

Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications

Psychology, Department of

2014

The Utility of an Efficient Outcomes Assessment System at University Counseling Centers

S. Mark Kopta
University of Evansville

Megan Lynn Petrik
Marquette University, megan.petrik@marquette.edu

Stephen M. Saunders
Marquette University, stephen.saunders@marquette.edu

Michael Mond
Johns Hopkins University

Glenn Hirsch
University of Minnesota

See next page for additional authors

Follow this and additional works at: https://epublications.marquette.edu/psych_fac



Part of the [Psychology Commons](#)

Recommended Citation

Kopta, S. Mark; Petrik, Megan Lynn; Saunders, Stephen M.; Mond, Michael; Hirsch, Glenn; Kadison, Richard; and Raymond, Danielle, "The Utility of an Efficient Outcomes Assessment System at University Counseling Centers" (2014). *Psychology Faculty Research and Publications*. 153.
https://epublications.marquette.edu/psych_fac/153

Authors

S. Mark Kopta, Megan Lynn Petrik, Stephen M. Saunders, Michael Mond, Glenn Hirsch, Richard Kadison, and Danielle Raymond

Marquette University

e-Publications@Marquette

Psychology Faculty Research and Publications/College of Arts and Sciences

This paper is NOT THE PUBLISHED VERSION.

Access the published version via the link in the citation below.

Journal of College Student Psychotherapy, Vol. 28, No. 2 (2014): 97-116. [DOI](#). This article is © Taylor & Francis (Routledge) and permission has been granted for this version to appear in [e-Publications@Marquette](#). Taylor & Francis (Routledge) does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Taylor & Francis (Routledge).

The Utility of an Efficient Outcomes Assessment System at University Counseling Centers

S. Mark Kopta

University Of Evansville, Evansville, Indiana

Megan L. Petrik

Marquette University, Milwaukee, Wisconsin

Stephen M. Saunders

Marquette University, Milwaukee, Wisconsin

Michael Mond

Johns Hopkins University, Baltimore, Maryland

Glenn Hirsch

University Of Minnesota, Minneapolis, Minnesota

Richard Kadison

Harvard University, Cambridge, Massachusetts

Danielle Raymond

University Of Evansville, Evansville, Indiana

Abstract

Due to increased demands placed on university counseling centers (UCCs) in recent years, there is a need for these centers to enhance the efficiency and effectiveness of their psychological services. Regularly monitoring client progress is one approach to increase the likelihood of positive clinical outcomes. This article describes the use of the Behavioral Health Measure–20 (BHM-20; Kopta & Lowry, 2002) in monitoring the progress of 13,803 clients at 23 UCCs across the United States from 2006–2011. Results show that the BHM-20, via the CelestHealth System–MH (CHS-MH), is an effective instrument for the electronic administration, scoring, and tracking of client progress. Overall, clients improved over the course of treatment and the benefit from treatment peaked between 7 to 10 sessions. Three case vignettes using the CHS-MH in patient care are presented. Implications for continuing the use of monitoring individual client change at UCCs are discussed.

KEYWORDS

clinical significance, college students, psychotherapy outcome, university counseling centers

Beginning in the 1980s, university counseling centers (UCCs) evolved from primarily servicing developmental issues (e.g., academic challenges, the home-to-college transition) to treating clinical syndromes (e.g., mood, anxiety, substance use disorders) and managing crises (Benton, Robertson, Tseng, Newton, & Benton, [1]; Kitzrow, [18]; Watkins, Hunt, & Eisenberg, [39]; Wolgast et al., [42]). For example, Erdur-Baker, Aberson, Barrow, and Draper ([10]) compiled data from 32 university and college counseling centers in the United States. Comparing the intake status of student clients in 1991 and 1997, they found that the 1997 clinical sample had more severe problems of greater chronicity.

The literature does not provide a clear reason for the increase in the severity of presenting problems to UCCs. Some have suggested that societal values are shifting away from instilling intrinsic goals in youth (e.g., finding a purpose in life) toward instilling extrinsic goals (e.g., earning a high income), leading to generational increases in psychopathology (Twenge et al., [38]). Other reasons might include reduced stigmatization of mental illness, making it easier for college students to seek psychological treatment, and increased capacity for students with serious and persistent mental illness to attend college due to improvements in psychotropic medications and illness management (Much & Swanson, [30]).

UCC staff are noticing the increased demands for college mental health services. In the 2011 National Survey of Counseling Center Directors (Gallagher, [12]), 228 counseling center directors provided information about the nature of college mental health services for 2.3 million students at their institutions. Nine out of 10 directors reported an increase in students seeking services for severe psychological problems, such as crisis issues requiring an immediate response, alcohol and illicit drug use, and eating disorders. Gallagher ([12]) also found that 23% of counseling center clients were on psychiatric medication, which is an increase from 20% in 2003, 17% in 2000, and 9% in 1994. Unfortunately, there has not been a proportional increase in resources and funding for UCCs (e.g., Guinee & Ness, [13]; Kraft, [21]), which has resulted in more stress for UCC staff (Gallagher, [12]) and debate about the scope of mental health services that UCCs should provide (Eells, [7]).

To meet the challenge of increased demands, UCCs strive to use existing resources more efficiently and effectively. They seek to gather information about clients to determine what type of treatment the person might require, how long treatment might last, and whether treatment seems to be working (Hansen, Lambert, & Forman, [14]). To do this, UCCs are beginning to employ methodologies developed by psychotherapy researchers, specifically instruments that can be regularly administered over the course of treatment in order to monitor treatment outcomes. Outcomes have traditionally been evaluated at the group level (e.g., to determine which treatment for depression is more effective). More recently, researchers have developed methods to monitor progress at the individual level. This information is much more relevant to the treating clinicians (e.g.,

"Is this particular client improving?"). Similar to managed care organizations, UCCs are finding that monitoring both group and individual outcomes proves useful in documenting the need for treatment, maximizing efficiency, and accounting for use of resources.

COMPARING GROUP OUTCOMES

To evaluate the statistical significance of change in treatment, psychotherapy researchers historically conducted null hypothesis significance testing (NHST), wherein the effects of treatment are evaluated by comparing changes in groups of patients from preintervention to postintervention. Numerous important NHST studies have established the general efficacy of psychotherapeutic treatments, such as the National Institute of Mental Health Treatment of Depression Collaborative Research Program (NIMH TDCRP) (Elkin et al., [9]), Project MATCH (Project Match Research Group, [34]), and others (e.g., Westen & Morrison, [40]).

Thanks to such research, mental health professionals can be confident that the services they provide are effective. However, NHST can only identify whether the obtained magnitude of change obtained by a group exceeds what would be expected by chance. NHST does not indicate information of most interest and importance to clients and clinicians, which is whether and to what extent an individual client is improving (Ogles, Lunnen, & Bonesteel, [33]).

TRACKING INDIVIDUAL OUTCOMES

By tracking changes in clients' mental health across sessions, treatment can be modified in order to improve outcomes. Research suggests that providing feedback to clinicians regarding client progress positively affects treatment outcomes, reduces the likelihood of clinical deterioration, and enhances the likelihood of positive outcomes even for clients who showed an initial poor response to treatment (Lambert, [22]; Lambert et al., [25]; Maruish, [29]; Reese, Norsworthy, & Rowlands, [35]).

Evaluating Clinical Significance

Jacobson and Truax ([17]) introduced a method to enable researchers and clinicians to determine whether any particular client achieved "clinically significant change." To do so, a client's change score on a measure is rated on two criteria. First, the change is rated in terms of whether it is statistically reliable (i.e., exceeding that expected by chance[1]). Second, the end-of-treatment score is rated in terms of whether it places the client within the normal range of functioning, meaning that the score is more within the range of scores obtained by a functional population than the range of scores of a dysfunctional population.

Jacobson and Truax ([17]) proposed that these two criteria be used to classify individuals into one of four categories. A client has "deteriorated" if the change score reliably moved in a negative direction over treatment. A client is considered "unchanged" if the score did not reliably change in either direction. A client is "improved" if the score reliably changed in the positive direction. Finally, a client is considered "recovered" if both the score changes reliably and the end-of-treatment score is in the range of the functional population.

Measures to Track Individual Progress

The past several decades have seen a profusion of methods and measures that UCCs can use to track and evaluate whether a particular client is improving (or deteriorating). Historically, measures were fairly narrow, such as measuring depressive symptoms in a study comparing depression treatments. Most modern instruments are fairly broad-based, and they measure a wide range of symptoms, distress, and impairment in order to capture the problems presented by a broad range of clientele. Three such instruments are the COMPASS system, the Outcome Questionnaire (OQ), and the Counseling Center Assessment of Psychological Symptoms (CCAPS).

A 123-item paper-and-pencil questionnaire, the COMPASS system (Sperry, Brill, Howard, & Grissom, [37]) evaluated change in three areas (i.e., subjective well-being, symptoms, and life functioning) over the course of treatment, based on the phase model of psychotherapy outcomes (Howard, Lueger, Maling, & Martinovich, [16]). According to the phase model, successful treatment entails three sequential stages of improvement. In the initial *rehabilitation* phase, the client obtains a greater sense of well-being because of increased optimism and hope. During the subsequent *remediation* phase, symptoms that are associated with specific diagnoses diminish (e.g., panic attacks, pessimistic thinking). In the final *rehabilitation* phase, functioning in work, relationships, and self-care improves. Studies have supported the idea that the phases occur sequentially and are contingent (e.g., Lutz, Lowry, Kopta, Einstein, & Howard, [28]). A clinical outcomes report was generated, so that results might be discussed by the therapist and client at the subsequent session. The length of the COMPASS would likely be too cumbersome in today's UCC setting, but the system provided a template for how to use regular outcomes assessment to evaluate a client's progress in treatment.

The OQ (Lambert, Harmon, Slade, Whipple, & Hawkins, [24]) also assesses multiple aspects of clients' mental health (i.e., psychological symptoms, interpersonal problems, social role functioning, and quality of life). It is available in a 30-item version and a 45-item version, and both have been shown to be reliable and valid. The OQ is completed via computer interface, and feedback about treatment progress is therefore readily available to clinicians and clinic directors, which is potentially helpful in alerting to deterioration in mental health and to crisis situations, such as suicidal or homicidal risk.

The CCAPS was created specifically for UCCs (Locke et al., [27]). There are two versions, one comprising 62 items and the other 34 items, and both have been shown to have good reliability and validity. The normative sample is quite large and is comprised of only college students. The CCAPS evaluates numerous areas of symptoms and distress that might be experienced by college students, including depression, anxiety, substance use problems, and academic distress.

THE PRESENT STUDY

This article describes the use of another instrument, the Behavioral Health Measure–20 (BHM-20; Kopta & Lowry, [19]), at 23 UCCs across the United States. The BHM-20 is included in the computer-based CelestHealth System-MH (CHS-MH; Bryan, Kopta, & Lowes, [3]) and was developed specifically for UCCs. It is broad-based and features the phase model of psychotherapy outcome. It is administered electronically, which allows immediate feedback to clinicians. The BHM-20 includes a Suicide Monitoring Scale (SMS), which enables immediate feedback about the status of clients that admit to suicidal ideation or risk. Finally, the BHM-20 categorizes clients into clinical significance categories, which is useful at the individual and the clinic-wide level. Using clinical significance methodology, the CHS-MH provides each UCC summative information regarding outcomes for its clients (e.g., percent improved, percent recovered).

In addition to evaluating the clients' response to treatment while using the CHS-MH, our study will present three case studies that highlight the use of the CHS-MH in individual client care.

METHOD

Participants

Data were available from 22,465 total counseling center clients seen at 23 different UCCs associated with a variety of public and private colleges and universities in the United States between June 2006 and November 2011. The UCCs were those who subscribed to an electronic outcomes assessment system offered by CelestHealth Solutions, LLC. Our sample consisted of the 13,803 clients who participated in a minimum of 2 and a maximum of 25 psychotherapy sessions ($M = 5.66$, $SD = 4.82$; median = 4.00). There were no diagnostic criteria

for inclusion in the study. There were neither criteria for treatment termination nor limits to treatment duration, although some therapists set limits as part of treatment strategy.

All clients were over 18 years of age and were college students. Most were female (62.5%) and about one third were male (35.3%). Some did not report their gender (2.2%). The majority of clients indicated that they are Caucasian (56.6%), followed by Asian American/Pacific Islander (5.9%), African American/Black (5.3%), Latina/Latino/Hispanic (5.1%), "Other" (4.0%), and biracial/multiracial (1.4%). One in five clients (20.0%) did not report their ethnicity. With regard to relationship status, the modal response was "Single, not dating" (44.7%), followed by "Dating" (17.6%), "Married" (12.0%), and "Separated" or "Divorced" (2.5%). Relationship status was not reported by 22.3% of clients.

In accordance with Jacobson and Truax ([17]), clinical significance analyses were conducted with the functional-at-intake subset removed. In other words, only those who scored below the cutoff separating functional from dysfunctional were included; this subset was determined specifically for each scale using Jacobson and Truax's formula and the norms for the BHM-20 (Kopta & Lowry, [19]; see Procedure section for more details). The size of the dysfunctional-at-intake groups were as follows: 9,503 patients on the Global Mental Health scale, 9,423 on the Well-Being scale, 7,004 on the Symptoms scale, and 9,821 on the Life Functioning scale.

Across the 23 UCCs, 425 different therapists provided psychotherapy. Clients were included in this study if their therapist either had a doctorate in psychology or social work (36.7%) or a master's degree in counseling psychology, clinical psychology, or social work (63.3%). Therapists also included trainees at various levels (e.g., predoctoral interns, social work interns, postdoctoral clinicians). Data from psychiatric visits and visits with bachelor-level counselors were excluded from this study. The therapists varied in theoretical orientation and treatment approach.

Measures

Behavioral Health Measure—20

Treatment progress was assessed using the BHM-20 (Kopta & Lowry, [19]). The BHM-20 is administered using either a tablet or desktop computer, and it takes most clients about 90 seconds to complete. Clients are instructed to answer questions "as they relate to the past two weeks." The items are averaged to create an overall scale called Global Mental Health (GMH), with higher scores indicating better mental health. The BHM-20 has three primary scales—well-being, psychological symptoms, and life functioning—that are designed to evaluate the three phases of outcome as delineated in the phase model (Howard et al., [16]). As with the GMH, the primary scales are created by averaging the items, with higher scores indicating better mental health.

The well-being (WB) scale comprises three items that measures distress, life satisfaction, and level of energy and motivation. Clients responded to the items on a 5-point (0–4) Likert scale, with higher scores indicating greater well-being.

The psychological symptoms (SYM) scale includes 13 items measuring some of the most common symptoms that clients present with to psychotherapy, such as symptoms of depression, anxiety, panic, mood swings, eating problems, and alcohol- or drug-use problems. There are two items assessing suicidality and one evaluating risk of violence. Clients responded to the items on a 5-point Likert scale ranging from 0 (*almost always*) to 4 (*never*). The items on the SYM are used to generate subscales, comprising two, three, or four items. This study included evaluation of the depression (DEP) and anxiety (ANX) subscales.

Instructions for the life functioning (LF) scale direct respondents to indicate "How have you been getting along in the following areas of your life?" over the past 2 weeks. The four items cover intimate relationships, social

relationships, work and/or school, and life enjoyment (such as recreation and leisure activities). Clients responded to the items on a 5-point Likert scale ranging from 0 (*terrible*) to 4 (*very well*).

Research has demonstrated that the BHM-20 has good reliability and construct validity. Kopta and Lowry ([19]) evaluated the psychometric properties of the scale using four samples (community adults, college students, college counseling clients, and psychotherapy outpatients). Internal consistency coefficients (coefficient alpha) ranged from .89 to .90 for the GMH score. The WB scale alpha coefficients ranged from .65 to .74; the SYM scale alpha coefficients ranged from .85 to .86; and the LF scale alpha coefficients ranged from .72 to .77.

Construct validity was assessed using discriminant analyses (Kopta & Lowry, [19]). Among all four samples, all four scales distinguished clinical from nonclinical groups. Concurrent validity was assessed by comparing the GMH scale to other measures, including the BASIS-32 (Eisen, Wilcox, Leff, Schaefer, & Culhane, [8]), the Outcome Questionnaire–45 (OQ-45) (Lambert & Finch, [23]), and the Symptom Checklist–R–90 (SCL-R-90) (Derogatis & Savitz, [6]). The analyses showed high correlations between the GMH scale and the other measures of global mental health functioning, with correlations ranging from .81 to .85 (Kopta & Lowry, [19]). Finally, the BHM-20 scales' sensitivity to change, to reflect improvement or deterioration in the mental health of respondents, was evaluated using both the college counseling and the psychotherapy outpatient samples. Comparing intake scores to session 3 scores, all four scales demonstrated significant change (Kopta & Lowry, [19]).

Suicide Monitoring Scale (SMS)

Embedded within the BHM-20, the SMS was developed in collaboration with suicide experts to aid clinicians in tracking and managing suicide risk over the course of treatment in a reliable and valid manner (Kopta, Mond, David, Potruzski, & Doll, [20]). The SMS consists of two items that assess the frequency of suicidal thoughts and estimate the risk of suicide. The first item asks, "In the past two weeks, how much have you been distressed by thoughts of ending your life?" If the client indicates "A little bit," "Sometimes," "Often," or "Almost always" (versus "Never"), then he or she is asked to "Indicate your overall risk of suicide" as either "Extremely high risk," "High risk," "Moderate risk," "Low risk," or "No risk." Based on their responses, clients are categorized into one of four risk levels—no risk, low risk, moderate risk, and high risk.

Bryan, Corso, Rudd, and Cordero ([2]) found that the SMS improved the detection of suicidal clients six-fold relative to standard assessment approaches used by primary care physicians. Kopta and colleagues ([20]) found the internal consistency for the SMS to be adequate in a sample of 912 college counseling clients (Cronbach's alpha = .72). Internal consistency for the SMS in this study was adequate for a two-item scale (Cronbach's alpha = .73).

Procedure

Data were collected as part of the normal clinical routine at all centers. Clients completed the BHM-20 at the beginning of each psychotherapy session via desktop, netbook, or tablet (i.e., iPad™) computer. Each client's responses were de-identified and securely transmitted electronically to CelestHealth Solutions, LLC. Responses were analyzed and scored. The data remain stored on secure servers at CelestHealth Solutions for data analytic purposes.

A report summarizing the various scale and subscale scores was immediately transmitted back to the therapist, so that, if desired, the therapist could discuss the report with the client. The report produces several color-coded graphs and tables. Dose-outcome graphs for the GMH scale, the three primary scales, and symptom subscales show the client's progress across sessions. The Behavioral Health Profile displays color-coded scale scores based on normative data. Using cut-off scores developed according to the clinical significance guidelines

described by Jacobson and Truax ([17]), scale scores are classified according to four levels of mental health distress—normal (no distress), low distress, medium distress, and high distress.

Statistical Analyses

Average intake and termination scores on the scales and subscales for all clients were compared. Following that, the individual change scores were analyzed using Jacobson and Truax's ([17]) methodology of clinical significance. For this approach, the clinical significance cut-off scores that separated the "dysfunctional" from "functional" individuals were calculated for the scales. The dysfunctional population was defined as college counseling clients; the functional population was defined as college students not in counseling (In individual cases, of course, some college counseling clients would not score as "dysfunctional" and many who are not clients would score within the "dysfunctional" range.). We inserted the means and standard deviations for the two groups as reported by Kopta and Lowry ([19]) into Jacobson and Truax's formula that is recommended for calculating these cut-off scores.

A Reliable Change Index (RCI; see Jacobson & Truax, [17]) for each scale was also calculated. The RCI is based on the test–retest reliability of the measure from which the scores are derived. If the magnitude of change for an individual exceeds the RCI, then the pre–post difference is considered to be reliable. See the "Evaluating clinical significance" section for the details of how these two criteria were used to categorize participants.

RESULTS

Intake and Final Scores on the BHM-20: Group Comparisons

Mean intake and final session scores for all clients on the GMH scale, three primary scales (WB, SYM and LF), and two symptom subscales (DEP and ANX) are shown in Table 1. For all six scales, the average termination score was significantly improved from the average intake score, with medium to large effect sizes for all of the paired-sample *t*-tests.

TABLE 1 BHM Scale and Subscale Scores at the Intake Session and at the Final Session (N = 13,803)

Scale	Intake Score M (SD)	Final Score M (SD)	t-value	r ²
Global Mental Health	2.56 (0.65)	2.92 (0.63)	68.1*	.25
Well-Being (WB)	1.76 (0.86)	2.24 (0.87)	62.0*	.21
Psychological Symptoms (SYM)	2.82 (0.67)	3.15 (0.65)	57.9*	.20
Depression Symptoms (DEP)	2.12 (0.96)	2.75 (0.83)	45.6*	.13
Anxiety Symptoms (ANX)	2.31 (0.88)	2.61 (0.93)	55.4*	.18
Life Functioning (LF)	2.16 (0.82)	2.49 (0.84)	60.9*	.21

Note. Possible responses to all scales are on a 0–4 scale, with higher scores indicating better mental health.

**p* < .001.

Intake and final session scores were examined after categorizing clients into the following treatment duration categories: 2–3 sessions, 4–6 sessions, 7–10 sessions, 11–15 sessions, and 16–25 sessions (see Table 2). Paired samples *t*-tests were statistically significant (*p* < .001) for all groups and across all scales. Effect sizes (*r*²) analyses revealed small to medium magnitude of differences for all comparisons.

TABLE 2 BHM Scale and Subscale Scores at Intake and Final Session Categorized According to Treatment Duration (N = 13,803)

Treatment Duration										
	2–3 Sessions (n = 6516)		4–6 Sessions (n = 3449)		7–10 Sessions (n = 1907)		11–15 Sessions (n = 1083)		16–25 Sessions (n = 848)	
Scale	Intake M (SD)	Final M (SD)	Intake M (SD)	Final M (SD)	Intake M (SD)	Final M (SD)	Intake M (SD)	Final M (SD)	Intake M (SD)	Final M (SD)
Global Mental Health	2.59 (0.66)	2.86 (0.64)	2.55 (0.65)	2.94 (0.64)	2.53 (0.65)	3.00 (0.59)	2.53 (0.64)	3.02 (0.58)	2.53 (0.63)	3.05 (0.56)
Well-Being (WB)	1.79 (0.87)	2.16 (0.87)	1.77 (0.85)	2.28 (0.89)	1.73 (0.86)	2.34 (0.83)	1.70 (0.86)	2.34 (0.87)	1.66 (0.84)	2.36 (0.87)
Psychological Symptoms (SYM)	2.85 (0.68)	3.09 (0.66)	2.81 (0.67)	3.16 (0.65)	2.80 (0.66)	3.23 (0.60)	2.80 (0.67)	3.24 (0.63)	2.81 (0.65)	3.27 (0.64)
Depression Symptoms (DEP)	2.34 (0.88)	2.67 (0.84)	2.30 (0.88)	2.77 (0.83)	2.27 (0.89)	2.85 (0.78)	2.25 (0.88)	2.86 (0.79)	2.25 (0.86)	2.88 (0.80)
Anxiety Symptoms (ANX)	2.20 (0.97)	2.52 (0.94)	2.13 (0.96)	2.63 (0.93)	2.12 (0.95)	2.71 (0.89)	2.14 (0.95)	2.77 (0.90)	2.15 (0.92)	2.83 (0.89)
Life Functioning (LF)	2.20 (0.85)	2.45 (0.84)	2.16 (0.81)	2.52 (0.86)	2.11 (0.82)	2.55 (0.81)	2.08 (0.78)	2.52 (0.84)	2.09 (0.76)	2.53 (0.84)
Suicide Monitoring (SMS)	3.64 (0.84)	3.86 (0.55)	3.60 (0.87)	3.87 (0.52)	3.59 (0.86)	3.89 (0.45)	3.59 (0.88)	3.88 (0.52)	3.54 (0.93)	3.90 (0.52)

8 Note. All paired samples *t*-tests were significant at $p < .001$ (*t*-scores ranged from 9.25 to 39.2; effect sizes correspondingly range from $r^2 = .006$ – .10).

Change scores from the intake session to the final session were computed for each scale to examine the differences in the magnitude of change across the duration categories. Omnibus ANOVA tests reveal that there were statistically significant differences between the treatment duration groups on all of the scales. Tukey LSD post-hoc tests revealed that change scores were of a greater magnitude (indicating more positive change) as they remained longer in treatment. After seven sessions, however, there were no significant differences in change scores. That is, change scores were similar between clients in the 7–10 session, 11–15 session, and 16–25 session categorizations. The same pattern was noted for all of the scales.

Clinical Significance Analyses: Categorizing Individual Clients

The BHM-20 scales' cutoff score, which determine whether the client is within the range of normal functioning, and RCI values, which indicate whether the client's improvement is statistically reliable, are shown in Table 3 (Kopta & Lowry, [19] provided the normative data.). Based on these scores and values, each client was categorized into one of five groups shown in Table 4. If the client's score at intake was greater than the cutoff score (indicating that he or she began treatment in the functional population), the client was categorized as "functional-at-intake." The following percentages of clients were functional-at-intake: 31.2% ($n = 4300$) on the GMH scale, 31.7% ($n = 4380$) on the WB scale, 49.3% ($n = 6799$) on the SYM scale, and 28.2% ($n = 28.2$) on the LF scale. These clients were then excluded from the remaining analyses.

TABLE 3 Clinical Significance Cutoff Scores and Reliable Change Indices

Scale	Cutoff Score	RCI Value
Global Mental Health (GMH)	2.93	0.81
Well-Being (WB)	2.16	1.28
Psychological Symptoms (SYM)	2.91	0.76
Life Functioning (LF)	2.64	1.02

TABLE 4 Clinical Significance Analyses

	Reliable Change			
	Deteriorated	No Change	Improved	Recovered ^a
Global Mental Health ($n = 9503$)	351 (3.7%)	4853 (51.1%)	4294 (45.2%)	3088 (71.9% of Improved)
Well-Being (WB; $n = 9423$)	106 (1.1%)	5846 (62.0%)	2553 (27.1%)	2121 (80.1% of Improved)
Psychological Symptoms (SYM; $n = 7004$)	148 (2.1%)	4048 (57.8%)	2791 (39.9%)	2445 (87.6% of Improved)
Life Functioning (LF; $n = 9821$)	144 (1.5%)	6439 (65.6%)	1890 (19.2%)	1467 (77.6% of Improved)

Note. Clinical significance analyses are conducted with the Functional-at-Intake subset removed. Totals (sum of Deteriorated, No Change, and Improved groups) do not add to 100% due to missing data.

^aParticipants categorized as "Recovered" are the subset of the participants categorized as "Improved" who crossed the cut-off score between the norm-based functional and dysfunctional population.

Clinical significance analyses were conducted with the remaining dysfunctional-at-intake clients (see Table 4). If the change value (the difference between the intake and termination scores) exceeded the RCI but was in the direction of deterioration (indicating more pathology), then the client was categorized as "deteriorated." Table 4 shows that fewer than 4% of clients deteriorated on the GMH scale. If the client's change score did not exceed

the RCI, the client was categorized as "no change," which was the case for 51% of the clients on the GMH scale. If the client's GMH change score reliably changed to indicate less pathology, then the client was categorized as "improved"; about 45% of clients fell into this category. Finally, of the clients categorized as "improved," those whose final score also exceeded the cutoff score were categorized as "recovered." The majority (71.9%–87.6%) of "improved" clients also scored as "recovered."

Clients were subdivided into treatment duration groups and compared according to clinical significance categorization. Figure 1 displays the percentage of clients who reliably improved across various treatment durations. Chi-square analyses revealed significant differences between the percentage of clients reliably improved for all of the scales (all p values $<.001$) across the treatment duration groups. However, effect sizes for these analyses were all of small magnitude ($\phi = .10 - .16$). It can be seen that the longer treatment continued, the more likely the client was to be categorized as improved. As suggested in the group analyses, the greatest change tended to occur between clients who remained in treatment from 2–3 sessions, 4–6 sessions, or 7–10 sessions, with the percentage of reliable improvement leveling off for clients remaining in treatment from 11–15 sessions and 16–25 sessions.

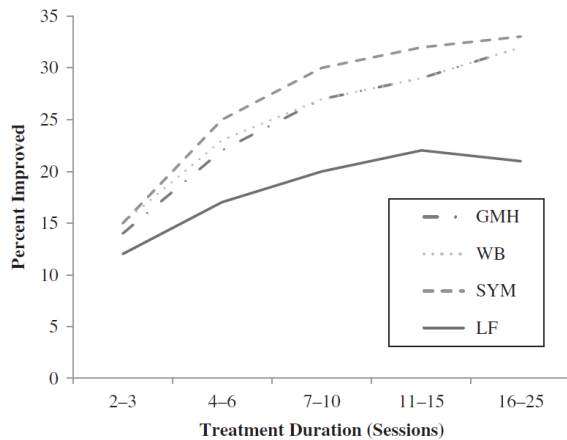


FIGURE 1 Percentage of clients reliably improving across BHM-20 scales across various lengths of treatment duration.

SMS Items

Of the 13,803 clients, 10,689 (77.4%) indicated "Never" being distressed by thoughts of ending their life; those clients did not respond to the second SMS item. The clients who had at least "A little bit" of distress indicated their overall risk of suicide (Table 5). Of those clients who indicated "A little bit" of distress about thoughts of ending their life, almost 60% indicated "No risk" of suicide. In contrast, of those who indicated "Almost always" experiencing distress about suicidal thoughts, 39% (i.e., 80 clients) indicated that their risk of suicide was either "High" or "Extremely high." Of the 3,114 clients who had an initial score of 3 or less on the SMS scale, meaning at least some risk for suicide, 91.1% had an improved SMS score by termination.

TABLE 5 Comparing Responses on the Two Items of the Suicide Monitoring Scale (SMS) at the Intake Session

	Q21: Indicate your overall risk of suicide (row %)				
	Extremely high risk	High risk	Moderate risk	Low risk	No risk
Q10: Thoughts of ending your life?					
Almost always (n = 206)	10.9%	28.1%	17.2%	21.9%	21.9%
Often (n = 595)	1.4%	9.9%	37.1%	35.7%	16.0%

Sometimes (n = 445)	0.4%	1.1%	18.2%	53.9%	26.4%
A little bit (n = 1,868)	0.2%	0.3%	1.7%	37.9%	59.9%
Never (n = 10,689)	(Clients who respond "Never" do not answer Q21)				

The percentage of clients endorsing suicide risk (SMS scores between 0–3) decreased from intake to the last session (22.5% to 7.9%), while the percentage endorsing no risk (SMS = 4) increased from the intake to last session (77.5% to 92.1%). A paired samples *t*-test comparing average first and last session scores on the initial SMS item (which all clients endorse) showed that clients report being at less risk for suicide at the final session ($M = 3.87, SD = 0.53$) than at intake ($M = 3.61, SD = 0.86; t[13,802] = -32.22, p < .001$). The effect size for this analysis was small ($r^2 = .07$).

To investigate the relationship between clients' change in suicide risk and duration in treatment, a one-way ANOVA compared SMS change scores between different treatment duration groups. There was a significant difference in SMS change scores across treatment duration categories ($F [4, 13798] = 6.32, p < .001$), though the effect size was small ($\eta^2 = .001$). Tukey LSD post hoc analysis revealed that reduction in suicide risk was greater for clients who remained in treatment longer. Clients who remained in treatment for 2–3 sessions had less reduction in suicide risk than clients who remained in treatment 7–10 and 16–25 sessions.

Case Vignettes

Practicing UCC therapists provided three case vignettes to demonstrate the usefulness of the CHS-MH in client care. Each vignette presents relevant BHM-20 scale information and therapist impressions of using the CHS-MH. Figures 2–4 can be found online.

Case 1

A college-age male was seen for 26 sessions for treatment of depression. As seen in Figure 2, his GMH and Depression scores indicate poor psychological functioning at intake. The client verbally denied any active suicidality at the first session even though his score on the SMS scale at intake was 1. Though reluctant to discuss suicidality, he stated that he believed he would be safe until the next session, which he willingly scheduled. However, he attempted to hang himself between the first and second sessions, although he decided to seek help at an emergency room during the attempt, which led to a brief psychiatric hospitalization afterward. He was unable to complete fall semester classes. When he returned to therapy at the beginning of the spring semester, his suicide risk remained high (SMS = 1). The therapist shared the following about using the CHS-MH to work with this client:

He was initially mistrustful of me and of the counseling process. I shared the SMS scale results with him at every session as a vehicle to discuss his suicidality. The client became accustomed to seeing his scores and having us discuss them as part of our work together. The client made considerable progress over time [seen in Figure 2]. He passed all his spring classes, became medication compliant, which he had not been for the first part of our work together, and his Depression, GMH, and SMS improved.

Case 2

A college-age female was seen for 18 sessions for treatment of grief and problematic alcohol use. This student used alcohol as a way to cope with grief over her mother's death several years earlier. As seen in Figure 3, an episode of heavy drinking around the third session had a very negative impact on most areas of her functioning. She showed gradual improvement, until another drinking relapse around session 14 caused a drop in her

Drug/Alcohol Use subscale during and subsequent to that date. The therapist indicated that reviewing the CHS-MH results with the patient was crucial in the treatment:

The data proved valuable in helping her see the impact of her drinking on other areas of her functioning. She showed some improvement in the GMH and other BHM scales prior to termination, but it was not clear if her attempts at controlled drinking would ultimately be successful. The BHM data helped define the cyclical nature of her mental health concerns, which we discussed as part of her termination at the time of her graduation.

Case 3

A freshman female was seen for six sessions for treatment of depression at the beginning of Spring semester. Figure 4 shows that she experienced increasing levels of depression and distress as the semester progressed. The therapist commented about using the CHS-MH with this client:

This is an example of how the treatment per se did not produce the hoped for outcome, but the BHM data were still useful. Typical cognitive-behavioral therapy plus starting antidepressant medication did not appear to have significant positive effects. Reviewing her BHM data helped me discuss the lack of treatment progress and aided in communicating with her mother when her suicide score went to red (high risk). The student decided to transfer to a college she felt would be a much better fit with her academic and social style. The BHM data were helpful in catching the lack of counseling progress earlier than I think may have happened without the data.

DISCUSSION

This study provides further evidence that students seeking counseling at UCCs benefit from treatment (cf. Choi, Bushkey, & Johnson, [5]; Nafsinger, Couillard, & Smith, [31]). Overall, clients' average termination score was significantly improved from their average intake score on all of the BHM scales. An examination of the change across different lengths of treatment indicated that improvement was more likely as treatment lasted longer—up to a point, which is consistent with the dose-effect model of psychotherapy outcome (Howard, Kopta, Krause, & Orlinsky, [15]). The results also support the phase model, which proposes that well-being, symptoms and life functioning improve at different rates. Consistent with the phase model (Howard et al., [16]), Figure 1 shows that scores on the life functioning scale were the slowest and least likely to show improvement.

Consistent with the dose-effect model, these results suggested that benefit tends to peak at between 7 to 10 sessions. After that point, there were no significant differences in change scores. Additionally, at that point the average client had achieved a GMH score of 3.00, which exceeds the clinical significance criterion of 2.93. This result supports the effectiveness of brief intervention, which is important because many UCCs have session limits (Gallagher, [11], [12]). Although treatment length is ideally determined by clinical need, the results suggest that 10 sessions might be a rational point at which to terminate treatment for many clients. If a client reaches the 10 session limit without making sufficient progress, then a case review could justify additional sessions. Regularly tracking a client's progress using a brief outcome measure like the BHM-20 is a useful way to determine when it may be appropriate to terminate treatment.

This study supports the idea that the BHM-20 is an effective instrument for the electronic administration, scoring, and tracking of client progress via the CHS-MH (Bryan et al., [3]). The broad-based measure, specifically developed for UCCs, also allows immediate feedback to clinicians on clients' progress and potential problems. The BHM-20 also allows categorization of clients into clinical significance categories, which is useful at both the individual and the clinic-wide level. The system offers each individual UCC summative information regarding client outcomes (e.g., percent improved, percent recovered).

Of particular utility to UCCs might be the specific BHM-20 focus on suicide risk monitoring in light of the recent national attention to suicide prevention on college campuses (Schwartz & Friedman, [36]) and Gallagher's ([12]) finding that 90% of UCC directors report an increase in students seeking services for severe psychological problems. The SMS enables immediate feedback about the status of clients who admit to suicidal ideation or risk. The SMS results provide encouraging data on the ability of UCCs to manage suicide risk with college clients; the longer clients remained in treatment, the more reduction in suicide risk was reported.

Study Limitations

There are several limitations to the present study. First, there was neither control of nor information about the type of treatment provided by the UCCs or by specific counselors. Second, there was no information about the presenting problems or diagnoses of the clients. Another limitation is that the data came from UCCs that elected to utilize the BHM-20 in their practice, which potentially limits the generalizability of the results. Further, there was no control group of comparably symptomatic students not receiving treatment, so we can't be certain whether improved scores would have happened due to time alone. All these limitations are part and parcel of effectiveness research, which seeks to understand how psychotherapy is actually conducted in real-life environments.

Although neither clients nor counseling centers were selected randomly, the large sample size suggests that these clients were likely representative of counseling center clients around the United States. For example, the gender distribution of this sample (62.5% female) was consistent with the 65% reported by Gallagher ([11], [12]) and the National Survey of Counseling Center Directors.

Future Directions

Psychotherapy researchers have tended to adopt a top-down approach to clinical practice. In this approach, science is generated and should be adopted by therapists. Hence researchers and policy makers generate treatment guidelines that are largely determined by findings derived from meta-analytic studies. However, clinics and clinicians, including UCCs across the United States, have instead begun to utilize their own research measures and methods in order to improve provision of services (e.g., Nafsinger et al., [31]; Whipple & Lambert, [41]). It is past time for researchers to pay closer attention to the needs and experiences of clinics in the provision of services and in the utility of research (Castonguay et al., [4]).

The CHS-MH represents a step in this direction. The system was developed in consultation with UCC directors and clinicians for the explicit purpose of improving service provision. It can help clinicians and directors monitor client progress, identify clients at risk of suicide, support adaptive treatment planning, and thereby enhance the likelihood of positive treatment outcomes (Lambert et al., [25]; Leon, Kopta, Howard, & Lutz, [26]; Lutz et al., [28]; Newnham & Page, [32]). The cost to the center is an annual license fee plus the cost of desktop, netbook, or tablet computers which are used by clients to complete the questionnaires. There is no charge for staff training; there are no server or support fees.

SUPPLEMENTAL MATERIAL

Supplemental data for this article can be accessed on the publisher's website.

Note

1. A small amount of change on a measure may not be deemed reliable because of the inherent unreliability of any measure, referred to as measurement error. That is, even if a person's level of depression or anxiety does not change, obtained scores on depression or anxiety at different times are likely to be different. To be deemed

reliable, then, a change in scores must exceed measurement error, which is usually calculated using the test–retest reliability of the measure.

REFERENCES

- Benton, S. A., Robertson, J. M., Tseng, W., Newton, F. B., & Benton, S. L. (2003). Changes in counseling center client problems across 13 years. *Professional Psychology: Research and Practice, 34*, 66–72. doi:10.1037/0735-7028.34.1.66
- Bryan, C. J., Corso, K. A., Rudd, M. D., & Cordero, L. (2008). Improving identification of suicidal patients in primary care through routine screening. *Primary Care and Community Psychology, 13*(4), 143–147. doi:10.1080/17468840802168268
- Bryan, C. J., Kopta, S. M., & Lowes, B. D. (2012). CelestHealth System: A new horizon for mental health treatment. *Integrating Science and Practice, 2*, 7–11.
- Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research: Approaches and applications. In M. J. Lambert (Ed.), *Handbook of psychotherapy and behavior change* (6th ed., pp. 85–133). Hoboken, NJ: John Wiley & Sons.
- Choi, K., Bushkey, W., & Johnson, B. (2010). Evaluation of counseling outcomes at a university counseling center: The impact of clinically significant change on problem resolution and academic functioning. *Journal of Counseling Psychology, 57*, 297–303. doi:10.1037/a0020029
- Derogatis, L. R., & Spitzer, R. L. (1982). The SCL-90-R, Brief Symptom Inventory, and matching clinical rating scales. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (2nd ed., pp. 679–724). Mahwah, NJ: Lawrence Erlbaum Associates.
- Eells, G. T. (2012). Who are we and where are we going? Reflections on counseling services' scope of care. *Journal of College Student Psychotherapy, 26*, 356–258. doi:10.1080/87568225.2012.711137
- Eisen, S., Wilcox, M., Leff, H., Schaefer, E., & Culhane, M. (1999). Assessing behavioral health outcomes in outpatient programs: Reliability and validity of the BASIS-32. *The Journal of Behavioral Health Services and Research, 26*, 5–17. doi:10.1007/BF02287790
- Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., ... Docherty, J. P. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Archives of General Psychiatry, 46*, 971–982. doi:10.1001/archpsyc.1989.01810110013002
- Erdur-Baker, O., Aberson, C. L., Barrow, J. C., & Draper, M. R. (2006). Nature and severity of college students' psychological concerns: A comparison of clinical and nonclinical national samples. *Professional Psychology: Research and Practice, 37*, 317–323. doi:10.1037/0735-7028.37.3.317
- Gallagher, R. P. (2010). *National survey of counseling center directors 2010*. Alexandria, VA: International Association of Counseling Services.
- Gallagher, R. P. (2011). *National survey of counseling center directors 2011*. Alexandria, VA: International Association of Counseling Services.
- Guinee, J. P., & Ness, M. E. (2000). Counseling centers of the 1990s: Challenges and changes. *The Counseling Psychologist, 28*, 267–280. doi:10.1177/0011000000282006
- 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*, 329–343. doi:10.1093/clipsy.9.3.329
- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose–effect relationship in psychotherapy. *American Psychologist, 41*, 159–164.
- Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology, 61*, 678–685.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19. doi:10.1037/0022-006X.59.1.12

- Kitzrow, M. A. (2003). The mental health needs of today's college students: Challenges and recommendations. *NASPA Journal*, 41(1), 167–181.
- Kopta, S. M., & Lowry, J. L. (2002). Psychometric evaluation of the Behavioral Health Questionnaire–20: A brief instrument for assessing global mental health and the three phases of psychotherapy outcome. *Psychotherapy Research*, 12, 413–426.
- Kopta, S. M., Mond, M., David, L., Potruzski, N., & Doll, L. (2010). *Assessing suicidality: Validation of the Suicide Monitoring Scale of the Behavioral Health Measure–20 and implications for psychotherapeutic strategies*. Paper presented at the annual meeting of the International Society for Psychotherapy Research, Monterey, CA.
- Kraft, D. P. (2011). One hundred years of college mental health. *Journal of American College Health*, 59, 477–481. doi:10.1080/07448481.2011.569964
- Lambert, M. J. (2010). *Prevention of treatment failure*. Washington, DC: American Psychological Association.
- Lambert, M. J., & Finch, A. E. (1999). The Outcome Questionnaire. In M. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (2nd ed., pp. 831–869). Mahwah, NJ: Lawrence Erlbaum Associates.
- Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing feedback to psychotherapists on their clients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology*, 61, 165–174. doi:10.1002/jclp.20113
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track client outcome? A meta-analysis. *Clinical Psychology: Science and Practice*, 10, 288–301. doi:10.1093/clipsy.bpg025
- Leon, S. C., Kopta, S. M., Howard, K. I., & Lutz, W. (1999). Predicting patients' responses to psychotherapy: Are some more predictable than others? *Journal of Consulting and Clinical Psychology*, 67, 698–704. doi:10.1037/0022-006X.67.5.698
- Locke, B. D., Buzolitz, J. S., Lei, P., 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. doi:10.1037/a0021282.
- Lutz, W., Lowry, J. L., Kopta, S. M., Einstein, D., & Howard, K. I. (2001). Prediction of dose-response relations based on client characteristics. *Journal of Clinical Psychology*, 57, 889–900. doi:10.1002/jclp.1057
- Maruish, M. (Ed.) (1999). *The use of psychological testing for treatment planning and outcomes assessment* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Much, K., & Swanson, A. L. (2010). The debate about increasing college student psychopathology: Are college students really getting "sicker?" *Journal of College Student Psychotherapy*, 24, 86–97. doi:10.1080/87568220903558570
- Nafsinger, M. A., Couillard, G. C., & Smith, T. B. (1999). Evaluating therapy outcome at a university counseling center with the college adjustment scales. *Journal of College Counseling*, 2, 3–13. doi:10.1002/j.2161-1882.1999.tb00137.x
- Newnham, E. A., & Page, A. C. (2010). Bridging the gap between best evidence and best practice in mental health. *Clinical Psychology Review*, 30, 127–142. doi:10.1016/j.cpr.2009.10.004
- Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review*, 21, 421–446.
- Project Match Research Group. (1997). Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. *Journal of Studies on Alcohol*, 58, 7–29.
- Reese, R. J., Norsworthy, L. A., & Rowlands, S. R. (2009). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, Research, Practice, Training*, 46, 418–431. doi:10.1037/a0017901
- Schwartz, L. J., & Friedman, H. A. (2009). College student suicide. *Journal of College Student Psychotherapy*, 23, 78–102. doi:10.1080/87568220902743058
- Sperry, L., Brill, P. L., Howard, K. I., & Grissom, G. R. (1996). *Treatment outcomes in psychotherapy and psychiatric interventions*. New York, NY: Brunner/Mazel, Inc.

- Twenge, J. M., Gentile, B., DeWall, C. N., Ma, D., Lacefield, K., & Schurtz, D. R. (2010). Birth cohort increases in psychopathology among young Americans, 1938-2007: A cross-temporal meta-analysis of the MMPI. *Clinical Psychology Review, 30*, 145–154. doi:10.1016/j.cpr.2009.10.005
- 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*, 319–337. doi:10.1177/1473325011401468
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology, 69*, 875–899. doi:10.1037/0022-006X.69.6.875
- Whipple, J. L., & Lambert, M. J. (2011). Outcome measures for practice. *Annual Review of Clinical Psychology, 7*, 87–111. doi:10.1146/annurev-clinpsy-040510-143938
- Wolgast, B. M., Rader, J., Roche, D., Thompson, C. P., von Zuben, F. C., & Goldberg, A. (2005). Investigation of clinically significant change by severity level in college counseling center clients. *Journal of College Counseling, 8*, 140–152. doi:10.1002/j.2161-1882.2005.tb00081.x