable configurations of units at each level, while focusing on total sample minimums necessary to obtain accurate estimates of therapist (Level 3) effects. For example, they found that 100 therapists seeing 15 patients each produced sufficient power, while also suggesting that number of clients and therapists can vary as long as the overall sample has at least 1500 clients. Other sources recommend total sample sizes of at least 1,000 (Hox & Maas, 2010; Raudenbush & Bryk, 2002). Similarly, prior research on three-level models suggests that in order to achieve sufficient power, 22 therapists seeing eight patients each are needed, while more therapists are needed if there are less therapists per patient (de Jong et al., 2010). Minimum sample sizes were reduced when randomization was used. Applying these guidelines to the present study, in which clients, therapists, and centers were not randomized, the subscale with the lowest final sample size (Alcohol Use) had 18,901 clients seen by 2,261 therapists at 116 centers. This results in an average of eight clients per therapist, and 19 therapists per center. Based on the previously outlined guidelines for power in MLM, this study is sufficiently powered at the highest level (centers) to detect accurate center-level effects and to test effects of center-level characteristics.

Results

Posttreatment CCAPS Scores

Center ICCs, or the variance in outcome accounted for by center (i.e., where the client was seen), are presented in Table 2. Centers accounted for an average of 1.93% of the differences in final CCAPS scores, after controlling for baseline. On specific subscales, centers explained between 1.32% (Eating Concerns) and 2.61% (Alcohol Use) of the differences in final CCAPS scores. By comparison, variance accounted for by the therapist ranged from 0% (Eating Concerns) to 2.33% (Distress Index) and variance accounted for by the client ranged from 95.29% (Distress Index) to 98.68% (Eating Concerns).

Parameter estimates and fit statistics from Model 2 for all CCAPS subscales are reported in Table 2. Negative beta values indicate lower posttreatment scores (e.g., positive outcomes by ending treatment in less distress). Across subscales, the pattern of explanatory significance was similar. On nearly all subscales, baseline distress for clients and center averages were significant, such that higher symptomatology at baseline predicted higher final CCAPS scores. Session frequency significantly predicted outcome on four subscales (Depression, Anxiety, Academic Distress, and Distress Index), such that more frequent sessions were associated with lower final CCAPS scores. For example, for every day fewer between appointments (i.e., for more frequent sessions closer together), clients at the center ended treatment .01 points lower on Depression. APA accreditation was significantly related to outcomes on only the Distress Index, such that receiving treatment at an APA accredited UCC was predictive of better outcomes (i.e., lower final scores). Session limits, IACS accreditation, center size, and integrated services were not significantly related to outcomes on any subscales.

The PREs for models with center-level baseline ranged from .11 (Eating Concerns and Alcohol Use) to .28 (Generalized Anxiety), indicating that average center baseline scores explained between 11% and 28% of center level variance in post-treatment scores. Proportional reduction in center-level

Table 2. Results From Multilevel Models Predicting Posttreatment Symptom Scores on the CCAPS

Value	Depr.	Anx.	Soc.Anx.	Academ.	Eating	Hostility	Alcohol	DI
Client ICC	96.37%	96.61%	97.22%	97.16%	98.68%	95.74%	95.86%	95.29%
Therapist ICC	1.79%	1.56%	1.17%	1.36%	0.00%	1.85%	1.53%	2.33%
Center ICC	1.84%	1.82%	1.61%	1.48%	1.32%	2.41%	2.61%	2.37%
Model R^2	22.94%	26.48%	33.21%	19.55%	26.01%	23.35%	25.72%	26.04%
Baseline PRE	0.21	0.28	0.22	0.18	0.11	0.20	0.11	0.21
Predictor PRE	0.33	0.43	0.32	0.34	0.15	0.29	0.18	0.37
LRT	15.98	20.73*	11.37	16.68	5.16	10.42	6.71	21.11*
Intercept	0.01	0.01	0.00	0.02	0.00	0.01	−0.01	0.01
SE	0.03	0.03	0.02	0.03	0.04	0.03	0.04	0.03
Client baseline	0.57**	0.64**	0.76**	0.6**	0.7**	0.54**	0.56**	0.65**
SE	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Center baseline	0.47**	0.7**	0.69**	0.67**	0.57**	0.51**	0.55**	0.58**
SE	0.09	0.12	0.14	0.13	0.15	0.11	0.14	0.11
Session limits	0.02	0.02	0.03	-0.02	0.03	0.03	0.04	0.02
SE	0.02	0.02	0.02	0.02	0.03	0.02	0.03	0.02
Session freq.	0.01*	0.01*	0.01	0.01*	0.01	0.00	0.00	0.01*
SE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IACS Accred.	0.03	0.03	0.01	0.00	0.01	0.04	0.03	0.02
SE	0.02	0.02	0.02	0.02	0.03	0.02	0.03	0.02
APA Accred.	−0.06	−0.06	−0.04	−0.07	−0.02	−0.06	−0.06	−0.07*
SE	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.02
Center size	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Integr. Services	−0.02	0.00	0.01	−0.01	0.01	0.01	0.01	0.00
SE	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.02
Clients, *n*	46,901	47,084	39,249	40,866	21,008	27,458	18,901	46,503
Therapists, *n*	2,361	2,361	2,358	2,358	2,299	2,337	2,261	2,361
Centers, *n*	116	116	116	116	116	116	116	116

Note. Intraclass correlation coefficients (ICCs) are reported from Model 1 (controlling for client and center baseline distress), representing the percentage of variance attributable to each level of grouping (client, therapist, center). Effect sizes (R^2) are reported from Model 2, representing the percentage of variance in outcome accounted for by all predictors. Proportional reduction in error (PRE) represents the proportion of center variance accounted for by center baseline CCAPS and all center predictors. Likelihood ratio tests (LRT) tested whether model fit improved between Models 1 and 2. Beta coefficients for parameters are reported. CCAPS = College Counseling Assessment of Psychological Symptoms; DI = Distress Index; SE = standard error.
*$p < .01$. **$p < .001$.

variance with the addition of all center-level predictors (including average baseline) ranged from .15 (Eating Concerns) to .43 (Generalized Anxiety). The percent of overall outcome variance accounted for by all predictors in Model 2 for each subscale was calculated via a generalized R^2 (see Table 2). The overall model predicted between 19.55% (Academic Distress) and

33.21% (Social Anxiety) of the variance in posttreatment scores but we observed that client baseline distress accounted for the vast majority of the percent variance explained by the overall model. For example, the overall Model 2 for Depression explained a total of 22.94% of the variance in outcome, with 22.46% of that coming from the predictive power of client baseline distress. None of the other predictors explained more than 1% of the outcome variance on any subscales. The LRT indicated an improvement in model fit ($p < .01$) when adding predictors for only Generalized Anxiety and the Distress Index.

Rate of Change

The ICCs for rate of change were calculated from Model 1 (Table 3). Centers accounted for an average of 3.24% of the differences in client rate of change across treatment. On specific subscales, after controlling for baseline severity, center membership explained between 1.19% (Academic Distress) and 6.89% (Alcohol Use) of the differences in outcomes. Variance explained by therapist ranged from 1.26% (Academic Distress) to 2.68% (Distress Index) and variance explained by client ranged from 90.87% (Alcohol Use) to 97.55% (Academic Distress).

Parameter estimates and fit statistics from Model 2 for all CCAPS subscales are reported in Table 3. As with the posttreatment models, negative beta values indicate more positive outcomes (i.e., per session, clients achieve change more quickly on the CCAPS). Also similar to the posttreatment analyses, the pattern of explanatory significance was generally equivalent across domains. On nearly all subscales, client baseline distress and center average baseline distress were significant, such that higher symptomatology at baseline predicted faster change (a positive outcome). Unlike the posttreatment analyses, neither session frequency nor APA accreditation significantly predicted outcome on any subscales. Session limits, IACS accreditation, center size, and integrated services were also not significantly related to outcomes on any subscales. The proportional reduction in error (PRE), or center-level variance, in rate of change for models with center-level baseline ranged from .01 (Alcohol Use) to .22 (Depression), indicating that average center baseline scores explained between 1% and 22% of center level variance in rate of change. PRE in rate of change with the addition of all center level predictors (including average baseline) ranged from .11 (Eating Concerns and Alcohol Use) to .30 (Depression and Academic Distress). We were unable to compute R^2 values. However, LRTs indicated that predictors did not improve model fit on any subscales, so we would not expect the effect sizes to be meaningful.

Table 3. Results From Multilevel Models Predicting Rate of Change on the CCAPS Across Treatment

Value	Depr.	Anx.	Soc.Anx.	Academ.	Eating	Hostility	Alcohol	DI
Client ICC	95.45%	95.71%	96.53%	97.55%	95.93%	92.11%	90.87%	94.82%
Therapist ICC	2.31%	2.07%	1.39%	1.26%	1.47%	2.22%	2.24%	2.68%
Center ICC	2.24%	2.21%	2.09%	1.19%	2.60%	5.67%	6.89%	3.03%
Model R^2	31.07%	23.72%	16.42%	19.17%	17.42%	35.22%	30.66%	27.20%
Baseline PRE	0.22	0.03	0.03	0.21	0.05	0.08	0.01	0.14
Predictor PRE	0.30	0.18	0.12	0.30	0.11	0.22	0.11	0.25
LRT	10.63	14.62	9.55	6.37	7.23	15.04	10.06	13.19
Slope	−0.46**	−0.39**	−0.29**	−0.32**	−0.45**	−0.43**	−0.47**	−0.36**
SE	0.02	0.02	0.01	0.02	0.03	0.02	0.03	0.02
Client baseline	−0.26**	−0.21**	−0.13**	−0.26**	−0.18**	−0.29**	−0.28**	−0.21**
SE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Center baseline	−0.26**	−0.14	−0.11	−0.22**	−0.20	−0.27**	−0.12	−0.24**
SE	0.05	0.07	0.08	0.06	0.10	0.08	0.11	0.07
Session limits	0.02	0.02	0.02	0.00	0.03	0.04	0.04	0.02
SE	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.01
Session freq.	0.00	0.00	0.00	0.00	0.00	0.00	−0.01	0.00
SE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IACS Accred.	0.02	0.03	0.01	0.00	0.01	0.03	0.03	0.02
SE	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.01
APA Accred.	−0.01	−0.01	0.00	−0.01	0.01	−0.01	0.00	−0.01
SE	0.02	0.02	0.01	0.01	0.02	0.02	0.03	0.01
Center size	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Integr. Services	0.02	0.02	0.02	0.01	0.03	0.03	0.04	0.02
SE	0.02	0.01	0.01	0.01	0.02	0.02	0.03	0.01
Clients, *n*	46,901	47,084	39,249	40,866	21,008	27,458	18,901	46,503
Therapists, *n*	2,361	2,361	2,358	2,358	2,299	2,337	2,261	2,361
Centers, *n*	116	116	116	116	116	116	116	116

Note. Intraclass correlation coefficients (ICCs) are reported from Model 1 (controlling for client and center baseline distress), representing the percentage of variance attributable to each level of grouping (session, client, therapist, center). Effect sizes (R^2) are reported from Model 2, representing the percentage of variance in outcome accounted for by all predictors. Proportional reduction in error (PRE) represents the proportion of center variance accounted for by center baseline CCAPS and all center predictors. Likelihood ratio tests (LRT) tested whether model fit improved between Models 1 and 2. SE = standard error. Beta coefficients for parameters are reported. CCAPS = College Counseling Assessment of Psychological Symptoms; DI = Distress Index; SE = standard error.
*p <.01, **p <.001.

Discussion

Two research questions were investigated. First, how much of the variance in outcomes was attributable to center effects? As a whole, analyses found no strong center effects, and instead revealed only a few small, nonzero effects. Second, did specific UCC characteristics or policies explain part of this proportion of variance in outcomes? Analyses revealed that the center-level predictors of interest, in combination with center baseline severity, did explain part of the center effect. However, on their own, these characteristics and policies do not show large effect sizes with regard to the total outcome variance. The results are discussed in more detail next.

Center Effects

Results showed that the center the client received treatment from explained a relatively small amount of outcome variance across all symptom domains. Controlling for initial distress at the client and center levels (Model 1), centers explained on average 1.93% of the variance in final CCAPS scores and 3.24% of the variance in rate of change across treatment. Perhaps the most parsimonious interpretation of these findings is that other factors may be more important. Consistent with previous research (Bohart & Wade, 2013; Wampold et al., 2017), ICCs in the current study demonstrated that grouping at the client-level explained much more variance than grouping at the center-level. This is in line with other CCMH findings, demonstrating that the center accounted for 2.1% of the total variance in CCAPS depression scores, whereas 64% of the variance was accounted for by differences at the client-level (Lefevor et al., 2017).

Supposing meaningful center effects do exist, it is possible that they were not captured by the calculated outcomes. Instead, perhaps other outcomes such as dropout and attendance rates are more likely to be influenced by UCC characteristics or policies.

Center Policies and Characteristics Linked with Outcome

The second research question asked if UCC characteristics or policies explained the center effect. The findings indicate that together (and when combined with average center baseline scores), they account for a substantial part of these effects (e.g., 43% of General Anxiety posttreatment scores). However, this should be contextualized around the small center effects we found. In essence, these predictors explained a large percent of a small amount of variance. A more meaningful measure of their importance is likely to be how each of them predict client change. Only two center-level predictors significantly explained the total variance in outcomes: session frequency and APA accreditation (and only for the Distress Index). The findings about session frequency explaining the center effect are consistent with the empirical link between UCC policy setting a routine session frequency, as well as research demonstrating a link between higher frequency and greater improvement (Erekson et al., 2015; Reese et al., 2011).

It should also be noted that client initial distress (included as a control) was significantly predictive of both posttreatment CCAPS score and rate of symptom change, illustrating that variables that tap into client factors continue to be robust and significant (Bohart & Wade, 2013). This client-level predictor explained a range of the client-level variance in outcomes (from 3.29% of predicting final score on Social Anxiety to 25.88% of predicating final score on Eating Concerns). The center-level version of this variable (average initial distress within a center) was included to statistically disentangle client and center effects, and greater average initial distress was significantly associated with higher final scores on all CCAPS subscales.

Limitations

Although the sample was large and represented a variety of UCCs, the interpretation and relevance of these findings may be restricted to UCCs who not only are CCMH members, but who also decide to contribute their data to CCMH's national repository. Results could be systematically biased because UCCs that do not contribute data (and who do belong to CCMH) may not do so because of lack of time, financial and staffing resources, technological capabilities, or low prioritization of research. These system-level characteristics could also be linked with routine practice and outcomes, and explain more of the center effect.

Another limitation manifested whereby a large number of clients (approximately 276,000) did not meet initial inclusion criteria. The biggest loss (approximately 209,000) occurred when clients were eliminated for at least one of the following issues: only one CCAPS administration, they could not be assigned a primary therapist, and/or they only had one appointment. Although the remaining sample is sizeable and heterogeneous, this data loss somewhat limits the conclusions that can be drawn. There may be something systematic at a center level that explains why so many clients were excluded. For example, UCCs not following best practices about consistent outcome monitoring (Gondek et al., 2016), and only administering the CCAPS once or twice.

Research Implications

As a whole, these results provide support to previous findings and point out several directions for future research. The explanatory power of session frequency is consistent with research conducted on clinical treatment models in collegiate mental health. A 2018 report published by CCMH identified UCCs that utilized a "treatment model" (e.g., clients are assigned to a clinician when an opening exists) and UCCs that utilized an "absorption model" (e.g., clinicians are assigned new clients regardless of availability). UCCs with a "treatment model" provided a significantly higher dosage (i.e., number of sessions) and frequency of treatment and produced more symptom change, compared to the "absorption model" (CCMH, 2019). These findings suggest that future studies on the center effect should look at clinical models, investigate variables that may have an impact beyond the clinical model, and examine variables that may interact with such models.

Apart from session frequency, the center-level predictors were not significantly related to outcome, including APA accreditation (for all except the Distress Index subscale), IACS accreditation, center size, integrated services, and session limits. With regard to session limits, there may be a discrepancy between what centers indicate is policy and what they actually do on a daily basis with clients. If exceptions are granted, that within-center variation may wash out a detectable center effect. Future studies might address if there is a data-based difference in the number of actual sessions used when comparing UCCs with and without reported session limits. Researchers could gather qualitative data about explicit policies, as well as when exceptions are made and under whose authority (e.g., allowing more frequent sessions for clients high in suicidality).

Another theme suggested that policies purported to be best practices were not supported by this study's results. Although IACS and APA accreditation designations are much sought after by centers (e.g., over 800 internship and post-doctoral fellowship sites have earned APA accreditation), no previous studies have looked at whether these accreditations translate into better services. In this study, IACS and APA accreditation were not associated with better outcomes (with the exception of the Distress Index subscale for APA accreditation). This is not to say that those accreditations have no benefit, but it may mean that as binary categories, these designations are too broad to have predictive validity. Future studies could identify key features used by accreditation agencies to determine if differences on those specific characteristics explain more about positive or negative outcomes.

In this same vein, it is interesting to note that schools indicating they routinely provide integrated services did not have routinely better outcomes. This suggests that either supplemental programs may not be as efficacious as intended, or that somehow the effect they are having is not being captured by this data. For instance, centers might begin treating a client with individual psychotherapy for alcohol use problems, but then later refer them to a specific substance abuse group or workshop. If they administer the CCAPS at the beginning of treatment but systematically do not to administer it for workshops, a client's change score would only reflect the individual therapy part of treatment. Similar to the accreditation variables, the nonsignificant finding may also be due to the categorical nature and lack of specificity of the predictor. Future research should aim to closely monitor the use and impact of all services received by clients.

Center size (average number of clients treated annually) also failed to predict client improvement. This suggests that the outcomes of a student at a large versus small UCC may not be systematically different. It may be more important to understand how UCCs are equipped to handle growing demand. For an additional meaningful center-level variable, future studies might use percent of the students being treated relative to the institution size.

Another possible explanation as to why the predictors accounted for such a small part of the outcome variance is that analyses failed to capture other, more important, contextual factors. Data could be gathered on other center-level variables that might strengthen effects, such as referral and fee policies, appointment reminders, and supervision practices. Also worthy of empirical attention are less "structural" types of variables that reflect important dynamic processes in the day-to-day functioning of a UCC, such as therapist burnout, organization climate, and financial resources. Researchers could also turn to more macro-level variables. Studies in community mental health centers have argued the importance of incorporating variables, such as the level of neighborhood poverty, to better understand the treatment setting surroundings (Delgadillo et al., 2016). Future studies could extend this into the UCC context by assessing environmental variables, such as average socioeconomic status of the student body, and incorporating other institutional variables such as location or demographics of the student body.

Future studies could also consider examining the predictive utility of a standardized caseload metric called the Clinical Load Index (CLI). This value represents the number of students served by total clinical hours available at the UCC. Higher CLI scores (a proxy for higher caseloads) have been linked with UCCs that serve more students, lower treatment doses (i.e., fewer overall appointments), less frequent treatment, and less symptom improvement (CCMH, 2020). The CLI may capture something different from the center size alone, and could speak to the level of demand from a particular student population, or of more internal factors like staffing and funds.

Finally, the current study (using data gathered in a naturalistic setting) could lay the ground work for empirical studies that would experimentally manipulate variables related to the associations we found between UCC variables and outcomes. For example, the current study found that more frequent sessions predicted a larger decrease in symptoms for depression and academic distress. Centers could randomly assign clients that present with either of those concerns to treatment groups with varying session frequencies.

Practice and Advocacy Implications

Within the MLM variance partitioning approach, researchers can better understand at which level to target efforts to improve treatment effectiveness.

Although the current study found effects at the center level to be small, they merit further attention because of the potentially substantial implications they hold for organizational funding, resources, and treatment planning. This is also critical at a time when centers face the increasing need to advocate for themselves to gain the resources needed to meet growing demand (Xiao et al., 2017). How best to advocate and where to spend resources is not always apparent. Higher education institutions should provide UCCs with the necessary resources (both time and financial) to conduct intensive self-studies to identify their own administrative practices that facilitate the best outcomes.

Another potential lesson is that blanket policies based on standard recommendations (e.g., "sessions should occur once a week") may not be helpful. The current study demonstrated the robustness of client factors in predicting outcomes, so centers could consider tailoring their policies based on the client's characteristics at intake (e.g., offering additional services to clients with more severe distress at baseline; McAleavey et al., 2019). Others have argued the need for a more personalized approach using risk stratification (based on disability, functional impairment, etc.) and client profiling to modify treatment recommendations (Delgadillo et al., 2016). It may also be useful to examine cross-level interactions, as the center effect may be better explained by a client's interpretation or perception of their UCCs characteristics and policies.

Based on the lack of findings related to integrated services, the current data do not support recent arguments that UCCs should provide additional services beyond counseling. It has been suggested that an overhaul of university healthcare and the integration of behavioral and medical care would greatly improve the detection and treatment of mental health problems, save costs, and improve other outcomes (e.g., dropping out of college; Alschuler et al., 2008). The current study's findings would not support doing so, and therefore decisions about supplemental services should be determined based on self-study at the UCC.

Finally, it is important to note that as organizations, UCCs work extremely hard on behalf of students and have an important role to play in influencing their clients' outcomes. The reported results should not be taken to mean that the center, its characteristics, accreditations, and/or policies are in any way meaningless to the therapeutic process. Rather, we hope that this study highlights the importance of studying the uniqueness of individual UCCs and the need for more research to be done in this area, at both a national and local level.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Dever M. Carney https://orcid.org/0000-0001-6073-1559

References

Alschuler, K., Hoodin, F., & Byrd, M. (2008). The need for integrating behavioral care in a college health center. *Health Psychology*, *27*(3), 388–393. https://doi.org/10.1037/0278-6133.27.3.388

American Psychological Association. (n.d.) *APA accreditation*. www.accreditation.apa.org

Bohart, A. C., & Wade, A. G. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 219–257). Wiley.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Center for Collegiate Mental Health. (2016). *2015 Annual report* (Publication No. STA 15-108). https://ccmh.psu.edu/files/2017/10/2015_CCMH_Report_1-18-2015-yq3vik.pdf

Center for Collegiate Mental Health. (2019). *2018 Annual report* (Publication No. STA 19-180). https://sites.psu.edu/ccmh/files/2019/09/2018-Annual-Report-9.27.19-FINAL.pdf

Center for Collegiate Mental Health. (2020). *2019 Annual report* (Publication No. STA 20-244). https://sites.psu.edu/ccmh/files/2020/03/2019-CCMH-Annual-Report_3.17.20.pdf

Coleman, J. J., Drinane, J. M., Owen, J., & Kopta, S. M. (2019). Establishing expectations: Exploring session limits in university counseling centers. *Journal of College Student Psychotherapy*, *35*(2), 1–15. https://doi.org/10.1080/87568225.2019.1650680

de Jong, K., Moerbeek, M., & van der Leeden, R. (2010). A priori power analysis in longitudinal three-level multilevel models: An example with therapist effects. *Psychotherapy Research*, *20*(3), 273–284. https://doi.org/10.1080/10503300903376320

Delgadillo, J., Moreea, O., & Lutz, W. (2016). Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. *Behaviour Research and Therapy*, *79*, 15–22. https://doi.org/10.1016/j.brat.2016.02.003

Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, *27*(29), 6137–6157. https://doi.org/10.1002/sim.3429

Erekson, D. M., Lambert, M. J., & Eggett, D. L. (2015). The relationship between session frequency and psychotherapy outcome in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, *83*(6), 1097–1107. https://doi.org/10.1037/a0039774

Falkenström, F., Grant, J., & Holmqvist, R. (2018). Review of organizational effects on the outcome of mental health treatments. *Psychotherapy Research*, *28*(1), 76–90. https://doi.org/10.1080/10503307.2016.1158883

Glisson, C., & Green, P. (2011). Organizational climate, services, and outcomes in child welfare systems. *Child Abuse & Neglect*, *35*(8), 582–591. https://doi.org/10.1016/j.chiabu.2011.04.009

Gondek, D., Edbrooke-Childs, J., Fink, E., Deighton, J., & Wolpert, M. (2016). Feedback from outcome measures and treatment effectiveness, treatment efficiency, and collaborative practice: A systematic review. *Administration and Policy in Mental Health and Mental Health Services Research*, *43*(3), 325–343. https://doi:10.1007/s10488-015-0710-5

Hammer, J. H., Perrin, P. B., & Spiker, D. A. (2019). Impact of integrated care and co-location of care on mental help-seeking perceptions. *Journal of Mental Health*, 1–6. https://doi.org/10.1080/09638237.2019.1581334

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2006). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, *9*(3), 329–343. https://doi.org/10.1093/clipsy.9.3.329

Hox, J. J., & Maas, C. J. M. (2010). Multilevel analysis. In *Encyclopedia of social measurement* (*Vol. 2*, pp. 785–793). http://dspace.library.uu.nl/handle/1874/23630

International Accreditation of Counseling Services (n.d.). *Get accredited*. https://iacsinc.org/get-accredited

Jaeger, B. (2017). *r2glmm: Computes R Squared for mixed (multilevel) models* (0.1.2) [R]. https://CRAN.R-project.org/package=r2glmm

Kim, D.-M., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research*, *16*(2), 161–172. https://doi.org/10.1080/10503300500264911

Lefevor, G. T., Janis, R. A., & Park, S. Y. (2017). Religious and sexual identities: An intersectional, longitudinal examination of change in therapy. *The Counseling Psychologist*, *45*(3), 387–413. https://doi.org/10.1177/0011000017702721

Locke, B. D., Buzolitz, J. S., Lei, P.-W., Boswell, J. F., McAleavey, A. A., Sevig, T. D., Dowis, J. D., & Hayes, J. A. (2011). Development of the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62). *Journal of Counseling Psychology*, *58*(1), 97–109. https://doi.org/10.1037/a0021282

Locke, B. D., McAleavey, A. A., Zhao, Y., Lei, P.-W., Hayes, J. A., Castonguay, L. G., Li, H., Tate, R., & Lin, Y.-C. (2012). Development and initial validation of the Counseling Center Assessment of Psychological Symptoms–34. *Measurement and Evaluation in Counseling and Development*, *45*(3), 151–169. https://doi.org/10.1177/0748175611432642

Lutz, W., Leon, S. C., Martinovich, Z., Lyons, J. S., & Stiles, W. B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*, *54*(1), 32–39. https://doi.org/10.1037/0022-0167.54.1.32

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

McAleavey, A. A., Lockard, A. J., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2015). Building a practice research network: Obstacles faced and lessons learned at the Center for Collegiate Mental Health. *Psychotherapy Research*, *25*(1), 134–151. https://doi.org/10.1080/10503307.2014.883652

McAleavey, A. A., Nordberg, S. S., Hayes, J. A., Castonguay, L. G., Locke, B. D., & Lockard, A. J. (2012). Clinical validity of the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62): Further evaluation and clinical applications. *Journal of Counseling Psychology*, *59*(4), 575–590. https://doi.org/10.1037/a0029855

McAleavey, A. A., Youn, S., Xiao, H., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2019). Effectiveness of routine psychotherapy: Method matters. *Psychotherapy Research*, *29*(2), 139–156. https://doi.org/10.1080/10503307.2017.1395921

Paulson, B., Everall, R., & Stuart, J. (2001). Client perceptions of hindering experiences in counselling. *Counseling and Psychotherapy Research*, *1*(1), 53–61. https://doi.org/10.1080/14733140112331385258

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D., & R Development Core Team. (2013). *nlme: Linear and Nonlinear Mixed Effects Models* (3.1) [R]. https://CRAN.R-project.org/package=nlme

R Development Core Team. (2014). *R: A language and environment for statistical computing*. (3.4.1). R Foundation for Statistical Computing. http://www.R-project.org/

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed., Vol. 1). Sage.

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*(5), 608–619. https://doi.org/10.1080/10503307.2011.598580

Schiefele, A.-K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., Kopta, M., Schulte, D., Saxon, D., Nielsen, S. L., & Lambert, M. J. (2017). Reliability of therapist effects in practice-based psychotherapy research: A guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(5), 598–613. https://doi.org/10.1007/s10488-016-0736-3

Wampold, B. E., Baldwin, S. A., Holtforth, M. G., & Imel, Z. E. (2017). What characterizes effective therapists? In L. G. Castonguay, & C. E. Hill (Eds.), *How and why are some therapists better than others? Understanding therapist effects* (pp. 37–54). American Psychological Association.

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work*. Routledge.

Wolgast, B. M., Lambert, M. J., & Puschner, B. (2004). The dose-response relationship at a college counseling center: Implications for setting session limits. *Journal of College Student Psychotherapy*, *18*(2), 15–29. https://doi.org/10.1300/J035v18n02_03

Xiao, H., Carney, D. M., Youn, S., Janis, R. A., Castonguay, L. G., Hayes, J. A., & Locke, B. D. (2017). Are we in crisis? National mental health and treatment trends in college counseling centers. *Psychological Services*, *14*(4), 407–415. https://doi.org/10.1037/ser0000130

Author Biographies

Dever M. Carney, MS, is a doctoral candidate in clinical psychology at Penn State University. She conducts practice-oriented research focused on psychotherapy process and outcome, with particular emphasis on how organizational policies influence treatment outcomes.

Louis G. Castonguay, PhD, completed his doctorate in clinical psychology at SUNY Stony Brook, a clinical internship at UC Berkeley, and a Post-doctorate at Stanford University. He is currently a liberal arts professor of psychology at Penn State University. His research focuses on the process and outcome of psychotherapy, as well as on practice-based evidence and practice-research networks.

Rebecca A. Janis, PhD, is a data analyst at the Center for Collegiate Mental Health. Her research focuses on methodological and statistical issues in the evaluation of psychotherapy process and outcome.

Brett E. Scofield, PhD, has devoted nearly his entire career to collegiate mental health, working as a clinician and administrator within numerous university-based counseling centers over the past 17 years. He currently holds leadership positions within Penn State Counseling and Psychological Services, as well as the Center for Collegiate Mental Health, a national practice-based research network of over 650 college counseling centers. Dr. Scofield has several publications in the area of college student mental health.

Jeffrey A. Hayes, PhD, is a professor in the Department of Educational Psychology, Counseling, and Special Education at Penn State, with a courtesy appointment in the Department of Psychology. He is also the current editor of *Psychotherapy Research* and maintains an independent psychotherapy practice in State College, PA. Dr. Hayes is the author of more than 100 publications focused primarily on psychotherapy and college student mental health.

Benjamin D. Locke, PhD, is the director of Penn State's Counseling and Psychological Services; the founder and executive director of the Center for Collegiate Mental Health, a practice–research network of over 650 counseling centers; and an affiliate faculty member in the counseling and clinical psychology departments at Penn State. He presents and consults widely about student mental health in higher education and has published dozens of peer-reviewed articles on the topic.