The Pennsylvania State University

The Graduate School

College of the Liberal Arts

THE EFFECT OF COUNSELING CENTER CHARACTERISTICS AND POLICIES ON PSYCHOTHERAPY OUTCOMES

A Thesis in

Psychology

by

Dever M. Carney

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 2018

The thesis of Dever M. Carney was reviewed and approved* by the following:

Louis G. Castonguay
Professor of Psychology
Thesis Adviser

Michael N. Hallquist
Assistant Professor of Psychology

Jeffrey A. Hayes
Professor of Education (Counselor Education)
Professor of Psychology

Melvin M. Mark
Professor of Psychology
Head of the Psychology Department

*Signatures are on file in the Graduate School.

ABSTRACT

Researchers agree that psychotherapy works, but not for everyone. Above and beyond contributions from the client, therapist, and therapy process, contextual factors may have a unique impact on the outcomes a person achieves in psychotherapy. At a contextual level, college counseling centers often make key administrative decisions that have the potential to systematically affect their clients' outcomes in psychotherapy. The current project used data collected through the Center for Collegiate Mental Health, a practice-research network with over 500 participating college and university counseling centers. Clients' symptoms were measured on the Counseling Center Assessment of Psychological Symptoms (CCAPS), a multidimensional instrument designed for repeated assessment in collegiate mental health settings. The final sample contained 105 centers, 1,601 therapists, and 29,028 clients, and outcome was operationalized as the latent difference score between CCAPS subscale scores at the beginning and end of treatment. Multilevel modeling was used to estimate the percent of the variance in outcome accounted for by the specific counseling center, and further sought to explain that "center effect" by examining the role of a number of specific administrative policies and characteristics like specific services, session limits, student to staff ratios, etc. (after controlling for key client variables). Results found a relatively small center effect, ranging from 1.50% (social anxiety subscale) to 3.32% (hostility subscale). Significant predictors of these center effects were treatment length, initial symptom severity, and the average initial symptom severity at a center, while the majority of other center variables examined were non-significant. This has potentially wide-ranging implications for counseling center policies and resource allocation.

TABLE OF CONTENTS

List of Tables……………………………………………………………………………………..vi
List of Figures………………………………………………………………………………vii
Acknowledgements…………………………………………………………………………..viii

Chapter 1. INTRODUCTION…………………………………………………………………..1
Client factors……………………………………………………………………………1
Therapist factors…………………………………………………………………………...1
Process variables…………………………………………………………………………..2
Contextual factors…………………………………………………………………………3
Contextual factors related to outcome……………………………………………………..5
Gaps in the literature………………………………………………………………………6
Current study………………………………………………………………………………8
Center for Collegiate Mental Health………………………………………………8
Hypotheses and research questions……………………………………………….9

Chapter 2. METHODS…………………………………………………………………………...13
Procedure………………………………………………………………………………...13
Sample……………………………………………………………………………………13
Clients……………………………………………………………………………13
Therapists………………………………………………………………………...14
Centers…………………………………………………………………………...15
Measures…………………………………………………………………………………15
Client-level symptom data……………………………………………………….15
Client- and therapist- level appointment data……………………………………17
Center-level data…………………………………………………………………17
Statistical Analyses………………………………………………………………………17
Operationalizing client outcomes………………………………………………..17
Predicting client outcomes……………………………………………………….18

Chapter 3. RESULTS…………………………………………………………………………….23
Depression………………………………………………………………………………..23
Anxiety…………………………………………………………………………………...25
Social Anxiety…………………………………………………………………………….27
Eating Concerns…………………………………………………………………………..29
Hostility…………………………………………………………………………………..31
Academic Distress………………………………………………………………………..33
Alcohol Use……………………………………………………………………………...35

Chapter 4. DISCUSSION……………………………………………………………………….38
Center effects……………………………………………………………………………..38
Center policies and characteristics affecting outcome…………………………………...40

Clinical and administrative implications……………………………………………...45
Limitations……………………………………………………………………………47
Future directions………………………………………………………………………48

References…………………………………………………………………………………..51

Appendix A: Tables 1-10……………………………………………………………....63-79
Appendix B: Figure 1……………………………………………………………………….80
Appendix C: CCAPS-34……………………………………………………………………81
Appendix D: SDS……………………………………………………………………………..82

List of Tables

Table 1: Summary of client-level variables and demographics

Table 2: Summary of center-level variables

Table 3: Results of MLM predicting change in Depression scores

Table 4: Results of MLM predicting change in Anxiety scores

Table 5: Results of MLM predicting change in Social Anxiety scores

Table 6: Results of MLM predicting change in Eating Concerns scores

Table 7: Results of MLM predicting change in Hostility scores

Table 8: Results of MLM predicting change in Academic Concerns scores

Table 9: Results of MLM predicting change in Alcohol scores

Table 10: Summary of predictors from final multilevel model for each subscale

List of Figures

Figure 1: Latent change score model for change on Eating Concerns

Acknowledgements

I would first like to express sincere gratitude and appreciation to the members of my committee, Drs. Jeff Hayes, Michael Hallquist, and Louis Castonguay. Thank you for the time and effort you gave to reviewing this project and for the enlightening discussions we shared. Your insightful comments, sharp suggestions, and expert feedback were instrumental to enhancing the final project, and for advancing my own growth as a scholar. In particular, I'd like to thank my advisor, Louis Castonguay, for supporting me throughout this process. Your vision, wealth of expertise, hard work, and guidance helped this project come to fruition.

I cannot express enough gratitude for my lab mates, friends, family, and partner for their unwavering support during a stressful time of my life. Henry, Rebecca, and Ryan, you all made the lab a wonderful place to work, and I couldn't ask for better people to share an office with. In particular, I'd like to thank Rebecca for being a wonderful work wife and helping me at every step of this process. Your guidance and assistance during the data cleaning, analyses, and writing were essential, and I'm honestly not sure how I would have done this without you. Thank you for your patience and selflessness, and for sharing your beautiful brain with me during many late nights at Moore. To Lia, Chloe, Thomas, Erin, Emily, Einat, and Ben, thank you for being amazing friends and for gracefully listening to me complain and worry for so many months.

To my wonderful family—Mom, Dad, Alex, Julia, Jackson, and Anna—I'm eternally grateful for your steady support, kindness, and understanding that I had to work even when we were on a family vacation in Mexico. And finally, thank you to my boyfriend, Ryan. Your calming presence, rock-solid support, and never-ending patience made it possible for me to pursue a project that was both challenging and rewarding. I am lucky to have you in my life.

Chapter 1: Introduction

It has become a clear consensus among researchers that psychotherapy works, but not for everyone (Castonguay & Beutler, 2006; Lambert & Ogles, 2004). While study findings vary, estimates place the overall rate of client improvement between 40-70%, compared to about 5-10% of clients who actually deteriorate. For decades, scientists have explored what elements in particular influence whether a client experiences a positive outcome (yielding a reduction in psychopathological symptoms and/or distress), experiences no change, or experiences a negative outcome (an increase in symptomatology and/or distress).

Client factors

In some cases, researchers have demonstrated that psychotherapy outcomes are related to client factors, or characteristics that are unique to the individual seeking treatment. To name a few, there is evidence that a client's expectations (Constantino, Arnkoff, Glass, Ametrano, & Smith, 2011), attendance patterns, reflective functioning (Elkeblad, Falkenström, Holmqvist, 2016), personality characteristics, perceptions about therapist credibility (Goates-Jones & Hill, 2008), and motivation (Ilagan, Vinson, Sharp, Ilagan, & Oberman, 2015) can affect how well they do in treatment (Bohart & Wade, 2013). One of the more robust findings is that a client's eventual outcome is related to their level of presenting distress or severity (Clarkin & Levy, 2004; Hansen & Lambert, 2003; Saxon & Barkham, 2012).

Therapist factors

A large number of studies have also found evidence that therapist differences play a role in influencing outcomes; a so-called "therapist effect" (Castonguay & Hill, 2017). Therapist effects refer to the variability in outcome attributable to therapists. In other words, not all therapists are equal and some better than others at fostering good outcomes with their clients. Previous studies have linked client change with a therapist's level of facilitative interpersonal skills (e.g., Anderson et al., 2016), deliberate practice outside of session (Chow et al., 2015), reflective functioning (Cologon, Schweitzer, King, & Nolte, 2017), and ability to build a therapeutic alliance (Baldwin, Wampold, Imel, 2007; Falkenström, Granström, & Holmqvist, 2014). Interestingly, most findings reported insignificant contributions to outcome from therapist age, gender, ethnicity, theoretical orientation, level of experience, professional degree, or the matching of client and therapist on key demographic variables (Shiner et al., 2017; Wampold, 2017). Summary estimates place the overall therapist effect on client outcomes around 5-8% (Castonguay & Hill, 2017), and it has also been suggested that therapist effects may interact with client variables. For example, one study reported that a greater therapist effect was found for more severe clients who start out treatment a higher level of symptomatology (Saxon & Barkham, 2012).

Process variables

Related to but conceptually distinct from client and therapist characteristics are a number of process variables that have also been linked to outcomes. Among them are several factors linked to the therapeutic relationship, such as alliance, empathy, and positive regard (Norcross, 2011). While the adherence to and competent delivery of treatment manuals appears not to be related to outcome (Webb, DeRubeis, & Barber, 2010), the use of some interventions have, such as specific cognitive therapy techniques (Crits-Christoph, Connolly Gibbons, & Mukherjee, 2013). Furthermore, therapeutic improvement has been linked to the interaction of client, therapist, and process variables (Castonguay, 2013).

Contextual factors

Despite this wide array of clinically meaningful factors, researchers still do not have a complete understanding of why some clients do better than others in therapy. In 2011, Norcross and Lambert put forth a theoretical suggestion that only about 60% of the variance in psychotherapy outcomes is currently explained. They estimated that client characteristics accounted for about 30% of the total variance, the specific treatment method accounted for about 8%, the therapeutic alliance for 12%, and therapist characteristics for 7%, leaving about 40% of the variance unaccounted for. Although these estimates should not be considered as a definitive breakdown of therapy outcome variance components, their suggestion stimulated the field to both test the assertion that 60% of the variance is explained by those factors, as well as further explore what might account for the unexplained differences across clients.

In the same way that the field asks how and why some therapists are better than others, there is a need to investigate how and why some treatment settings are better than others. Arguments have been made that contextual factors beyond the client, therapist, and dyadic process may account for some of that remaining variance, and several researchers have maintained the necessity of incorporating data about the therapy setting into analyses of psychotherapy outcomes (Wampold & Imel, 2015). McLeod (2013) stated that contextual factors (like the decision to disseminate new treatments within a therapy practice) have implications for client outcomes, and further suggested that the effectiveness of therapy may depend less on specific treatments and the skillfulness of therapists, and more so on the ways in which those concepts and skills are adapted and rebuilt within a local practice. In a similar vein, Powell and Beidas (2016) argued for the necessity of empirical research on both the inner context (intraorganizational characteristics like culture and policy) and outer context (funding, leadership, inter-organizational environment, advocacy and support) of behavioral health settings to inform administrative best practices, as well as facilitate system-level efforts to implement evidence-based treatments. Baldwin and Imel (2013) remarked that most therapist effect studies occur in naturalistic settings and often do not randomly assign therapists, and therefore more studies should examine potential confounding "third" variables that might account for observed therapist effects, such as referral policies within a given therapy practice.

Falkenström, Grant, and Holmqvist (2018) reviewed the current literature on organizational effects, and argued that the investigation of contextual factors and psychotherapy outcome is essential, and remarked that "it is remarkable that not more research has been devoted to this." They demonstrated that across the 19 studies reviewed, all found some evidence of an organizational effect, providing evidence that the treatment setting may be an important source of variance in outcomes. They point out that studies should continue to examine the explanatory power of organizational climate and culture, but also noted the predictive promise shown by differences in patient populations and differences in treatment processes. They also note that large practice research networks (like the one in the current study) may be particularly well-suited to the examination of center effects, due to the size and complexity of the data being gathered.

Various studies have explored contextual factors in both community outpatient and college counseling settings, through an examination of organizational policies that may affect how treatment is delivered. The findings described here are largely descriptive and lack ties to client outcomes and symptom change, but highlight important between-organization differences.

For example, a little more than half of college counseling centers impose limits on the number of sessions allowed per client (Reetz, Bershad, LeViness, & Whitlock, 2015), often to accommodate increasingly higher demand and longer waitlists. However, one study found that on average, college counseling centers with time-limited services not only had longer waitlists but were still unable to serve a higher proportion of the student body (Gyorky, Royalty, & Johnson, 1994). In this context where resources are often strained, it is also common practice to offer brief services but then refer students to external providers (Owen et al., 2007; Stone & McMichael, 1996) The decision-making process for referrals has been detailed by several researchers, and there appears to be heterogeneity (especially in the college counseling world) with regard to how, when, and why clinicians refer patients out of the therapy practice (Gage & Gyorky, 1990; Lacour & Carter, 2002; Lawe, Penick, Raskin and Raymond, 1999; Quintana, Yesenosky, Kilmartin, & Macias, 1991). Researchers have also advocated for further research on organizational-level barriers to treatment in the college student population, such as affordability, availability, accessibility, and acceptability; some of which may be reflected in counseling center policies (Marsh & Wilcoxon, 2015).

Contextual factors related to outcomes

A few studies have examined how policies might be linked to client outcomes. For example, Erekson, Lambert, and Egget (2015) demonstrated that more frequent therapy sessions was associated with clients achieving clinically significant gains more quickly, but not necessarily achieve more change overall. Others provided evidence that the majority of clients required 14 sessions or more to achieve clinically significant change, suggesting that lower session limits in counseling centers may not be optimally beneficial, and that longer treatments may be more beneficial (Wolgast, Lambert, & Puschner, 2003). Findings also show that briefer treatment may be adequate for some, but that the majority of clients will not make the necessary gains (Bohart & Wade, 2003). However, some studies have shown that session limits are not related to clients seeking treatment at college counseling centers (Uffleman & Hardin, 2002), clients' attendance (Gallagher, 2005), or overall outcomes (Orlinsky, Ronnestad, & Willutski, 2004). The mixed findings merit further investigation. Another study looked at the relationship between student to full-time equivalent (FTE) staff ratios and outcomes in college counseling centers, and found that in general, smaller (i.e. "better") student to staff ratios were not linked with more positive outcomes. However, findings also suggested that despite ratios falling behind recommended thresholds, most centers were still able to provide adequately effective treatment (Elreda, 2014). Another study within the CCMH framework did not explicitly look at policies, but used statistical techniques that allowed researchers to examine contributions from center membership. In a multilevel exploration of psychotherapy change for religious and sexual minorities, they found that differences between clients explained much of the variance than differences between centers (Lefevor, Janis, & Park, 2017).

Taken together, these studies represent an important beginning towards understanding how centers, at a higher level, can affect the outcomes their clients achieve.

Gaps in the literature

Despite this important contribution, the studies listed above that have linked center policies or characteristics to client outcomes are relatively few in number. What's more, these investigations have methodologically and statistically failed to delineate the part of the outcome variance that is uniquely due to center variables. Despite the repeated call for investigation of contextual variables, studies are slow to emerge in this area. One of the challenges in parsing apart the variance in psychotherapy outcomes is the hierarchical nature of the constructs being studied. The fact that clients are "nested" within therapists and therapists are "nested" within therapy practices or centers violates traditional regression assumptions of independent observations. In recent decades, it has become standard practice to statistically account for these clustered relationships using multilevel modeling (MLM), which derives therapist and organizational effects that correspond to an intraclass class correlation (ICC) coefficient (Raudenbush & Bryk, 2002). Of the studies mentioned above that link center variables with client psychotherapy outcomes, only two used multilevel modeling to account for inherently hierarchical data structures (Erekson, Lambert, Egget, 2015; Lefevor, Janis, & Park, 2017).

In the literature, issues often cited with utilizing MLM are too-small sample sizes and insufficient power to detect effects. Conducting analyses with too few clients per therapist or too few therapists per center can lead to biased parameter estimates and inflated rates of type I errors (Schiefele et al., 2017). Researchers have stated the need for larger sample sizes (de Jong et al., 2010; Kim et al., 2006; Soldz, 2006; Williams, 2016), though the difficulties regarding lack of resources and system-level constraints are often acknowledged as impediments to securing adequate sample sizes. These issues have occasionally been addressed with simulation studies (e.g. Bell et al., 2014), but these have historically delivered results that are inconsistent with other research studies. So far, very few projects have been able to muster the sample sizes to meet recommended guidelines (Schiefele et al., 2017). In the studies mentioned above linking center variables to outcomes, researchers were often dealing with small to medium samples, ranging from 89 to 1925 clients.

These studies also do not attempt to investigate the unique contribution of the center itself (i.e. a “center effect”), in the similar manner to how therapist effects have been calculated. This is likely due to difficulties acquiring enough units at the center level to conduct multilevel modeling. While understandable, this represents an issue in the field where a potentially important source of variance is being underexplored. It also means that the majority of study findings are generated within only one center, and while theoretically important, are not generalizable to the population as a whole (e.g. Erekson et al., 2003; Mahon et al., 2015; Wolgast et al., 2004).

Current study

To fill in the gaps in the center effect literature, the current study will attempt to address the aforementioned issues by utilizing a large, representative, and heterogeneous sample, and accounting for unique center contributions through multilevel modeling.

Center for Collegiate Mental Health. The current study will investigate organizational-level variables with a naturalistic sample derived from the Center for Collegiate Mental Health (CCMH). CCMH is a practice research network involving over 480 college and university counseling center members, each of them contributing client-, therapist-, and center-level data every year (Castonguay, Locke, & Hayes, 2011; McAleavey, Lockard, Castonguay, Hayes, & Locke, 2015). Examining the possible impact of center-related variables is particularly timely for this infrastructure. As a higher numbers of students seek treatment every year (Watkins, Hunt, & Eisenberg, 2012; Xiao et al., 2017), college counseling centers across the country experience growing demands for services. Using a nationally representative dataset to identify center-level features that are associated with good and poor outcomes has the potential of better understanding and working toward improving the effectiveness of treatments in this setting. Ultimately, this line of research can help guide center directors and policymakers towards best administrative practices, as well as provide evidence to support centers advocating for more resources.

With its rich dataset containing numerous types of clinical and demographic data collected from a nationwide membership, CCMH is uniquely positioned to provide insight on how a wide array of center variables are related to clients' improvement. Predictors in the current study will be a number of center variables that have already received empirical support with regards to client outcomes, as well as some that have not been previously tested.

Hypotheses and research questions. The current study will estimate the variance components contributing to client differences in outcomes, by assessing what portion of the differences can be attributed to client, therapist, and organizational factors associated with the agency in which treatment is provided. This project is largely exploratory in nature, and ultimately aims to extend previous research, as well as explore questions that have not been addressed by previous studies. Two main research questions will be explored:

1) How much of the variance in psychotherapy outcomes can be attributed to center effects?
2) Which specific center characteristics or policies help explain this proportion of variance in psychotherapy outcomes?

Analyses to answer this second question will be guided by a set of hypotheses that predict a replication of findings, and a second set of research questions based on previously mixed findings or no prior research. With regard to the replication of findings, it is hypothesized that:

1) A client's initial level of severity will be significantly related to client outcomes, such that higher pre-treatment distress will predict greater change (Hansen & Lambert, 2003; Hansen, Lambert, & Forman, 2002).
2) A longer treatment length for clients will be associated with a greater reduction in symptoms at post-treatment (Bohart & Wade, 2003).
3) A center's average level of initial severity will be significantly related to outcomes, such that higher pre-treatment distress will predict greater change (this is extrapolated from findings at the client level).
4) The presence of limits will predict a smaller decrease in symptoms at post-treatment (Bohart & Wade, 2003).
5) Session frequency will not be significantly related to client outcomes (Erekson et al., 2015).
6) A longer treatment length for clients will be associated with a greater reduction in symptoms at post-treatment (this is extrapolated from findings at the client level).

Beyond the above predictions guided by previous findings, another center-level variable has garnered mixed or unexpected results in the literature and thus lends itself to further exploration here. Since 1981, for example, the International Association of Counseling Services (IACS) has warned against large student to FTE staff ratios in counseling centers, because this purportedly prohibits the best access to mental health care (Spivack et al., 2010). However, one study found that although centers frequently fall behind the IACS guidelines, the majority of centers were still able to provide adequate care and achieve good client outcomes (Elreda, 2014). This unexpected finding leads to the following research question to be explored:

7) Will student to FTE staff ratio be related to client outcomes?

To the best of our knowledge, no identified studies have examined a connection between a center's APA or IACS accreditation status and psychotherapy outcomes. Because these statuses are a recognition of meeting and maintaining high educational and professional standards in the field, it is conceivable that the accreditation status of the treatment setting will be related to outcomes. The accreditations also require certain standards of care be met (e.g. rigorous internal documentation, using evidence-based practices), so this status might serve as a proxy variable, capturing other (positive) aspects of the counseling context. Similarly, no studies have been conducted on counseling center size and outcome. Size may be related to other predictors (e.g. student-staff ratio), but it also possible that a client's experience of receiving treatment at a small versus large counseling center (and the feeling of being recognizable and attended to, or just one of many clients) may impact how they attend, engage with, and benefit from treatment.

Also to the best of our knowledge, no studies have examined whether the integration of various services in the collegiate mental health treatment setting are linked with client outcomes. Efforts towards such integration has been a movement in other fields (e.g. medicine, social work; Balasubramanian et al., 2017; Stanhope & Staussner, 2017), and is currently the object of much focus and attention in the public policy sector. For example, the Substance Abuse and Mental Health Services Administration (SAMHSA) created the Center for Integrated Health Solutions (CIHS), which has spent over $25 million dollars in recent years to facilitate the integration movement (SAMHSA-HRSA Center for Integrated Health Solutions, 2011). It therefore seems indicated to investigate the relationship between integrated services and outcomes.

Finally, no identified studies have examined whether institutional-level variables (such as NCAA athletic division and whether a school is public or private) affect client outcomes. Athletic division can often be related to student culture and number of scholarship athletes, so it is possible that this variable may capture something about the differences in students seeking services at the counseling center. Athletes (especially in Division I) have been shown to be at a higher risk for alcohol abuse and disordered eating, and may therefore shift the distribution and presentation of treatment-seeking students at schools (Engel et al., 2003). Likewise, school type (public or private) may capture something about the student body related to academic or financial stress, as this varies across school type. Within this context, the following exploratory research questions that will be addressed:

8) To what extent will counseling center APA accreditation status be related to client outcomes?
9) To what extent will counseling center IACS accreditation status be related to client outcomes?
10) To what extent will the number of clients at a counseling center be related to client outcomes?
11) To what extent will the offering of a variety of integrated services be related to client outcomes?
12) To what extent will the type of school where a center is situated in be related to client outcomes?
13) To what extent will the athletic division of the school where a center is situated be related to client outcomes?

Chapter 2: Methods

Procedure

As mentioned above, data were collected through the nationwide practice-research network of CCMH. In this bottom-up infrastructure, participating centers collect data in a naturalistic setting via standardized instruments and electronic medical record (EMR) systems, which are then contributed to a national data repository. De-identified data are uploaded on an annual basis, which are subsequently cleaned and consolidated by research staff at Penn State University. Each center secures and maintains study approval from their local Institutional Review Board.

Sample

Data were used from the CCMH data repository spanning the 2014-2015 and 2015-2016 academic years. Although CCMH represents a collaborative network of over 480 university counseling centers, not all member centers contribute data on an annual basis. The final sample contained 105 centers, with a total of 1,601 therapists and 29,028 clients.

Clients. The clients in this dataset are all graduate or undergraduate students presenting for treatment at college and university counseling centers across the country. A data reduction process required that clients were only included if they attended at least two individual therapy sessions, and if they completed at least two symptom questionnaires (to ensure a change score could be computed), within at least two weeks of their first and last sessions respectively (to ensure the change score captures the beginning and end of treatment). Because the dataset spans two years, some students received multiple courses of treatment. If clients had more than one course of therapy (indicated by appointments separated by more than 90 days), only their first course of treatment was retained. This accounted for the fact that clients might be expected to demonstrate different change patterns if they are returning to the same center for additional courses of treatment.

At the start, the dataset contained 162,888 clients. The final sample contained 29,028 clients who met inclusion criteria, and it should be noted that the greatest decrease in sample size was setting a minimum of two CCAPS, where the loss in clients was about 105,000 clients. The final sample was 37.95% female, 18.24% male, 0.28% self-identify, and 0.70% other. The average age was 22.7 years (*SD*= 5.11), and the demographic breakdown of ethnicities and other client variables are reported in Table 1. Clients represented a range of years in school, sexual orientations, and religious denominations.

Therapists. Therapists in this dataset represent treatment providers across all contributing centers that were designated in a center's local electronic medical record (EMR) system as the therapist for at least one client in the data. Inclusion criteria for the current study required therapists to have at least five clients meeting the above criteria, ensuring that multilevel modeling with therapist as a grouping factor could be done. Some therapists completed an optional survey of demographic items each year. Demographic summaries are not reported for therapists, due to the low response rate on these self-reported items (e.g. only 46.3% of therapists completed the age item). It was decided that these numbers would not be representative of the overall therapist population in this sample.

At the start, there were 4,459 unique therapists in the data. After data reduction, the dataset contained 1,601 therapists, with an average of 17.4 clients (*SD*= 15.11) per therapist.

Centers. Centers who contributed data were only included if they had a minimum of 25 clients with valid data, to ensure that center data was representative and reliable at an organizational level. Centers were also excluded if they had not contributed data on any one of the predictor variables. At the start, the dataset contained 148 centers, and the final dataset contained 105 centers, with an average of 276.46 clients (range = 25 to 2,419; *SD* = 321.68) and 15.89 therapists per center (range = 2 to 55; *SD* =11.30). The average treatment length was 6.89 sessions (*SD*= 1.88), the average session frequency was 1 session every 15.56 days (*SD*= 3.24). A summary of center-level variables is presented in Table 2.

Measures

This study utilized client-level appointment data, client-level symptom data, and center-level aggregate data.

Client-level symptom data. Client symptoms across several domains were assessed via the College Counseling Assessment of Psychological Symptoms (CCAPS), which is a multidimensional measure specifically designed to assess collegiate mental health concerns. The long form of the CCAPS contains 62 items, and has excellent psychometric properties (Locke et al., 2011). The short form retains 34 of the original items, and was developed after counseling centers reported that the CCAPS-62 was too lengthy for routine clinical assessment and monitoring (Locke et al., 2012). The CCAPS-34 demonstrates good convergent validity with comparison measures. It also demonstrates good test-retest reliability, with 1-week and 2-week reliability coefficients ranging from r = .79 (alcohol use) to r = .87 (depression). The magnitudes of these convergent correlations are similar to what were found for the CCAPS-62, and when administered as a stand-alone instrument, the CCAPS-34 supports the factor structure that emerged for the CCAPS-62 (Locke et al., 2012).

The CCAPS-34 has seven empirically-derived subscales: Depression (6 items), Generalized Anxiety (6 items), Social Anxiety (5 items), Academic Distress (4 items), Eating Concerns (3 items), Alcohol Use (4 items), and Hostility (6 items). Because the CCAPS-34 contains only items found in the CCAPS-62, administrations of the CCAPS-62 can be scored as a short form using only the items from the CCAPS-34, which was done in the current study to include as much client data as possible. Centers are able to decide locally whether to administer the long or short form of the CCAPS, and the CCAPS-34 is administered more frequently than the CCAPS-62 (likely due to perceived time constraints). On the CCAPS-34, clients are asked to rate themselves on a number of items relating to the past two weeks. Reports are made on a Likert scale, ranging from 0 (*not at all like me)* to 4 (*extremely like me*). Total subscale scores are derived by calculating the average of all items which load onto that particular subscale. After reverse-coding certain items, higher subscale scores indicate more distress, and lower scores indicate lower distress, with scores ranging from 0 to 4. Subscales and the corresponding CCAPS items are listed in Appendix A.

Because this study is mostly exploratory in nature, no single domain of functioning on the CCAPS-34 was the focus of investigation. Pre- and post-treatment scores on items from each subscale were used to calculate latent change scores (detailed below), representing that client’s psychotherapy outcomes. Subscale scores from each domain were also included as client-level predictor to control for the initial severity of symptoms.

Client- and therapist- level appointment data. Appointment data were collected by the center's local EMR system, and include information such as date, therapist, attendance status (e.g. attended, no show, cancellation), and appointment type (e.g. was this a group or individual therapy appointment). Treatment length (i.e. the number of attended sessions) was also included as a client-level predictor. Client-level variables were used to match clients and therapists, as well as calculate aggregate predictor variables at the center level (e.g. average session frequency).

Center-level data. Center-level data were collected via the Standardized Data Set (SDS), a standardized questionnaire created from the collective intake materials of multiple counseling centers (Castonguay, Locke, & Hayes, 2011, CCMH, 2012). At the center level, the SDS is completed annually by center directors or administrators at each site, and has items capturing organizational characteristics and policies. It provides information about various policies and features of each center, including: APA accreditation status, IACS accreditation status, four types of integrated services (health, learning, drug and alcohol, and psychiatric care), and session limits. They also provided information about the institution in which the center is housed: athletic NCAA division, enrolment size, type (public, private, or other). A number of aggregated predictors were also computed for each center: average treatment length, average session frequency, center size, and student to staff ratio. Relevant SDS items are listed in Appendix B.

Statistical Analyses

Operationalizing client outcomes. We were primarily interested in client outcome after engaging in psychotherapy, as operationalized by a pre- to post-treatment symptomatology and as measured by the change in scores on each CCAPS-34 subscale. Studies have demonstrated that raw change scores (computed by subtracting *time2 – time1)* tend to be less reliable than the component variables used to calculate them (Alison, 1990; Kessler, 1977; King, 2006) and do not take into account the nearly universal phenomenon of regression to the mean between *time1* and *time2*. One solution is to utilize the structural equation modeling (SEM) framework to generate latent change score models (LCSM; also called latent difference scores). Latent change score modeling is well-suited to examining longitudinal relationships, as well as detecting complex covariance patterns in large sets of continuous variables (McArdle, 2009, McArdle & Ghisletta, 2012).

In the current study, LCSM was applied to the first and last instances of the CCAPS-34 items to calculate a latent change score of symptoms from pre-treatment (time1) to post-treatment (time2). Analyses were conducted in the "lavaan" package (Rosseel, 2012) in the R programming language (version 3.4.1; R Development Core Team, 2014), and used a robust estimator and maximum likelihood estimation to handle missing CCAPS data, which was assumed to be missing at random. For each subscale, observed scores on CCAPS-34 items belonging to that subscale were freed at time1 and constrained to zero at time2, so that the intercepts of the change scores represented change in symptoms at time2 not accounted for by time1.

As an example, Figure 1 illustrates the current study's path diagram of the latent change score for the Eating Concerns symptom domain. Similar latent change score models were constructed for each of the other domains.

Predicting client outcomes. Multilevel modeling (MLM) was used to examine the research questions, controlling for the fact that clients are non-independent and nested within therapists, and that therapists are nested centers (Bryk & Raudenbush, 2002). Analyses were conducted using maximum likelihood estimation with the "nlme" package (Pinheiro, Bates, DebRoy, Sarkar, & R Development Core Team, 2013) in the R programming language. A series of six models were built with three levels: clients within therapists (level 1), therapists within centers (level 2), and counseling centers (level 3), allowing the examination of unique contributions from each source to client psychotherapy outcomes.

The research questions were addressed in two steps: first to examine the effect of center membership on outcomes without any predictors, and then adding predictors to identify which variables account for variance explained by the center. The steps are explicated in further detail below. The first research question asked how much of the variance in client outcomes was attributable to the center. To answer this in Model 1, client outcome on a given subscale was modeled as a fully unconditional (null) multilevel regression with a random intercept and center and therapist included as a grouping variable, which allowed us to determine how variance in client outcome was allocated across the client, therapist, and center levels. Therapists were included at Level 2 to account for nesting, but no predictors were modeled at this level. The kinds of variables that have historically explained some of the differences between therapists are not currently available within the CCMH infrastructure (e.g. a therapist's facilitative interpersonal skills or deliberative practice). A generic equation for Model 1 is presented here, as it was structurally equivalent across all subscales:

$$L1\_Client : [Subscale]_{tij} = p_{0ij} + e_{tij}$$
$$L2\_Therapist : p_{0ij} = b_{00j} + u_{0ij}$$
$$L3\_Center : b_{00j} = g_{000} + u_{00j}$$

Here, [*Subscale*]$_{tij}$ represents the change in that subscale score at time *t* for client *i* at center *j*. Random intercepts were included at the therapist and center level (u_{0ij} and u_{00j}), allowing therapists and centers to have a unique deviation from the average change score, in addition to a term of residual variance (e_{tij}).

In Model 2, change in symptom domain scores was modeled as a function of a random intercept and client-level variables. No center-level predictors were added, but two client-level predictors were entered to control for client factors that have been known to impact outcomes: a client's initial severity and that client's treatment length, or "dose" (Bohart & Wade, 2013). These two client-level predictors were centered around the counseling center mean to disentangle the client-level and center-level contextual effects (Raudenbush & Bryk, 2002). This allows for comparison between models with and without center variables, but still taking into account contributions at the client level. A generic equation for Model 2 is presented here, as it was structurally equivalent across all subscales:

$$
\begin{aligned}
&L1\_Client:[Subscale]_{tij} = p_{0ij} + p_{1ij}(initial.sev) + p_{2ij}(tx.length.) + e_{tij}\\
&L2\_Therapist: p_{0ij} = b_{00j} + u_{0ij}\\
&L3\_Center: b_{00j} = g_{000} + u_{00j}
\end{aligned}
$$

Random intercepts were again included at the therapist and center level, allowing therapists and clients to vary on their average change. Variance attributable to each level (or the ICC) was calculated as a ratio of specific-level error variance divided by the total error variance (Raudenbush & Bryk, 2002) from a null model without any predictors. The variation associated with level 3 divided through the total variance is the "center effect."

Next, the second research question was explored to examine which center policies and characteristics might explain variance attributable to the center. In order to systematically test alternative assumptions about which center factors might influence outcomes, an iterative model building approach was used to test the significance of each center-level predictor entered into the model. In successive models, predictors were added in groups based on the level of supporting evidence from the literature—first, variables that have shown a relationship with outcome (initial severity at client and aggregated center level, treatment length at the client and aggregated center level, center average session frequency, session limits; Model 3), then a variable with limited and unexpected findings (student to staff ratio; Model 4), and finally, variables that have yet to be explored in the college counseling literature (APA accreditation status, IACS accreditation status, center size, NCAA athletic division, school type, and integration of additional services; Model 5).

The Akaike Information Criteria (AIC) index is reported, where lower values are indicative of a better model fit (McCoach & Black, 2008). Predictors were retained based on a combination of model fit and significance tests. If the addition of fixed effects significantly improved model fit (as determined by lower AIC values), predictors with significant *t* values were retained in the overall model. At the final step, likelihood-ratio tests (LRT) were used to compare model fit with and without additional predictors to assess whether adding predictors improved the overall model fit (Bolker et al., 2009). The LRT tests the null hypothesis that there is 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 also report the percent of variance explained by each predictor, to indicate the magnitude of the effect.

A final model (Model 6) contained only significant predictors retained from the other models. We report Kenward-Rogers adjusted p-values for all retained predictors in the final model, which corrects for inflated rates of type I errors (Kenward & Roger, 1997; 2009). To understand the magnitude of the effects, we calculated the variance explained by each predictor by subtracting the client-level intercept variance in a model with the additional predictor from the previous model without that predictor, and divided the remainder by the variance from the model without the predictor (Raudenbush & Bryk, 2002).

Chapter 3: Results

Depression

Parameter estimates and fit statistics for all models of depression are reported in Table 3. In Model 1, an intraclass correlation coefficient (ICC) was calculated for the variance accounted for by centers, before adding any predictors. Differences between centers accounted for 1.87% of the variance in client outcomes on depression, and Model 1 fit indices yielded AIC= 55642, -2LL= -27817.

In Model 2, differences between centers (controlling for client initial severity and treatment length) accounted for 2.63% of differences in outcome. The addition of client-level predictors significantly improved the model fit, AIC= 48922, -2LL= -24455. A client's treatment length significantly predicted their outcome (π_{2ij}=0.00, t = -9.30, p <.001), such that longer treatment was associated with a bigger decrease in scores (a positive outcome). Client initial severity on depression also predicted outcome (π_{1ij}=-0.38, t = -86.39, p <.001), such that higher depression symptomatology at baseline was associated with a decrease in latent depression scores post-treatment.

In Model 3, the first wave of center variables was added, and this represented a better fit than Model 2, AIC=48904, -2LL= -24442. In the presence of two significant client-level predictors, the average initial severity of all clients at a center significantly predicted outcome, (γ_{001}=-0.38, t = -5.20, p <.001), such that higher average symptomatology was associated with more change. Session frequency significantly predicted outcome, (γ_{002}=0.00, t = 1.97, p <.05), such that more frequent sessions were associated with more change. Not significantly associated with outcomes were center average treatment length (γ_{003}=0.01, t = -1.59, p =.12), and session limits (γ_{004}=0.00, t = 0.13, p =.90).

In Model 4, this did not fit the data better than Model 3, AIC= 48904, -2LL=-24443. Student-staff ratio was added and did not significantly predict outcomes in the presence of the other predictors, (γ_{005}=0.00, t = -0.58, p =.56).

In Model 5, this did not represent a better fit than Model 4, AIC=48909, -2LL=-24437. The third wave of variables were added, and none significantly predicted outcomes in depression: integrated health services (γ_{006}=-0.03, t = -0.95, p =.34), integrated learning services (γ_{007}=0.05, t = 0.99, p =.32), integrated drug and alcohol services (γ_{008}=-0.01, t = -0.41, p =.68), integrated psychiatric services (γ_{009}=0.01, t = 1.44, p =.15), APA internship accreditation status (γ_{0010}=-0.01, t = -0.36, p =.72), IACS accreditation status (γ_{0011}=0.01, t = 0.35, p =.73), center size (γ_{0012}=0.00, t = -1.65, p =.10), NCAA athletic division (γ_{0013}=0.02, t = 1.50, p =.14), and school type (γ_{0014}=0.04, t = 1.55, p =.12).

The final Model 6 retained only the significant predictors, including client initial severity on depression (π_{1ij}= -0.38, t = -86.40, p <.001) explaining 21.73% of the client-level variance, client treatment length (π_{2ij}=0.00, t = -9.28, p <.001) explaining 0.29% of client-level variance, center average initial severity (γ_{001}=-0.38, t = -5.20, p <.001) explaining 28.76% of the center-level variance, and average session frequency (γ_{002}=-0.35, t = -4.82, p <.001) explaining 3.95% of the center level variance. Taken together, Model 6 for depression improved overall fit (compared to the Model 1) as demonstrated by the significant reduction calculated from the LRT, -2LL= -24443, $\chi^2(4)$=6747.18, p <.0001, AIC= 48902. The final Model 6 is represented by the following equation, where DeprΔ_{tij} represents the pre- to post-treatment latent change score on depression.

$$
\begin{aligned}
&L1\_Client: Depr\Delta_{tij} = \pi_{0ij} + \pi_{1ij}(initial.sev) + \pi_{2ij}(tx.length.) + e_{tij} \\
&L2\_Therapist: \pi_{0ij} = \beta_{00j} + u_{0ij} \\
&L3\_Center: \beta_{00j} = \gamma_{000} + \gamma_{001}(avg.initial) + \gamma_{002}(seshfreq) + u_{00j}
\end{aligned}
$$

Anxiety

Parameter estimates and fit statistics for all models of anxiety are reported in Table 4. In Model 1, an ICC was calculated for the variance accounted for by centers before adding any predictors. Differences between centers accounted for 1.71% of the variance in client outcomes on anxiety, and Model 1 fit indices yielded AIC= 54224, -2LL= -27108.

In Model 2, differences between centers (controlling for client initial severity and treatment length) accounted for 2.06% of variance in outcomes. The addition of client-level predictors significantly improved the model fit, AIC= 50147, -2LL= -25068. Client treatment length significantly predicted outcome (π_{2ij}=0.00, t = - 9.82, p <.001), such that longer treatment was associated more change (a positive outcome), although the effect was small. Client initial severity on anxiety also predicted outcome (π_{1ij}=-0.28, t = -65.23, p <.001), such that higher symptomatology at baseline was associated with a decrease in symptoms post-treatment.

In Model 3, the first wave of center variables was added and this represented a better fit than Model 2, AIC=50142, -2LL= -25061. In the presence of two significant client-level predictors, the average initial severity of all clients at a center significantly predicted outcome, (γ_{001}=-0.29, t = -3.36, p <.001), such that higher average symptomatology was associated with more change. Center average treatment length significantly predicted outcome (γ_{003}=-0.01, t = -

1.98, p <.05), such that longer average treatment lengths at a center predicted more change. Not significantly associated with outcomes were average session frequency (γ_{002}=0.00, t = 1.22, p =.22) and session limits (γ_{004}=0.00, t = -0.39, p =.70).

Model 4 did not fit the data better than Model 3, AIC= 50142, -2LL=-25061. Student-staff ratio was added and did not significantly predict outcomes in the presence of the other predictors, (γ_{005}=0.00, t = -0.57, p =.57).

Model 5 not represent a better fit than Model 4, AIC=50144, -2LL=-25055. The third wave of variables were added, and none significantly predicted outcomes: integrated health services (γ_{006}=-0.03, t = -0.82, p =.41), integrated learning services (γ_{007}=0.03, t = 0.67, p =.50), integrated drug and alcohol services (γ_{008}=0.00, t = 0.19, p =.85), integrated psychiatric services (γ_{009}=0.01, t = 1.56, p =.12), APA accreditation status (γ_{0010}=-0.04, t = -1.86, p =.07), IACS accreditation status (γ_{0011}=0.03, t = 1.52, p =.13), center size (γ_{0012}=0.00, t = -1.01, p =.31), NCAA athletic division (γ_{0013}=0.01, t = 1.11, p =.27), and school type (γ_{0014}=0.02, t = 0.93, p =.36).

Model 6 retained only the significant predictors, including client initial severity on anxiety (π_{1ij}= -0.28, t = -65.22, p <.001) explaining 13.79% of the client-level variance, client treatment length (π_{2ij}=-0.01, t = -9.83, p <.001) explaining 0.32% of the client-level variance, center average initial severity (γ_{001}=-0.29, t = -3.28, p <.001) explaining 15.60% of the center-level variance, and center session severity (γ_{002}=-0.01, t = -3.28, p <.05) explaining 8.25% of the center-level variance. Taken together, Model 6 for anxiety improved overall fit (compared to the Model 1) as demonstrated by the significant reduction calculated from the LRT, -2LL= -25062,

$\chi^2(4)$= 4092.23, p <.0001, AIC= 50140. The final Model 6 is represented by the following equation, where AnxΔ_{tij} represents the pre- to post-treatment latent change score on anxiety.

$$
\begin{aligned}
&L1\_Client: Anx\Delta_{tij} = \pi_{0ij} + \pi_{1ij}(initial.sev) + \pi_{2ij}(tx.length.) + e_{tij} \\
&L2\_Therapist: \pi_{0ij} = \beta_{00j} + u_{0ij} \\
&L3\_Center: \beta_{00j} = \gamma_{000} + \gamma_{001}(avg.initial) + \gamma_{002}(avg.txlength) + u_{00j}
\end{aligned}
$$

Social Anxiety

Parameter estimates and fit statistics for all models of social anxiety are reported in Table 5. In Model 1, an ICC was calculated for the variance accounted for by centers, before adding any predictors. Differences between centers accounted for 1.25% of the variance in client outcomes on social anxiety, and Model 1 fit indices yielded AIC= 35838, -2LL= -17915.

In Model 2, differences between centers (controlling for client initial severity and treatment length) accounted for 1.50% of differences in outcome. The addition of client-level predictors significantly improved the model fit, AIC= 33216, -2LL= -16602. Client treatment length significantly predicted outcome (π_{2ij}=0.00, t = -12.02, p <.001), such that longer treatment was associated more reduction in symptom scores. Client initial severity on social anxiety also predicted outcome (π_{1ij}=-0.16, t = -50.28, p <.001), such that higher symptomatology at baseline was associated with more change.

In Model 3, the first wave of center variables was added and this represented a better fit than Model 2, AIC=33220, -2LL= -16600. In the presence of two significant client-level predictors, the only significant center-level predictor was the average treatment length (γ_{003}=0.00, t = -0.69, p <.001). None of the other center variables were significantly associated with client outcomes: average initial severity of all clients at a center (γ_{001}=-0.10, t = -1.69, p

=.09), average session frequency (γ_{002}=0.00, t = 0.48, p =.633), and session limits (γ_{004}=0.01, t = 1.03, p =.30).

Model 4 did not fit the data better than Model 3, AIC= 33219, -2LL=-16601. Student-staff ratio was added and did not significantly predict social anxiety outcomes in the presence of the other client level predictors, (γ_{005}=0.00, t = -0.40, p =.69). In this model, average treatment length was no longer significant (γ_{003}=0.00, t = -1.07, p =.28).

Model 5 did not represent a better fit than Model 4, AIC=33226, -2LL=-16598. The third wave of variables were added, and none significantly predicted outcomes: integrated health services (γ_{006}=-0.01, t = -0.67, p =.51), integrated learning services (γ_{007}=0.03, t = 0.95, p =.34), integrated drug and alcohol services (γ_{008}=-0.02, t = -1.03, p =.31), integrated psychiatric services (γ_{009}=0.00, t = 0.84, p =.40), APA accreditation status (γ_{0010}=-0.02, t = -1.17, p =.24), IACS accreditation status (γ_{0011}=0.03, t = 1.79, p =.08), center size (γ_{0012}=0.00, t = 0.02, p =.98), NCAA athletic division (γ_{0013}=0.01, t = 0.88, p =.38), and school type (γ_{0014}= -0.01, t = -0.72, p =.47).

Model 6 retained only the significant predictors, including client initial severity on social anxiety (π_{1ij}= -0.16, t = -50.28, p <.001) explaining 8.63% of the client-level variance, and client treatment length (π_{2ij}=0.00, t = -12.02, p <.001) explaining 0.52% of the client-level variance. The final Model 6 was identical to Model 2, and represented a better model fit than Model 1, as demonstrated by the significant reduction calculated from the LRT, -2LL= -16602, $\chi^2(2)$=2625.34, p <.0001, AIC= 33216. The final Model 6 is represented by the following equation, where SocanxΔ_{tij} represents the pre- to post-treatment latent change score on social anxiety.

$$L1\_Client: SocAnx\Delta_{tij} = \pi_{0ij} + \pi_{1ij}(initial.sev) + \pi_{2ij}(tx.length.) + e_{tij}$$
$$L2\_Therapist: \pi_{0ij} = \beta_{00j} + u_{0ij}$$
$$L3\_Center: \beta_{00j} = \gamma_{000} + u_{00j}$$

Eating Concerns

Parameter estimates and fit statistics for all models of eating are reported in Table 6. For Model 1, an ICC was calculated for the variance accounted for by centers, before adding any predictors. Differences between centers accounted for 1.24% of the variance in client outcomes on eating, and Model 1 fit indices yielded AIC= 50783, -2LL= -25388.

In Model 2, differences between centers (controlling for client initial severity and treatment length) accounted for 1.66% of differences in outcome. The addition of client-level predictors significantly improved the model fit, AIC= 45528, -2LL= -22758. Client treatment length significantly predicted eating concerns outcome (π_{2ij}=0.00, t = -3.33, p <.001), such that longer treatment was associated with a slight decrease in scores. Client initial severity on eating concerns also predicted outcome (π_{1ij}=-0.26, t = -75.78, p <.001), such that higher symptomatology at baseline was associated with a decrease in symptoms.

In Model 3, the first wave of center variables was added and this represented a better fit than Model 2, AIC=45503, -2LL= -22741. In the presence of two significant client-level predictors, the average initial severity of all clients at a center significantly predicted outcome, (γ_{001}= -0.40, t = -5.91, p <.001), such that higher average symptomatology was associated with a bigger decrease in symptoms. No significant associations with outcomes were found for center average treatment length (γ_{003}=0.00, t = -0.91, p =.37), center session frequency (γ_{002}=0.00, t = 1.13, p =.26), and session limits (γ_{004}=0.18, t =1.23, p =.22).

Model 4 did not fit the data better than Model 3, AIC= 45503, -2LL=-22744. Student-staff ratio was added and did not significantly predict outcomes in the presence of the other predictors, (γ_{005}=0.00, t = 0.66, p =.51).

Model 5 did not represent a better fit than Model 4, AIC=45514, -2LL=-22741. The third wave of variables were added, and none significantly predicted outcomes: integrated health services (γ_{006}= -0.01, t = -0.42, p =.68), integrated learning services (γ_{007}= 0.00, t = 0.06, p =.95), integrated drug and alcohol services (γ_{008}= 0.00, t = 0.15, p =.88), integrated psychiatric services (γ_{009}= 0.00, t = 0.81, p =.42), APA accreditation status (γ_{0010}= -0.01, t = -0.66, p =.51), IACS accreditation status (γ_{0011}= 0.04, t = 2.03, p <0.05), center size (γ_{0012}= 0.00, t = -0.41, p =.69), NCAA athletic division (γ_{0013}= 0.00, t = -0.08, p =.94), and school type (γ_{0014}= 0.00, t = -0.18, p =.86).

Model 6 retained only the significant predictors, including client initial severity on eating concerns (π_{1ij}= -0.26, t = -75.59, p <.001) explaining 17.39% of the client-level variance, client treatment length (π_{2ij}=0.00, t = 3.24, p <.01) explaining 0.04% of the client-level variance, and center average initial severity (γ_{001}= -0.38, t = -5.48, p <.001) explaining 28.15% of the center-level variance. IACS accreditation status was no longer significant and was dropped from the final model (γ_{0011}= 0.03, t = 1.90, p =.06). Model 6 represented a better model fit than Model 1, as demonstrated by the significant reduction calculated from the LRT, -2LL= -22742, $\chi^2(4)$=5291.22, p <.0001, AIC= 45500. The final Model 6 for eating concerns is represented by the following equation, where EatΔ_{tij} represents the pre- to post-treatment latent change score on eating concerns.

$$L1\_Client: Eat\Delta_{tij} = \pi_{0ij} + \pi_{1ij}(initial.sev) + \pi_{2ij}(tx.length.) + e_{tij}$$
$$L2\_Therapist: \pi_{0ij} = \beta_{00j} + u_{0ij}$$
$$L3\_Center: \beta_{00j} = \gamma_{000} + \gamma_{001}(avg.initial) + u_{00j}$$

Hostility

Parameter estimates and fit statistics for all models of hostility are reported in Table 7. In Model 1, ICC was calculated for the variance accounted for by centers, before adding any predictors. Differences between centers accounted for 1.88% of the variance in client outcomes on hostility, and fit indices yielded AIC= 58333, -2LL= -29162.

In Model 2, differences between centers (controlling for client initial severity and treatment length) accounted for 3.32% of differences in outcome. The addition of client-level predictors significantly improved the model fit, AIC= 46378, -2LL= -23183. Client treatment length significantly predicted outcome (π_{2ij}=-0.01, t = -12.50, p <.001), such that longer treatment was associated a decrease in symptom scores on hostility. Client initial severity on hostility also predicted outcome (π_{1ij}=-0.41, t = -121.68, p <.001), such that higher symptomatology at baseline was associated with a decrease in symptoms post-treatment.

In Model 3, the first wave of center variables was added and this represented a better fit than Model 2, AIC=46362, -2LL= -23171. In the presence of two significant client-level predictors, the average initial severity of all clients at a center significantly predicted outcome, (γ_{001}= -0.32, t =-4.81, p <.001), such that higher average symptomatology was associated with more change. Not significantly associated with outcomes were center average treatment length (γ_{003}=0.00, t = -0.11, p =.92), session frequency (γ_{002}=0.00, t = 1.50, p =.14), and session limits (γ_{004}=0.00, t = -0.04, p =.97).

Model 4 did not fit the data better than Model 3, AIC= 46360, -2LL=-23172. Student-staff ratio was added and did not significantly predict hostility outcomes in the presence of the other predictors, (γ_{005}=0.00, t = -0.83, p =.41).

Model 5 did not represent a better fit than Model 4, AIC=46369, -2LL=-23169. The third wave of variables were added, and none significantly predicted outcomes: integrated health services (γ_{006}= -0.02, t = -0.47, p =.64), integrated learning services (γ_{007}= 0.02, t = 0.30, p =.77), integrated drug and alcohol services (γ_{008}= -0.01, t = -0.63, p =.53), integrated psychiatric services (γ_{009}= 0.01, t = 0.73, p =.47), APA accreditation status (γ_{0010}= -0.03, t = -1.23, p =.22), IACS accreditation status (γ_{0011}= 0.03, t = 1.43, p =.16), center size (γ_{0012}= 0.00, t = 0.31, p =.76), NCAA athletic division (γ_{0013}= 0.02, t = 1.39, p =.17), and school type (γ_{0014}= 0.00, t = -0.17, p =.87).

Model 6 retained only the significant predictors, including client initial severity on hostility (π_{1ij}= -0.41, t = -121.68, p <.001) explaining 35.59% of the client-level variance, client treatment length (π_{2ij}= -0.01, t = -12.41, p <.001) explaining 0.59% of the client-level variance, and center average initial severity (γ_{001}= -0.32, t = -4.97, p <.001) explaining 24.89% of the center-level variance. It represented a better model fit than Model 1 and Model 2, as demonstrated by the significant reduction calculated from the LRT, -2LL= -23172-, $\chi^2(3)$=11980.44, p <.0001, AIC= 46358. The final Model 6 for hostility is represented by the following equation, where HostilityΔ_{tij} represents the pre- to post-treatment latent change score on hostility.

$$L1\_Client: Hostility\Delta_{tij} = \pi_{0ij} + \pi_{1ij}(initial.sev) + \pi_{2ij}(tx.length.) + e_{tij}$$
$$L2\_Therapist: \pi_{0ij} = \beta_{00j} + u_{0ij}$$
$$L3\_Center: \beta_{00j} = \gamma_{000} + \gamma_{001}(avg.initial) + u_{00j}$$

Academic Distress

Parameter estimates and fit statistics for all models of academic distress are reported in Table 8. In Model 1, an ICC was calculated for the variance accounted for by centers before adding any predictors. Differences between centers accounted for 1.90% of the variance in client outcomes on academic distress, AIC= 35692, -2LL= -17842.

In Model 2, differences between centers (controlling for client initial severity and treatment length) accounted for 2.34% of differences in outcome. The addition of client-level predictors significantly improved the model fit, AIC= 31467, -2LL= -15728. Client treatment length significantly predicted outcome (π_{2ij}= 0.00, t = -4.37, p <.001), such that longer treatment was associated with a slight decrease in academic distress scores. Client initial severity on academic concerns also predicted outcome (π_{1ij}= -0.30, t = -67.57, p <.001), such that higher symptomatology at baseline was associated with a decrease in academic distress scores post-treatment.

In Model 3, the first wave of center variables was added this represented a better fit than Model 2, AIC=31438, -2LL= -15709. In the presence of two significant client-level predictors, the average initial severity of all clients at a center significantly predicted outcome, (γ_{001}= -0.42, t = -5.51, p <.001), such that higher average symptomatology was associated with more change. Session frequency significantly predicted outcome, (γ_{002}= 0.00, t = 2.50, p <.05), such that more frequent sessions were associated with a decrease in academic distress scores post-treatment. Not significantly associated with outcomes were center average treatment length (γ_{003}=0.00, t = -1.00, p =.32), and session limits (γ_{004}= -0.03, t = -1.95, p =.05).

Model 4 did not represent a better fit for the data compared to model Model 3, as indicated by larger fit indices: AIC= 31439 and -2LL=-15710. Student-staff ratio was added and did not significantly predict outcomes in the presence of the other predictors, (γ_{005}=0.00, t = -1.38, p =.17).

Model 5 did not represent a better fit than Model 4, AIC=31441, -2LL=-15704. The third wave of variables were added, and none significantly predicted outcomes: integrated health services (γ_{006}= -0.01, t = -0.25, p =.80), integrated learning services (γ_{007}= 0.03, t = 0.84, p =.41), integrated drug and alcohol services (γ_{008}= 0.00, t = -0.25, p =.81), integrated psychiatric services (γ_{009}= 0.00, t = 0.54, p =.59), APA accreditation status (γ_{0010}= -0.02, t = -1.49, p =.14), IACS accreditation status (γ_{0011}= 0.01, t = 0.45, p =.66), center size (γ_{0012}= 0.00, t = -1.78, p =.08), NCAA athletic division (γ_{0013}= 0.01, t = 1.46, p =.15), and school type (γ_{0014}= 0.01, t = 0.65, p =.52).

Model 6 retained only the significant predictors, including client initial severity on academic distress (π_{1ij}= -0.30, t = -67.57, p <.001) explaining 14.36% of the client-level variance, client treatment length (π_{2ij}= 0.00, t = -4.49, p <.001) explaining 0.06% of the client-level variance, center average initial severity (γ_{001}= -0.42, t = -5.74, p <.001) explaining 34.11% of the center-level variance, and session frequency (γ_{002}= 0.01, t = 2.59, p <.01) explaining 6.61% of the center-level variance. Model 6 represented a better model fit than Model 1, as demonstrated by the significant reduction calculated from the LRT, -2LL= -15711,

$\chi^2(4)$=4261.75, p <.0001, AIC= 31438. The final Model 6 is represented by the following equation, where AcademicΔ_{tij} represents the pre- to post-treatment latent change score on academic distress.

$$\begin{aligned}
&L1_Client : Academic_{tij} = p_{0ij} + p_{1ij}(initial.sev) + p_{2ij}(tx.length.) + e_{tij} \\
&L2_Therapist : p_{0ij} = b_{00j} + u_{0ij} \\
&L3_Center : b_{00j} = g_{000} + g_{001}(avg.initial) + g_{002}(seshfreq) + u_{00j}
\end{aligned}$$

Alcohol Use

For this final series of models on alcohol use, parameter estimates and fit statistics are reported in Table 9. In Model 1, an ICC was calculated for the variance accounted for by centers before adding any predictors. Differences between centers accounted for 0.89% of the variance in client outcomes on alcohol, and Model 1 fit indices yielded AIC= 47706, -2LL= -23849.

In Model 2, differences between centers (controlling for client initial severity and treatment length) accounted for 1.83% of differences in outcome. The addition of client-level predictors significantly improved the model fit, AIC= 35419, -2LL= -17704. Client treatment length significantly predicted outcome (π_{2ij}= 0.00, t = -7.34, p <.001), such that longer treatment was associated more change (a positive outcome). Client initial severity on alcohol use also predicted outcome (π_{1ij}= -0.40, t = -124.62, p <.001), such that higher symptomatology at baseline was associated with a decrease in latent scores on alcohol use.

In Model 3, the first wave of center variables was added and this represented a better fit than Model 2, AIC=35397, -2LL= -17688. In the presence of two significant client-level predictors, the average initial severity of all clients at a center significantly predicted outcome, (γ_{001}= -0.29, t = -5.24, p <.001), such that higher average symptomatology was associated with more change. Not significantly associated with outcomes were: session frequency center (γ_{002}= 0.00, *t* = 0.03, *p* =.98), average treatment length (γ_{003}=0.00, *t* = 0.78, *p* =.44), and session limits (γ_{004}= 0.01, *t* = 0.50, *p* =.62).

Model 4 did not fit the data better than Model 3, AIC= 35393, -2LL=-17689. Student-staff ratio was added and did not significantly predict alcohol outcomes in the presence of the other predictors, (γ_{005}=0.00, *t* = -0.41, *p* =.68).

Model 5 did not represent a better fit than Model 4, AIC=35403, -2LL=-17686. The third wave of variables were added, and none significantly predicted alcohol outcomes: integrated health services (γ_{006}= -0.02, *t* = -0.87, *p* =.39), integrated learning services (γ_{007}= -0.01, *t* = -0.21, *p* =.83), integrated drug and alcohol services (γ_{008}= 0.00, *t* = -0.31, *p* =.76), integrated psychiatric services (γ_{009}= 0.00, *t* = 0.37, *p* =.71), APA accreditation status (γ_{0010}= -0.02, *t* = -0.96, *p* =.34), IACS accreditation status (γ_{0011}= 0.02, *t* = 1.50, *p* =.14), center size (γ_{0012}= 0.00, *t* = 0.16, *p* =.87), NCAA athletic division (γ_{0013}= 0.00, *t* = 0.00, *p* =1.0), and school type (γ_{0014}= -0.02, *t* = -1.54, *p* =.13).

Model 6 retained only the significant predictors, including client initial severity on alcohol use (π_{1ij}= -0.40, *t* = -124.62, *p* <.001) explaining 36.25% of the client-level variance, client treatment length (π_{2ij}= 0.00, *t* = -7.20, *p* <.001) explaining 0.21% of the client-level variance, and center average initial severity (γ_{001}= -0.30, *t* = -5.79, *p* <.001) explaining 30.45% of the center-level variance. It represented a better model fit than Model 1 as demonstrated by the significant reduction calculated from the LRT, -2LL= -17689, χ^2(3)=12321.20, *p* <.0001, AIC=

35391. The final Model 6 for alcohol is represented by the following equation, where AlcoholΔ_{tij} represents the pre- to post-treatment latent change score on alcohol use.

$$
\begin{aligned}
&L1\_Client : Alcohol\Delta_{tij} = \pi_{0ij} + \pi_{1ij}(initial.sev) + \pi_{2ij}(tx.length.) + e_{tij} \\
&L2\_Therapist : \pi_{0ij} = \beta_{00j} + u_{0ij} \\
&L3\_Center : \beta_{00j} = \gamma_{000} + \gamma_{001}(avg.initial) + u_{00j}
\end{aligned}
$$

Chapter 4: Discussion

The goal of the current study was to investigate two main research questions: (1) what effect do counseling centers have on the outcomes of their psychotherapy clients, and (2) what characteristics and policies of these counseling centers might explain such an effect. As a whole, the analyses did not reveal strong center effects. Within the context of the effects observed, however, support was found for several of the hypotheses and predictions about specific center policies or characteristics that might affect outcomes.

Center effects

After exploring the first research question, centers explained a relatively small amount of the differences between client outcomes across all symptom domains. Once we controlled for the unique contribution of client variables that been previously linked with outcome (as part of Model 2 analyses), centers explained on average 2.19% of the variance in outcomes, ranging from 1.50% (social anxiety) to 3.32% (hostility). Perhaps the most parsimonious interpretation of these findings is that where clients are seen (within the CCMH infrastructure) does not truly matter. Indeed, it is possible that there is a fairly high level of homogeneity in terms of the standard of care and operational functioning among the university counseling centers that have chosen to join and remain in this PRN. Within this context, who the clients see as their therapist may be more important than the setting of the counseling centers where the treatment is provided. This is supported, at least for some measures of outcome, by a recent investigation also conducted in the CCMH infrastructure (Youn, Castonguay, Janis, Hayes, & Locke, submitted for publication). They found that while the therapist effects ranged between 1 to 4% of the variance across most domains of the CCAPS, they explain approximately 9% and 16% of rate of change for hostility and alcohol use respectively. Consistent with psychotherapy research in general (Bohart & Wade, 2013), it may also be that client characteristics, during and outside of therapy sessions, mostly accounted for the outcome changes in this study. Wampold (2017) estimated that therapy factors account for 13% of variance in outcomes, and the rest is likely due to the client. This is in line with another study's findings within the CCMH framework. Lefevor et al. (2017) demonstrated that when predicting depression scores, center accounted for 2.1% of the total variance, whereas 64% of the variance was accounted for by differences between clients within centers. Moreover, it is possible that what happens between the client and the therapist during treatment is also more relevant to client's improvement than contextual variables related to the counseling center setting. To date, however, no process studies have been completed in the CCMH PRN infrastructure.

It is also possible that center effects are in fact large and significant, but that effect is not being detected due to the current study's choice of dependent variable. Within this study, it may be that the calculated outcome (latent change score on a given subscale) is not capturing the meaningful change that would naturally take place during a client's treatment course. In the current study, clients were assessed on all subscales, but some of those symptom domains may not capture what they were addressing in treatment. A client treated for depression, for example, might not be expected to show substantial change on the hostility or eating subscales.

Furthermore, it is conceivable that that center effects are in fact real and significant, but that they are not captured only, or primarily, by symptom changes of clients. Other studies have suggested using rate of change as a way to operationalize client outcome instead of, or in addition to, symptom change (Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009; Youn et al., submitted for publication). Studies have demonstrated that clients may achieve similar raw symptom changes, but some respond to treatment more rapidly than others (Bohart & Wade, 2003). This has implications for resource utilization within centers, and therefore would be a fruitful avenue for researchers and clinicians to pursue.

Center policies and characteristics affecting outcome

After exploring the second question, client initial severity and treatment length were significantly associated with outcomes, illustrating that variables which tap into client factors continue to be robust and significant (Bohart & Wade, 2003). These client level predictors explained a range from 0.06% (treatment length predicting academic change) to 35.25% (initial severity predicating alcohol change) of the client-level variance in outcomes, demonstrating the potential for both weak and strong effects. Center-level versions of these variables (initial severity and treatment length aggregated at the center level) were included to statistically disentangle client and center effects, and were significantly associated with outcome on nearly all of the subscales (except social anxiety).

The additional center predictors introduced to the models were intended to represent center policies and characteristics, and were not significantly associated with the majority of outcomes. These included student to staff ratio, APA internship accreditation status, IACS accreditation status, center size, athletic division, integrated services (health, learning, drug and alcohol, or psychiatric), session limits, and school type (public/private).

A series of patterns describe which predictors explained variance accounted for by center and which did (see Table 10 for a summary of predictor findings). One theme that emerged highlights the heterogeneity that exists not only between centers, but likely also within centers.

Despite centers reporting explicit policies on paper, it seems likely that there is a discrepancy between what centers indicate is their explicit policy on paper, and what they do on a daily basis with their clients. One possible explanation is that either therapists or administrators are making exceptions to the rules. For instance, the current study found that the presence of session limits was not significantly associated with outcomes. However, one survey reported that only 82% of clients seen at counseling centers were actually subjected to the limits (Gyorky et al., 1994). With therapists and clients operating outside of organizational policy, the within center variation may wash out a detectable center effect. Future studies might address whether there is a data-based difference in the number of sessions used when comparing schools with and without reported session limits. If practice follows policy, we would expect there to be a clear difference. If, however, exceptions are made and policies are not strictly followed, the data would show that the presence of a reported "policy" (e.g. for session limits) does not capture the reality of a center's operations. One way to measure this would be to survey counseling centers about policy adherence, gathering information about their explicit policies as well exceptions made under whose authority, how often, and in what circumstances. Within-center heterogeneity also extends to policies like session frequency, where there may be a large within center difference, due to therapists or center administrators making exceptions for special cases (e.g. allowing more frequent sessions for clients high in suicidality, student athletes, and trauma or assault victims).

Another theme that emerged suggested that policies purported to be best practices were not supported by the current study. First, IACS and APA accreditation statuses did not predict better or worse outcomes. While these are designations that are much sought after (e.g. over 400 internship sites have achieved APA accreditation), no studies have looked at whether these accreditations translate into a better quality of care or better client outcomes. This is not to say that those accreditations have no benefit, but instead that there are perhaps qualities of counseling centers that are not tied to meet accreditation guidelines that explain more about positive or negative outcomes (e.g. therapist burnout, organization climate, financial resources).

Similarly, center average treatment length was not significantly associated with change scores on six out of the seven subscales (excepting anxiety). Results here may parallel the “good enough” model, where a faster rate of change has been observed in shorter treatments compared to longer ones, leading to the same magnitude of change from pre- to post-treatment but fewer sessions utilized (Barkham et al., 2006). This has been explained by client and therapist working together to make the best from what they get in terms of service conditions. So, a center effect may be washed out if clients and therapists will adapt to whatever treatment lengths are logistically feasible within centers.

In this same vein, it is interesting to note that in the current atmosphere of increased movements towards integration, schools indicating the presence of integrated services did not predict better outcomes. This was true even for domain-congruent pairings, such as change on academic distress score and integrated learning services, or change on alcohol use scores and the presence of integrated drug and alcohol services. This suggests that these 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 with individual treatment for alcohol use, but then refer students to specific substances abuse groups or workshops. If they administer the CCAPS at the beginning of treatment but systematically do not to administer the CCAPS for those workshops, a client’s change score would only reflect the start of their treatment. Thus the data would not capture gains and changes facilitated by integrated services outside of individual treatment. Center size (defined by the number of clients presenting to a center) was also found to be unassociated with client outcomes. This suggests that while increasingly numbers of students seeking treatment may strain resources, it may be more important to understand how centers are equipped to handle incoming students. This is not captured by looking at the number of clients alone. For an additional meaningful variable, studies might use percent of the study body being served in the counseling center, to capture center size relative to the institution.

Student to staff ratio was not significantly associated with outcomes, which replicates previous findings (Elreda, 2014) but seems to be inconsistent with IACS recommendations to the field. They warning against large ratios, where the working assumption is that with too high of a student to staff ratio the counseling center will not be able to meet client demand and thus clients and treatment will suffer as a result. One possible explanation for why ratios were not linked with outcomes in the current study is that some schools within the CCMH PRN may have additional resources outside of the main counseling center (e.g. a women's or LGBTQ center), thus diffusing the mental health needs across multiple places and reducing burden on the counseling center. Additional on-campus mental health resources are not captured in the current data, which introduces a potential confound. If centers have a large student to staff ratio, but immediately refer students with particular concerns out to other on-campus resources, the high student to staff ratio does not present a problem for long-term treatment. Those students who were referred out inflate the ratio, but would not tax or burden the center in the way that IACS suggests (e.g. adding to a long waitlist, forcing centers to institute session limits, contributing to high, long-standing caseloads). That this and the above predictors are not significant marks a departure from what are thought of as positive center policies and characteristics, and therefore more work is needed to parse apart these effects.

Lastly, a theme emerged such that institutional variables where the center was located (school type and athletic division) were not significantly predictive of outcome for any subscales. This suggests that while public versus private, and Division I, II, and III schools represent different academic and social contexts, this may not be as relevant with regards to students' mental health care. This may be a situation where more variance is explained at the client level, due to individual differences in mental health and attitudes toward treatment.

Another possible explanation as to why centers account for such a small part of the variance is that it the "center effect" is better explained by contextual factors not examined in the current study. Data could be captured on other organizational-level variables that might strengthen the center effect, such as referral policies, fee structures, appointment reminder systems, on-campus perception of the counseling center, and center supervision practices. Researchers could also turn to other contextual variables on a more macro scale. Studies in community mental health centers have argued for the importance of incorporating variables such as the level of poverty surrounding the neighborhood where the center is located to understand as much as possible about the treatment context setting (Delgadillo, Asaria, Ali, & Gilbody, 2016). Future studies could extend this into the college counseling context by assessing variables such as average socioeconomic status of students attending the institution, and incorporating other institutional variables such as location (e.g. city, rural), academic rigor, and attitude among students towards seeking treatment.

Finally, it is interesting to note that although the effects are small, results demonstrated convergence across subscales regarding which predictors explain outcomes. With only a few exceptions, the predictors showed the same patterns across domains of functioning, including the direction and (almost always) magnitude of the effect. This suggests that it is perhaps more important for centers to attend to other client characteristics (such as initial severity), regardless of what domain of concerns they bring to treatment. However, there were slight differences between subscales with regards to how much of a contribution center makes to explaining outcome variance. For example, centers explain 1.5% of the variance in social anxiety, but 3.32% of the variance in hostility and 2.63% of the variance in depression. Depression and hostility represent domains with higher likelihood of threat-to-self and threat-to-others behaviors, so the way centers manage and mitigate those threats (e.g. with behavioral response teams or additional resources) may account for why those center effects are higher.

Clinical and administrative implications

Within this variance partitioning approach, researchers can better understand at which level to target efforts to improve treatment effectiveness. While the effects at the center level found in the current study are small, they merit further research 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 increasing demand (Xiao et al., 2017). But how to advocate and where to spend those resources? It is in the best interests of counseling centers to be conducting research on these issues at their own site, as well as observing research findings distributed from a national or international level. The current study found that some seemingly positive center policies (e.g. no session limits) and characteristics (e.g. low student to staff ratio) were not in fact predictive of better outcomes, which could lead to the assumption that they may not be best practices. However, on a center by center basis, these could be particularly effective for a given local population. Research should continue to be conducted nationally, but also on a local scale so that centers can determine their own best administrative practices.

Another potential lesson is that blanket policies for all clients based on national recommendations (e.g. “sessions should occur once a week”) may not be the best strategy. The current study demonstrated the robustness of client factors in predicting outcomes, so centers could consider tailoring policies to meet the needs of specific clients, based on client characteristics at baseline. For example, one study within the CCMH framework suggested that more severe clients require additional services to make the same treatment gains as less severe clients. Specifically, researchers recommended that centers focus on individuals with highly elevated scores at intake, to help bring those clients’ post-treatment functioning to within normal range (McAleavey, 2017). Other researchers have argued the need for a more personalized approach to therapy, using risk stratification (based on variables like disability, employment, age, functional impairment, depression, and outcome expectancy) and client profiling to make treatment recommendations and guidelines, rather than relying on a rigid clinic policy (Delgadillo, Moreea, & Lutz, 2016). This also reflects an ongoing movement in the United Kingdom and other parts of Europe, called the “stepped care” model, which is characterized by the use of series of assessments and increasingly intense interventions to address a client’s specific needs in the most effective yet least costly manner (London School of Economics and Political Science, 2006).

Limitations

The present study included a large, naturalistic, multisite sample with a high number of units at each level of examination. By investigating center policies and characteristics in relation to therapeutic outcomes, this study has replicated and extended several findings in this area. However, there are limitations to the generalizability of our results. Our sample was comprised of mostly young adults who received treatment at their college or university's counseling center. Although the sample was large and the participants represent a variety of demographic groups, it is not a truly representative sample of treatment-seeking adults. Consequently, interpretation of these findings may be restricted to the experiences of students at college counseling centers. Findings about organizational policies and contextual characteristics might vary outside of this context (e.g. in community mental health centers, settings outside of the United States, or with adults of different ages).

Another limitation manifested during the data reduction process whereby a varying number of clients were lost at each step; most sizably by requiring that clients have at least two valid CCAPS administrations. Nearly 105,000 clients had one CCAPS administration but did not have two, and were therefore removed during the cleaning process because change scores could not be calculated. This is problematic because it further restricts the representativeness of the dataset and somewhat limits the conclusions that can be drawn. Additionally, there may be something systematic at a center level that explains why so many clients did not meet inclusion criteria. For instance, centers may routinely administer the CCAPS only at intake. Research has shown that providing consistent outcome monitoring feedback to therapists is helpful (Gondek et al., 2016), so centers that choose not to use the CCAPS in this manner may not be adhering to best practices.

Future directions

Future studies should consider expanding the independent variables that can help explain a center effect, as well as the dependent variables used to assess outcome and change. With regard to expanding the predictors, analyses could include interaction terms to explore the relationship between other variables. For example, including a center size by school enrollment interaction as proxy for the percent of the student body being served in the clinic. This may capture something different than student-staff ratio, and could speak to the level of demand and mental health symptoms of a particular student population. Future studies could also consider cross-level interaction terms, like size of center by initial severity. It could be that larger centers are better equipped to handle more severe, high resource-utilizing cases, but it could also be that centers cannot accommodate the extra needs when working with many other clients, and so these more severe clients get lost in the shuffle. Studies could parse this apart in future analyses.

Another line of studies could use rate of change as the outcome measure, instead of symptom change. Some studies have shown that clients might experience the same amount of symptom change, but some achieve those gains much quicker than others (Bohart & Wade, 2003). This has implications for future research designs, but also on impacts counseling centers a more practical level. If clients are able to achieve change more quickly (or slowly), this affects the pool of available therapy hours at a given center. These two last points (about interaction terms and outcome measures) suggest that perhaps that we need to be more specific in terms of what we investigate and how we investigate it to have a fairer and more realistic assessment of potentially meaningful contextual variables.

Researchers could also address the issue of outcome variables not capturing what clients addressed in treatment by only looking at relevant measures for each client. This could be accomplished by using diagnostic data, so if a client was diagnosed with primary Major Depressive Disorder the outcome variable selected would be the depression subscale. Future studies could also use specific symptom domains of concern, as identified by the therapist. Within the CCMH practice research network after initial consultation, clinicians routinely complete a checklist of concerns they believe are distressing for the client (e.g. anxiety, academic concerns, relationship problem, stress). This Clinician Index of Client Concerns (CLICC) could be used to pair concerns identified on the CLICC with the appropriate symptom domain on the CCAPS, thus ensuring a conceptual match across what was covered in treatment and what is being used as the outcome variable (CCMH, 2012).

Future studies should also consider doing all predictions within the SEM framework, instead of using SEM to generate latent change scores and then plugging those scores into a regular multilevel regression model. By conducting all analyses within the SEM framework, no error is lost and represents a purer prediction model. At the time of this study, three-level multilevel SEM is not supported by the lavaan package used in the current project, but will likely be possible in the future and could therefore be utilized to strengthen the statistical analyses.

Because CCMH gathers longitudinal data, future studies could also look at centers who have changed policies from one year to the next, to see if there are changes associated with outcomes in a more casual manner. In these quasi-experimental studies, researchers could examine centers whose policies have changed from one year to the next, and measure any changes in client outcomes that follow. For example, if a center began instituting fees for therapy, they may see an increase in their attendance rates in subsequent years.

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 counseling center policies/characteristics 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 a more or less frequent treatment group. Empirically testing these policies would facilitate progress towards identifying best clinical and administrative practices.

References

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology, 20,* 93-114.

Anderson, T., McClintock, A. S., Himawan, L., Song, X., & Patterson, C. L. (2016). A prospective study of therapist facilitative interpersonal skills as a predictor of treatment outcome. *Journal of Consulting and Clinical Psychology, 84*(1), 57-66. doi: http://dx.doi.org/10.1037/ccp0000060

Balasubramanian, B. A., Cohen, D. J., Jetelina, K. K., Dickinson, M., Davis, M., Gunn, R., … Green, L. A. (2017). Outcomes of integrated behavioral health with primary care. *Journal of the American Board of Family Medicine, 30*(2), 130-139. doi: 10.3122/jabfm.2017.02.160234

Baldwin, S. A. & Imel, A. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258-297). New York: Wiley.

Baldwin S. A., Wampold, B. E., & Imel, A. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*(6), 842-852. doi: 10.1037/0022-006X.75.6.842

Barkham, M., Connell, J., Stiles, W., Miles, J. N., Margison, F., Evans, C.,et al. (2006). Dose–effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74,* 160–167.

Bell, B. A., Morgan, G. B., Schoeneberger, J. A., Kromrey, J. D., & Ferron, J. M. (2014). How low can you go? An investigation of the influence of sample size and model complexity on point and interval estimates in two-level linear models. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 10*(1), 1-11. doi:10.1027/1614-2241/a000062.

Bohart, A. C. & Wade, A. G. (2013). The client in psychotherapy. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 219-257). New York: Wiley.

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

Castonguay, L. G., & Hill, C. E. (2017). *How and why are some therapists better than others? Understanding therapist effects.* Washington, D.C.: American Psychological Association.

Castonguay, L. G., Locke, B. D., & Hayes, J. A. (2011). The Center for Collegiate Mental Health: An example of a practice-research network in university counseling centers. *Journal of College Student Psychotherapy, 25*(2), 105-119. doi: 10.1080/87568225.2011.556929

Castonguay, L. G. & Beutler, L. E. (2006). Principles of therapeutic change: A task force on participants, relationships, and techniques factors. *Journal of Clinical Psychology, 62*(6), 631-638. DOI: 10.1002/jclp.20256

Clarkin, J. F. & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change* (5^{th} ed., pp 194-226). Hoboken, NJ: Wiley.

Center for Collegiate Mental Health. (2012, July). 2012 Standardized Data Set (Publication No. STA 12–01 MPC129989).

Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thornton, J. A., Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy*, *52*(3), 337-345. http://dx.doi.org/10.1037/pst0000015

Cologon, J., Schweitzer, R. D., King, R., & Nolte, T. (2017). Therapist reflective functioning, therapist attachment style and therapist effectiveness. *Administration and Policy in Mental Health and Mental Health Services Research, 44*, 614-625. DOI 10.1007/s10488-017-0790-5

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology, 67*(2), 184-192, DOI: 10.1002/jclp.20754

Crits-Christoph, P., Gibbons, M. C. G., & Mukherjee, D. (2013). Psychotherapy process-outcome research. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 298-340). New York: Wiley.

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. doi: 10.1080/10503300903376320.

Delgadillo, J., Asaria, M., Ali, S., & Gilbody, S. (2016). On poverty, politics and psychology: The socioeconomic gradient of mental health utilisation and outcomes. *The British Journal of Psychiatry, 209*, 429-430. doi: 10.1192/bjp.bp.115.171017.

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

Elkeblad, A., Flakenström, F., Holmqvist, R. (2016). Reflective functioning as predictor of working alliance and outcome in the treatment of depression. *Journal of Consulting and Clinical Psychology, 84*(1), 67-78. doi: 10.1037/ccp0000055

Elreda, D. (2014). *Examining how the ratio of students on campus to counseling center staff relates to dosage and counseling center treatment effectiveness*. (Unpublished doctoral dissertation). The Pennsylvania State University, State College, PA.

Engel, S. G., Johnson, C., Powers, P. S., Crosby, R. D., Wonderlich, S. A., Wittrock, D. A., & Mitchell, J. E. (2003). Predictors of disordered eating in a sample of elite Division I college athletes. *Eating Behaviors, 4*, 333-343. doi:10.1016/S1471-0153(03)00031-X

Erekson, D. M., Lambert, M. J., Egget, 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. http://dx.doi.org/10.1037/a0039774

Falkenstrom, F., Granström, F. & Holmqvist, R. (2014). Working alliance predicts psychotherapy outcome even when controlling for prior symptom improvement. *Psychotherapy Research, 24*(2), 146-159. http://dx.doi.org/10.1080/10503307.2013.847985

Falkenstrom, F., Grant, J. & Holmqvist, R. (2018). Review of organizational effects on the outcome of mental health treatments. *Psychotherapy Research, 28*(1), 76-90.

Gage, L. A. & Gyorky, Z.K. (1990). Identifying appropriate clients for time-limited counseling. *Journal of College Student Development, 31*, 476-477.

Gallagher, R. P. (2005). National survey of counseling center directors. The International Association of Counseling Services, Inc. Monograph Series Number 8O. Retrieved from http://www.collegecounseling.org/pdf/2005_survey.pdf.

Goates-Jones, M. & Hill, C. E. (2008). Treatment preference, treatment-preference match, and psychotherapist credibility: Influence on session outcome and preference shift. *Psychotherapy, 45*(1), 61-74. http://dx.doi.org/10.1037/0033-3204.45.1.61

Gondek, D., Edbrooke-Childs, J., Fink, E., Deighton, J., Wolpert, M. (2012). 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.

Gyorky, Z. K., Royalty, G. M., Johnson, D. H. (1994). Time-limited therapy in university counseling centers: Do time-limited and time-unlimited centers differ? *Professional Psychology: Research and Practice, 25*(1), 50-54. doi: 10.1037/0735-7028.25.1.50

Hansen, N. B. & Lambert, M. J. (2003). An evaluation of the dose-response relationship in naturalistic settings using survival analysis. *Mental Health Services Research, 5*, 1-11.

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology Science and Practice, 9*(3), 329-343.

Ilagan, G., Vinson, M. L., Sharp, J. L., Ilagan, J. & Oberman, A. (2015). Exploring outcomes and initial self-report of client motivation in college counseling center. *Journal of American College Health, 62*(3). doi: 10.1080/07448481.2014.1003379

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53,* 983-997. doi:10.2307/2533558

Kenward, M. G., & Roger, J. H. (2009). An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics & Data Analysis, 53,* 2583-2595. doi:10.1016/j.csda.2008.12.013

Kessler, R. C. (1977). Use of change scores in criteria in longitudinal survey research. *Quality and Quantity, 11,* 43-66.

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.

King, L. A., King, D. W., McArdle, J. J., Saxe, G. N., Doron-LaMarca, S., Orazem, R. J. (2006). Latent difference score approach to longitudinal trauma research. *Journal of Traumatic Stress, 19*(6), 771-785.

Lacour, M. M. & Carter, E. F. (2002). Challenges of referral decisions in college counseling. *Journal of College Student Psychotherapy, 17*(2), 39-52. doi: 10.1300/J035v17n02_04

Lambert, M.J. & Ogles, B.M. (2004). The efficacy and effectiveness of psychotherapy. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (pp. 139–193). New York: Wiley.

Lawe, C.F., Penick, J.M., Raskin, J.D., & Raymond, V.V. (1999). Influences on decisions to refer at university counseling centers. *Journal of College Student Psychotherapy, 14 (1)*, 59-68.

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

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

Locke, B. D., McAleavey, A. A., Zhao, Y., Lei, P., Hayes, J. A., Castonguay, L. G., Li, H., Tate, R., Lin, Y. (2012). Development and initial validation of the Counseling Center Assessment of Psychological Symptoms-34. *Measurement and Evaluation in Counseling and Development*, *45*(3).

London School of Economics and Political Science. Centre for Economic Performance. Mental Health Policy Group (2006). The depression report: A new deal for depression and anxiety disorders [online]. London, UK: LSE Research Online.

Mahon, M., Laux, J. M., Ritchie, M. H., Piazza, N. J., Tiamiyu, M. F. (2015). Brief therapy at a university counseling center: Working alliance, readiness to change, and symptom severity. *Journal of College Counseling*, *18*(3), 233-243. doi: 10.1002/jocc.12017

Marsh, C. N. & Wilcoxon, S. A. (2015). Underutilization of mental health services among college students: An examination of system-related barriers. *Journal of College Student Psychotherapy, 29*(3). doi: 10.1080/87568225.2015.1045783.

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. doi: 10.1080/10503307.2014.883652.

McAleavey, A. A., Youn, S., Xiao, H., Castonguay, L. G., Hayes, J., Locke, B. D. (2017). Effectiveness of routine psychotherapy: Method matters. (in press).

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605.

McArdle, J. J. & Ghisletta, P. (2012). Latent curve models and latent change score models estimated in R. *Structural Equation Modeling, 19*(4), 651-682.

McCoach, D. B., & Black, A. C. (2008). Model fit in multilevel modeling. *Multilevel analysis of educational data.* Greenwich, CT: Information Age Publishing.

McLeod, J. (2013). Qualitative research: Methods and contributions. In M. J. Lambert (Ed.), *Bergin & Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 49-84). New York: Wiley.

Norcross, J. C. & Lambert, M. J. (2011). Evidence-based therapy relationships. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based responsiveness*, *2*, 3-23. New York: Oxford University Press.

Owen, J., Devdas, L. & Rodolfa, E. (2007). University counseling center off-campus referrals: An exploratory investigation. *Journal of College Student Psychotherapy, 22*(2), 13-29. doi: 10.1300/J035v22n02_03.

Orlinsky, D. E., Ronnestad, M. H., Willutzki, U. (2004). Fifty years of psychotherapy process-outcome research: Continuity and change. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change (5th ed.).* Hoboken, NJ: Wiley.

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

Powell, B. J. & Beidas, R. S. (2016). Advancing implementation research and practice in behavioral health systems. *Administration and Policy in Mental Health and Mental Health Services Research, 43*, 825-833. doi: 10.1007/s10488-016-0762-1

Quintana, S.M., Yesenosky, J., Kilmartin, C., & Macias, D. (1991). Factors affecting referral decisions in a university counseling center. *Professional Psychology: Research and Practice*, *22 (1),* 90-97.

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

Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical Linear Models: Applications and data analysis methods (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Reetz, D. R., Bershad, C., LeViness, P., & Whitlock, M. (2016). The Association for University and College Counseling Center Directors annual survey. Retrieved from http://www.aucccd.org/director-surveys- public.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36.

SAMHSA-HRSA Center for Integrated Health Solutions. (2011). *Integrated care models.* Retrieved from www.integration.samhsa.gov/integrated-care-models.

Saxon, D. & Barkham, M. (2012). Patterns of therapist variability: Therapist effects and the contribution of patient severity and risk. *Journal of Consulting and Clinical Psychology*, *80*, 535–546.

Schiefele, A., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., … 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,* 598-613.

Shiner, B., Westgate, C. L., Harik, J. M., Watts, B. V., & Schnurr, P. P. (2017). Effect of patient-therapist gender match on psychotherapy retention among United States veterans with posttraumatic stress disorder. *Administration and Policy in Mental Health and Mental Health Services Research, 44,* 642-650.

Soldz, S. (2006) Models and meanings: Therapist effects and the stories we tell. *Psychotherapy Research, 16*(02), 173-177. doi: 10.1080/10503300500264937.

Spivack, J., Beale, C., Burks, S., Burns, W., Bolin, M., Collins Eaglin, J., . . .Weitzman, L. (2010). Accreditation Standards for University and College Counseling Services. Retrieved from http://www.iacsinc.org/IACS%20STANDARDS%20rev%2010-3-11.pdf

Stanhope, V., & Straussner, A. L. A. (2017). *Social work and integrated health care: From policy to practice and back*. New York, NY: Oxford University Press.

Stone, G. L. & McMichael, J. (1996). Thinking about mental health policy in university and college counseling centers. *Journal of College Student Psychotherapy, 10*(3), 3-27. doi: 10.1300/J035v10n03_02.

Uffleman, R. A. & Hardin, S. I. (2002). Session limits at university counseling centers: Effects on help-seeking attitudes. *Journal of Counseling Psychology, 49*(1), 127-132.

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,* 37-54. American Psychological Association.

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

Watkins, D. C., Hunt, J. B., & Eisenberg, D. (2012). Increased demand for mental health services on college campuses: Perspectives from administrators. *Qualitative Social Work, 11*(3), 319-337. doi: 10.1177/1473325011401468.

Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2).

Williams, N. J. (2016). Multilevel mechanisms of implementation strategies in mental health: Integrating theory, research, and practice. *Administration and Policy in Mental Health and Mental Health Services Research, 43*, 783-798. doi: 10.1007/s10488-015-0693-2.

Wolgast, B. M., Lambert, M. J., & Puschner, B. (2004). The dose-response relationship at a college counseling center. *Journal of College Student Psychotherapy, 18*(2), 15-29. doi: 10.1300/J035v18n0203.

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).

Youn, S., Castonguay, L.G., Janis, R. A., Hayes, J.A., & Locke, B.D. (submitted for publication). *Therapist effects: Unpacking the meaning behind differences.*

Appendix A

Table 1.

Summary of client-level variables and demographics

Variable (unit)	Mean	SD	Min	Max
Age (years)	22.7	5.11	18	60
Tx length (# sessions)	9.14	7.59	2	196
Initial severity (CCAPS score)	1.83	0.60	0	4
	n	%		
Gender, SDS_88				
Female	11017	37.95%		
Male	5296	18.24%		
Self-identify	81	0.28%		
Other	202	0.70%		
NA	12432	42.83%		
Ethnicity, SDS_95				
African American/Black	1316	4.53%		
American Indian or Alaskan Native	62	0.21%		
Asian/Asian American	1264	4.35%		
Hispanic/Latino(a)	1430	4.93%		
Native Hawaiian or Pacific Islander	24	0.08%		
Multi-racial	787	2.71%		
White	11658	40.16%		
Self-identify	276	0.95%		
NA	12211	42.07%		

Note. n= 29,028 clients. "Self-identify" was a response where clients had the option to write-in how they self-identified. NA = missing data.

Table 2.

Summary of center-level variables

Center predictor (unit)	Mean	SD	Min	Max
Average initial severity (CCAPS)	1.47	0.62	0.00	3.73
Average tx length (# sessions)	6.89	1.88	3.51	15.3
Session frequency (days b/w sessions)	15.56	3.24	8.14	24.32
Enrollment ratio (student/staff)	518.84	292.48	33.12	1761.86

	n	%
APA Accredited		
Yes	39	37.14
No	65	61.9
NA	1	1.00
IACS Accredited		
Yes	58	55.24
No	46	43.81
NA	1	1.00
Session limits		
Yes	39	37.14
No	65	61.9
NA	1	1.00
NCAA Athletic Division		
I	66	62.86
II	16	15.24
III	15	14.29
None	6	5.71
NA	2	1.90
School type		
Public	72	68.57
Private	30	28.57
Combined	2	1.90
NA	1	1.00

Note. n centers= 105. NA = data missing.

Table 3.

Results of MLM predicting change in Depression scores

	Parameter estimates (SE)					
Model parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	**-0.6*** **	**-0.59*** **	**-0.64*** **	**-0.70*** **	**-0.86*** **	**-0.71*** **
	(.01)	(.01)	(.07)	(.06)	(.08)	(.05)
Client tx length		**0.00*** **	**0.00*** **	**0.00*** **	**0.00*** **	**0.00*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Client init. severity		**-0.38*** **	**-0.38*** **	**-0.38*** **	**-0.38*** **	**-0.38*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Center avg tx length			0.01	---	---	---
			(.01)	---	---	---
Center avg init. severity			**-0.38*** **	**-0.33*** **	**-0.37*** **	**-0.35*** **
			(.07)	(.08)	(.08)	(.07)
Session limits			0.00	---	---	---
			(.02)	---	---	---
Center avg session freq			**0.00***	**0.01***	**0.01** **	**0.01** **
			(.00)	(.00)	(.00)	(.00)
Student-staff ratio				0.00	---	---
				(.00)	---	---
Services						
Health					-.03	---
					(.03)	---
Learning					0.05	---
					(.05)	---
Drug & Alcohol					-0.01	---
					(.02)	---
Psychiatric					0.01	---
					(.01)	---
IACS Accreditation					0.01	---
					(.02)	---
APA Accreditation					-0.01	---
					(.02)	---
Center size					0.00	---
					(.00)	---
Athletic Division					0.02	---
					(.01)	---
School type					0.04	---

					(.03)	---
Random effects						
Residual	0.67	0.59	0.59	0.59	0.59	0.59
Level 1- Intercept	0.08	0.07	0.07	0.07	0.07	0.07
Level 2- Intercept	0.09	0.1	0.08	0.08	0.07	0.08
Goodness of fit						
-2LL	-27817	-24455	-24442	-24443	-24437	-24443
AIC	55642	48922	48904	48904	48909	48902

Note. Model 1 is a null model. Model 2 added client predictors. Model 3 added the first wave of center predictors. Model 4 added the second wave with student staff ratio. Model 5 added the third wave of center predictors. Model 6 retained only the significant fixed effects. --- indicates the predictor was not retained due to non-significance. **Bold** indicates the predictor was kept. *p<.05, **p<.01, ***p<.001. N centers = 102, n therapists = 1,601, n clients = 27,123.

Table 4.

Results of MLM predicting change in Anxiety scores

	Parameter estimates (SE)					
Model parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	**-0.53***** (.01)	**-0.52***** (.01)	**-0.50***** (.07)	**-0.42***** (.05)	**-0.52***** (.08)	**-0.44***** (.04)
Client tx length		**0.00***** (.00)	**0.00***** (.00)	**0.01***** (.00)	**0.00***** (.00)	**0.01***** (.00)
Client init. severity		**-.28***** (.00)	**-.28***** (.00)	**-.28***** (.00)	**-.28***** (.00)	**-0.28***** (.00)
Center avg tx length			**-0.01*** (.01)	**-.01*** (.01)	**-.01*** (.01)	**-.01*** (.01)
Center avg init. severity			**-0.29***** (.09)	**-0.29**** (.09)	**-0.32***** (.08)	**-0.29**** (.09)
Session limits			0.00 (.02)	--- ---	--- ---	--- ---
Center avg session freq			0.00 (.00)	--- ---	--- ---	--- ---
Student-staff ratio				0.00 (.00)	--- ---	--- ---
Services						
Health					-0.03 (.03)	--- ---
Learning					0.03 (.05)	--- ---
Drug & Alcohol					0.00 (.02)	--- ---
Psychiatric					0.01 (.01)	--- ---
IACS Accreditation					0.03 (.02)	--- ---
APA Accreditation					-0.04 (.02)	--- ---
Center size					0.00 (.00)	--- ---
Athletic Division					0.01	---

					(.01)	---
School type					0.02	---
					(.02)	---
Random effects						
Residual	0.65	0.61	0.61	0.61	0.60	0.61
Level 1- Intercept	0.07	0.07	0.07	0.07	0.07	0.07
Level 2- Intercept	0.09	0.09	0.08	0.08	0.07	0.08
Goodness of fit						
-2LL	-27108	-25068	-25061	-25061	-25055	-25062
AIC	54224	50147	50142	50142	50144	50140

Note. Model 1 is a null model. Model 2 adds client predictors. Model 3 adds the first wave of center predictors. Model 4 adds the second wave with student staff ratio. Model 5 adds the third wave of center predictors. Model 6 retained only the significant fixed effects. --- indicates the predictor was not retained due to non-significance. **Bold** indicates the predictor was kept. *p<.05, **p<.01, ***p<.001. N centers = 102, n therapists = 1,601, n clients = 27,049.

Table 5.

Results of MLM predicting change in Social Anxiety scores

	Parameter estimates (SE)					
Model parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	**-0.35*** **	**-0.33*** **	**-0.34*** **	**-0.30*** **	**-0.34*** **	**-0.33*** **
	(.01)	(.01)	(.05)	(.04)	(.04)	(.01)
Client tx length		**0.00*** **	**0.00*** **	**0.00*** **	**0.00*** **	**0.00*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Client init. severity		**-0.16*** **	**-0.16*** **	**-0.16*** **	**-0.16*** **	**-0.16*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Center avg tx length			**0.00*** **	0.00	---	---
			(.00)	(.00)	---	---
Center avg init. severity			-0.10	---	---	---
			(.06)	---	---	---
Session limits			0.01	---	---	---
			(.01)	---	---	---
Center avg session freq			0.00	---	---	---
			(.00)	---	---	---
Student-staff ratio				0.00	---	---
				(.00)	---	---
Services						
Health					-.01	---
					(.02)	---
Learning					0.03	---
					(.03)	---
Drug & Alcohol					-0.02	---
					(.01)	---
Psychiatric					0.00	---
					(.00)	---
IACS Accreditation					0.03	---
					(.02)	---
APA Accreditation					-0.02	---
					(.02)	---
Center size					0.00	---
					(.00)	---
Athletic Division					0.01	---

					(.01)	**---**
School type					-0.01	**---**
					(.02)	**---**
Random effects						
Residual	0.46	0.44	0.44	0.44	0.44	0.44
Level 1- Intercept	0.04	0.04	0.04	0.04	0.04	0.04
Level 2- Intercept	0.05	0.05	0.05	0.05	0.05	0.05
Goodness of fit						
-2LL	-17915	-16602	-16600	-16601	-16598	-16602
AIC	35838	33216	33220	33219	33226	33216

Note. Model 1 is a null model. Model 2 adds client predictors. Model 3 adds the first wave of center predictors. Model 4 adds the second wave with student staff ratio. Model 5 adds the third wave of center predictors. Model 6 retained only the significant fixed effects. --- indicates the predictor was not retained due to non-significance. **Bold** indicates the predictor was kept. *p<.05, **p<.01, ***p<.001. N centers = 102, n therapists = 1,601, n clients = 27,277.

Table 6.

Results of MLM predicting change in Eating Concerns scores

	Parameter estimates (SE)					
Model parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	**-0.20*** **	**-0.20*** **	**-0.21*** **	**-0.20*** **	**-0.20*** **	**-0.21*** **
	(.01)	(.01)	(.05)	(.02)	(.04)	(.01)
Client tx length		**0.00*** **	**0.00** **	**0.00** **	**0.00** **	**0.00** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Client init.severity		**-0.26*** **	**-0.26*** **	**-0.26*** **	**-0.26*** **	**-0.26*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Center avg tx length			0.00	---	---	---
			(.00)	---	---	---
Center avg init. severity			**-0.40*** **	**-0.40*** **	**-0.37*** **	**-0.38*** **
			(.07)	(.07)	(.07)	(.07)
Session limits			0.18	---	---	---
			(.02)	---	---	---
Center avg session freq			0.00	---	---	---
			(.00)	---	---	---
Student-staff ratio				0.00	---	---
				(.00)	---	---
Services						
Health					-0.01	---
					(.02)	---
Learning					0.00	---
					(.04)	---
Drug & Alcohol					0.00	---
					(.02)	---
Psychiatric					0.00	---
					(.00)	---
IACS Accreditation					**0.04***	0.03
					(.04)	(.01)
APA Accreditation					-0.01	---
					(.02)	---
Center size					0.00	---
					(.00)	---
Athletic Division					0.00	---

					(.01)	---
School type					0.00	---
					(.02)	---
Random effects						
Residual	0.61	0.55	0.55	0.55	0.55	0.55
Level 1- Intercept	0.00	0.00	0.00	0.00	0.00	0.00
Level 2- Intercept	0.06	0.07	0.05	0.06	0.06	0.06
Goodness of fit						
-2LL	-25388	-22758	-22741	-22744	-22741	-22742
AIC	50783	45528	45503	45503	45514	45500

Note. Model 1 is a null model. Model 2 adds client predictors. Model 3 adds the first wave of center predictors. Model 4 adds the second wave with student staff ratio. Model 5 adds the third wave of center predictors. Model 6 retained only the significant fixed effects. --- indicates the predictor was not retained due to non-significance. **Bold** indicates the predictor was kept. *p<.05, **p<.01, ***p<.001. N centers = 102, n therapists = 1,601, n clients = 27,121.

Table 7.

Results of MLM predicting change in Hostility scores

	Parameter estimates (SE)					
Model parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	**-0.43*** **	**-0.40*** **	**-0.49*** **	**-0.39*** **	**-0.47*** **	**-0.41*** **
	(.01)	(.01)	(.08)	(.02)	(.06)	(.01)
Client tx length		**-0.01*** **	**-0.01*** **	**-0.01*** **	**-0.01*** **	**-0.01*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Client init. severity		**-0.41*** **	**-0.41*** **	**-0.41*** **	**-0.41*** **	**-0.41*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Center avg tx length			0.00	---	---	---
			(.01)	---	---	---
Center avg init. severity			**-0.32*** **	**-0.31*** **	**-0.33*** **	**-0.32*** **
			(.07)	(.07)	(.07)	(.07)
Session limits			0.00	---	---	---
			(.02)	---	---	---
Center avg session freq			0.00	---	---	---
			(.01)	---	---	---
Student-staff ratio				0.00	---	---
				(.00)	---	---
Services						
Health					-0.02	---
					(.03)	---
Learning					0.02	---
					(.05)	---
Drug & Alcohol					-0.01	---
					(.02)	---
Psychiatric					0.01	---
					(.01)	---
IACS Accreditation					0.03	---
					(.02)	---
APA Accreditation					-0.03	---
					(.03)	---
Center size					0.00	---
					(.00)	---
Athletic Division					0.02	---

					(.01)	---
School type					0.00	---
					(.03)	---
Random effects						
Residual	0.70	0.56	0.56	0.56	0.56	0.56
Level 1- Intercept	0.07	0.07	0.07	0.07	0.07	0.07
Level 2- Intercept	0.10	0.11	0.09	0.09	0.08	0.09
Goodness of fit						
-2LL	-29162	-23183	-23171	-23172	-23169	-23172
AIC	58333	46378	46362	46360	46369	46358

Note. Model 1 is a null model. Model 2 adds client predictors. Model 3 adds the first wave of center predictors. Model 4 adds the second wave with student staff ratio. Model 5 adds the third wave of center predictors. Model 6 retained only the significant fixed effects. --- indicates the predictor was not retained due to non-significance. **Bold** indicates the predictor was kept. *p<.05, **p<.01, ***p<.001. N centers = 102, n therapists = 1,601, n clients = 27,121.

Table 8.

Results of MLM predicting change in Academic Concerns scores

	Parameter estimates (SE)					
Model parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	**-0.25*** **	**-0.24*** **	**-0.29*** **	**-0.31*** **	**-0.39*** **	**-0.34*** **
	(.01)	(.01)	(.05)	(.04)	(.05)	(.04)
Client tx length		**0.00*** **	**0.00*** **	**0.00*** **	**0.00*** **	**0.00*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Client initial severity		**-0.30*** **	**-0.30*** **	**-0.30*** **	**-0.30*** **	**-0.30*** **
		(.00)	(.00)	(.00)	(.00)	(.00)
Center avg tx length			0.00	---	---	---
			(.00)	---	---	---
Center avg initial severity			**-0.42*** **	**-0.39*** **	**-0.34*** **	**-0.42*** **
			(.08)	(.08)	(.09)	(.07)
Session limits			-0.03	---	---	---
			(.01)	---	---	---
Center avg session freq			**0.01***	**0.00***	**0.01*** **	**0.01*** **
			(.00)	(.00)	(.00)	(.00)
Student-staff ratio				0.00	---	---
				(.00)	---	---
Services						
Health					-0.01	---
					(.02)	---
Learning					0.03	---
					(.03)	---
Drug & Alcohol					0.00	---
					(.01)	---
Psychiatric					0.00	---
					(.00)	---
IACS Accreditation					0.01	---
					(.01)	---
APA Accreditation					-0.02	---
					(.02)	---
Center size					0.00	---
					(.00)	---
Athletic Division					0.01	---

					(.02)	---
School type					0.01	---
					(.01)	---
Random effects						
Residual	0.46	0.23	0.43	0.43	0.43	0.43
Level 1- Intercept	0.05	0.04	0.04	0.04	0.04	0.04
Level 2- Intercept	0.06	0.07	0.05	0.05	0.05	0.05
Goodness of fit						
-2LL	-17842	-15728	-15709	-15710	-15704	-15711
AIC	35692	31467	31438	31438	31441	31438

Note. Model 1 is a null model. Model 2 adds client predictors. Model 3 adds the first wave of center predictors. Model 4 adds the second wave with student staff ratio. Model 5 adds the third wave of center predictors. Model 6 retained only the significant fixed effects. --- indicates the predictor was not retained due to non-significance. **Bold** indicates the predictor was kept. *p<.05, **p<.01, ***p<.001. N centers = 102, n therapists = 1,601, n clients = 27,459.

Table 9.

Results of MLM predicting change in Alcohol scores

	Parameter estimates (SE)					
Model parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Fixed effects						
Intercept	**-0.18*****	**-0.17*****	**-0.21*****	**-0.18*****	**-0.15*****	**-0.18*****
	(.01)	(.01)	(.05)	(.01)	(.04)	(.01)
Client tx length		**0.00*****	**0.00*****	**0.00*****	**0.00*****	**0.00*****
		(.00)	(.00)	(.00)	(.00)	(.00)
Client init. severity		**-0.40*****	**-0.40*****	**-0.40*****	**-0.40*****	**-0.40*****
		(.00)	(.00)	(.00)	(.00)	(.00)
Center avg tx length			0.00	---	---	---
			(.00)	---	---	---
Center avg init. severity			**-0.29*****	**-0.30*****	**-0.31*****	**-0.30*****
			(.00)	(.05)	(.06)	(.05)
Session limits			0.01	---	---	---
			(.01)	---	---	---
Center avg session freq			0.00	---	---	---
			(.00)	---	---	---
Student-staff ratio				0.00	---	---
				(.00)	---	---
Services						
Health					-0.02	---
					(.02)	---
Learning					-0.01	---
					(.03)	---
Drug & Alcohol					0.00	---
					(.01)	---
Psychiatric					0.00	---
					(.01)	---
IACS Accreditation					0.02	---
					(.01)	---
APA Accreditation					-0.02	---
					(.02)	---
Center size					0.00	---
					(.00)	---
Athletic Division					0.00	---

					(.01)	---
School type					-0.02	---
					(.02)	---
Random effects						
Residual	0.58	0.46	0.46	0.46	0.46	0.46
Level 1- Intercept	0.05	0.00	0.00	0.00	0.00	0.00
Level 2- Intercept	0.05	0.06	0.05	0.05	0.05	0.05
Goodness of fit						
-2LL	-23849	-17704	-17688	-17689	-17686	-17689
AIC	47706	35419	35397	35393	35403	35391

Note. Model 1 is a null model. Model 2 adds client predictors. Model 3 adds the first wave of center predictors. Model 4 adds the second wave with student staff ratio. Model 5 adds the third wave of center predictors. Model 6 retained only the significant fixed effects. --- indicates the predictor was not retained due to non-significance. **Bold** indicates the predictor was kept. *p<.05, **p<.01, ***p<.001. N centers = 102, n therapists = 1,601, n clients = 27,359.

Table 10.

Summary of predictors from final multilevel model for each subscale

	Subscale						
Model parameters	Depr.	Anxiety	Soc.Anx.	Eating	Hostility	Academ.	Alcohol
Client tx length	***	***	***	**	***	***	***
Client init. severity	***	***	***	***	***	***	***
Center avg tx length	---	*	---	---	---	---	---
Center avg init. severity	***	**	---	***	***	***	***
Center session limits	---	---	---	---	---	---	---
Center avg session freq	*	---	---	---	---	**	---
Student-staff ratio	---	---	---	---	---	---	---
Services	---	---	---	---	---	---	---
Health	---	---	---	---	---	---	---
Learning	---	---	---	---	---	---	---
Drug & Alcohol	---	---	---	---	---	---	---
Psychiatric	---	---	---	---	---	---	---
IACS Accreditation	---	---	---	---	---	---	---
APA Accreditation	---	---	---	---	---	---	---
Center size	---	---	---	---	---	---	---
Athletic Division	---	---	---	---	---	---	---
School type	---	---	---	---	---	---	---

*indicates p-value was <.05, **indicates p-value was <.01, ***indicates p-value was <.001.

Appendix B

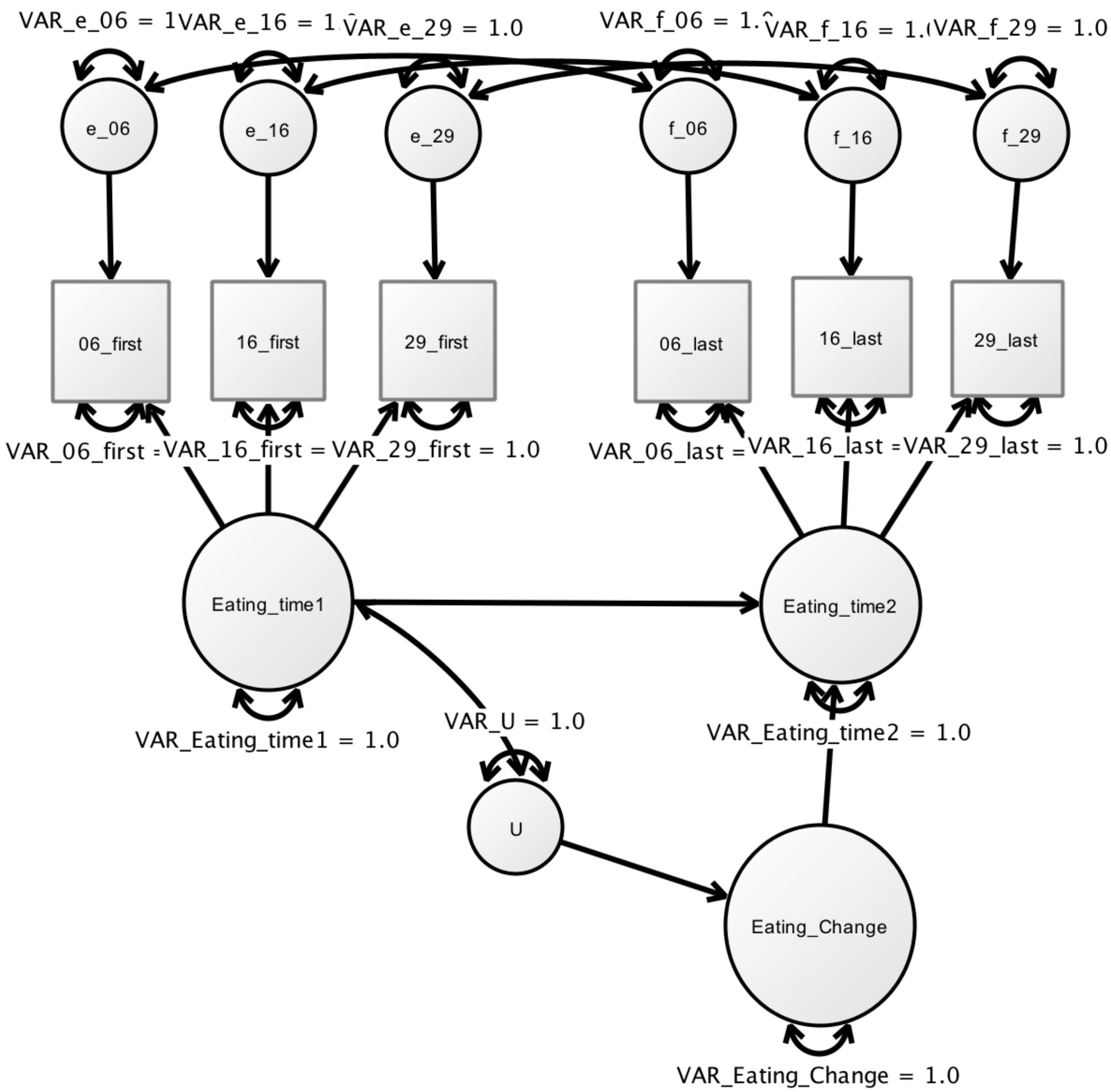

Figure 1. Latent change score model for change on Eating Concerns. U = within-person variance. 06_first = observed score on item 6 at the first instance of the CCAPS administration. 06_last = observed score on item 6 at the last instance of the CCAPS administration. e_06= error term. This pattern continues for the other items in the scale.

Appendix C

CCAPS-34 Items and Corresponding Subscales

<table>
<tr><th>Scale</th><th>Item</th></tr>
<tr><td>Depression</td><td>I don’t enjoy being around people as much as I used to
I feel isolated and alone
I feel worthless
I feel helpless
I feel sad all the time
I have thoughts of ending my life</td></tr>
<tr><td>Generalized Anxiety</td><td>My heart races for no good reason
I am anxious that I might have a panic attack in public
I have sleep difficulties
My thoughts are racing
I have spells of terror or panic
I feel tense</td></tr>
<tr><td>Social Anxiety</td><td>I am shy around others
I make friends easily
I am concerned that other people do not like me
I feel uncomfortable around people I don't know
I feel self-conscious around others</td></tr>
<tr><td>Academic Distress</td><td>I feel confident I can succeed academically
I am not able to concentrate as well as usual
It's hard to stay motivated for my classes
I am unable to keep up with my school work</td></tr>
<tr><td>Eating Concerns</td><td>I feel out of control when I eat
I think about food more than I would like to
I eat too much</td></tr>
<tr><td>Hostility</td><td>I have difficulty controlling my temper
I sometimes feel like breaking or smashing things
I get angry easily
I am afraid I may lose control and act violently
I frequently get into arguments
I have thoughts of hurting others</td></tr>
<tr><td>Alcohol Use</td><td>I drink alcohol frequently
When I drink alcohol I can't remember what happened
I drink more than I should
I have done something I have regretted because of drinking</td></tr>
</table>

Appendix D

Standardized Data Set (SDS) Items and Response Options

Question	**Response Options**
Center ID	Auto-generated by EMR software
Does your counseling center have a currently accredited APA pre-doctoral training program (American Psychological Association)?	Yes No
Is your counseling center currently accredited by IACS (International Association of Counseling Services?)	Yes No
Which services are integrated with your counseling center? (check all that apply)	Career Services Disability Services Drug & Alcohol Treatment Program Employee Assistance Program Learning Services Health Services Testing Services (e.g., standardized testing) Other (please specify)
What psychiatric services are provided by your center? (do not include psychiatric services through health services unless you are integrated)	None Part time, in house Full time, in house Part time, off campus consultant Other (please specify)
Does your center have an annual individual psychotherapy session limit?	Yes No
Is your institution private, public, or combined?	Private Public Combined
Please indicate which athletic division your institution currently belongs to:	None Division I Division II Division III