



AUTHORS

Stacy J. Priniski, M.S.
*University of
Wisconsin-Madison*

Erin Winterrowd, Ph.D.
Regis University

Abstract

Academic outcomes assessment in student affairs is integral for both service improvement and demonstrating the unit's value to the university's academic mission. However, identifying the right measures is challenging. We implemented three common measures (pre-post self-reported academic functioning, retrospective perceptions of service impact, and semester grades) within a single counseling center client sample ($N = 368$) and examined the impact of measure selection on the representativeness of client subsamples and the conclusions that might be drawn about service effectiveness. Students' perceptions of academic outcomes suggested greater impact than pre-post or grade measures overall but all three showed positive effects for clients identified as academically at-risk at baseline. No single measure captured a fully representative sample of clients. Rather than providing evidence for one "best" measure, results point to the importance of using multiple measures to assess academic outcomes. Implications for best practices in service outcomes assessment are discussed.

Proof in the Pudding: Implications of Measure Selection in Academic Outcomes Assessment

The American College Personnel Association's (1994) release of the Student Learning Imperative (re)sparked a dedication to improving assessment practices and a corresponding call to document the impact of student affairs services on student learning and development (Reynolds & Chris, 2008; Upcraft & Schuh, 1996). Pressure for student affairs divisions to demonstrate their value to the university's academic mission has increased simultaneously (Nafziger, Couillard, & Smith, 1999; Varlotta, 2012). However, with a plethora of assessment approaches and measures available it can be difficult to determine the best way to assess service impact. In this article we explore the ways in which practitioners from one unit (campus mental health services) have measured academic outcomes, and we investigate how measure selection influences conclusions about service effectiveness.

Outcomes Assessment in Campus Counseling Centers

Campus counseling centers (CCCs) provide a useful context for studying academic outcomes assessment for two reasons. First, the mechanisms by which CCC services might influence academic functioning are evident in the literature. CCCs improve students' psychological well-being (e.g., Minami et al., 2009), and psychological well-being is an important predictor of academic well-being (e.g., Miller & Markman, 2007; Stallman, 2010). Second, literature on the relationship between counseling and academics includes a variety of academic outcome measures with mixed results that give insight into the potential differences among them.

CORRESPONDENCE

Email
spriniski@wisc.edu

Lambert and Hawkins' (2004) conceptual model of CCC assessment provides a useful framework for measure selection. The model characterizes outcome measures by content (the construct of interest), source (e.g., client, therapist), method of data collection (e.g., self-report, behavioral), and time orientation (e.g., state vs. trait measures). Importantly, the model considers psychometric strength (reliability, validity, sensitivity to change), applicability, and practicality, emphasizing that not all measures are equally suited to capture a given outcome. In light of this model, we review three commonly used

measures of service impact on academic functioning: two self-report measures (pre-post self-reports of academic functioning and retrospective self-reports of counseling's impact) and one institutional measure (grades), and then report a field test of those measures within a single CCC sample.

Assessing Academic Outcomes

Pre-post self-reports. A common practice for CCCs is to assess the impact of services on academic functioning using pre-post measurements of school-related “symptoms” (e.g., difficulty keeping up with schoolwork, thoughts of leaving college). These measures fit seamlessly into existing assessments of self-reported psychological symptoms at most CCCs, and a number of validated questionnaires that include academic functioning are available (e.g., Counseling Center Assessment of Psychological Symptoms (CCAPS); Locke et al., 2011). Studies that assess academic outcomes using pre-post measures generally find that clients’ academic functioning improves over a set number of appointments (six, on average) whereas academic functioning remains unchanged over a similar time period among non-clients (DeStefano, Mellot, & Petersen, 2001; Lockard, Hayes, McAleavey, & Locke, 2012; Nafziger et al., 1999).

Retrospective self-reports. Another commonly used approach is to ask students directly whether they feel services helped them academically, typically after a set number of appointments or at the end of the semester. These retrospective self-report measures are often created in-house so most have not been validated in the literature. However, existing studies suggest that they are internally reliable and strongly correlated with other learning outcomes of counseling (Winterrowd, Priniski, Achter, & Abhold, 2016) and may be better tailored to specific student affairs units (Erwin & Sivo, 2001). National surveys of CCC directors indicate that over 60% of centers collect these measures and most find that clients report that counseling has a positive impact on their academic functioning (Gallagher, 2011) with the few published studies also supporting that conclusion (Winterrowd, et al., 2016; Reynolds & Chris, 2008; Turner & Berry, 2000).

Grade point average. Grade point average (GPA) information can be asked from students directly or taken from institutional records, with the latter being more common in the literature. Researchers typically measure changes in GPA from before service delivery to after, or test the relationship between extent of participation in services (e.g., number of appointments) and grades, controlling for prior performance (e.g., high school GPA). As an academic outcome of student services, GPA resonates with campus administration and students alike, although it is unclear whether GPA is sensitive to the kinds of changes that counseling is intended to create (Illovsy, 1997; Lockard et al., 2012). Most studies show null effects (Lee, Olson, Lock, Michelson, & Odes, 2009; Illovsy, 1997), although studies examining the impact of counseling for academically “at-risk” students (underprepared first-year students, Cholewa & Ramaswami, 2015; students on academic probation, Wlazelek & Coulter, 1999) find positive impacts of counseling on GPA.

Summary

Together, this body of literature highlights three key points about measure selection. First, researchers and practitioners can and do choose from a wide variety of academic measures to assess service impact. Second, the methods of data collection employed with each outcome measure differ in ways that impact the sample of clients evaluated. For example, pre-post self-reports can only be collected from clients who attend a certain number of appointments, and change in GPA can only be collected from students with continuous enrollment in credit-bearing courses. It is unclear whether different measures of academic outcomes capture a representative subsample of clients—which raises concerns about the validity of the conclusions that are drawn from them. Finally, it appears that different measures lead to different conclusions about the relationship between counseling and academics. Specifically, studies using pre-post self-report measures (e.g., Lockard et al., 2012) and retrospective self-reports (e.g., Winterrowd, et al., 2016) found positive effects of counseling on academic functioning whereas the results of studies that utilized GPA had mixed results (e.g., Cholewa & Ramaswami, 2015; Lee et al., 2009).

In this article we explore the ways in which practitioners from one unit (campus mental health services) have measured academic outcomes, and we investigate how measure selection influences conclusions about service effectiveness.

The Lambert and Hawkins (2004) model emphasizes the importance of measure selection and the dimensions upon which measures can vary. Implicit in their discussion is the idea that diverse measures can result in diverse conclusions about service effectiveness, highlighting the importance of using multiple measures of learning outcomes in counseling and in higher education generally (e.g., Astin & Antonio, 2012; Schuh, 2011; Suskie, 2009). The mixed results of the previous studies appear to support that assertion. However, because pre-post self-reports, retrospective reports, and grades have never been compared within a single sample, it is unclear whether inconsistent results in the literature reviewed here are due to differences in the measures themselves or other factors (e.g., differences in samples, timeframe of assessment, quality of services provided).

Current Study

In the current study we investigated the impact of measure selection in service evaluation using a comprehensive framework for assessment of academic outcomes within a single counseling center client sample.

In the current study we investigated the impact of measure selection in service evaluation using a comprehensive framework for assessment of academic outcomes within a single counseling center client sample. We compared the impact of counseling services indicated by three commonly used measures of academic outcomes—pre-post self-reports of academic functioning, retrospective reports of counseling impact on academics, and semester grades—for all clients, generally, and for clients identified at baseline as academically at-risk in particular (see Table 1 for a summary of measure features). In line with the Lambert and Hawkins (2004) model, the validity of the measures and the practical implications of measure selection were also of interest. Therefore, we examined differences in subsample characteristics to determine whether each measure captured a representative sample of the client population.

Table 1. *Characteristics of Three Common Measures of Academic Outcomes on the Dimensions of the Lambert and Hawkins (2004) Model.*

	Pre-Post Self-Reports	Retrospective Self-Reports	Grade Point Average (GPA)
Content	changes in academic functioning over the course of counseling	perceptions of counseling's impact on academic functioning	course performance
Source	student	student	university records
Method of data collection	self-report	self-report	institutional
Time orientation	state measure; varies day to day	state measure; varies day to day	trait measure; varies semester to semester
Psychometrics	several validated measures available; sensitive to change	typically in-house measures (not validated), but existing data suggests reliability and validity; sensitive to change	highly externally valid; less sensitive to change
Applicability	high	high	high

We hypothesized that pre-post self-reports, retrospective self-reports, and grades would each yield unique results within a single sample of counseling center students. Such results would suggest that the mixed findings of prior studies might be due to differences in the measures themselves and demonstrate the importance of measure selection in assessing the precise aspects of academic functioning each student service intends to support.

Method

Participants

Data were collected from 368 undergraduate students who received counseling services during the fall semester at a midsize predominantly undergraduate institution in the Midwest. Participants identified as White (89%), African American/Black (4%), Asian American/Asian (2%), Multiracial (2%), American Indian or Alaskan Native (1%), Hispanic/Latino(a) (1%), or other self-identified ethnicities (1%; 0.3% unreported). Women were 69% of the sample, men 30%, transgender individuals 0.3%, and other self-identified genders 0.5%. Participants were 25% first-year students, 25% sophomores, 20% juniors, and 29% seniors (1% unreported), with a mean age of 21.12 ($SD = 3.98$; 2% unreported).

Measures

Demographics and Presenting Concerns. From the counseling center intake paperwork we collected information about participating clients' gender, race/ethnicity, age, and year in school. We also noted whether clients selected "school and grades" as one of their reasons for seeking counseling services (on a 29-item presenting concerns checklist). This was used as a baseline measure of academic functioning and one indicator of being academically at-risk.

Counseling Center Assessment of Psychological Symptoms- Academic Distress Scale. The counseling center administered the long form of the Counseling Center Assessment of Psychological Symptoms (CCAPS-62; Locke et al., 2011) at intake and the short form (CCAPS-34; Locke et al., 2012) at the fifth appointment. The CCAPS is a self-report questionnaire that measures changes in psychological well-being generally and across various mental health subscales. We used the four-item Academic Distress Subscale to provide a baseline measure of academic functioning as well as examine changes in academic distress ("It's hard to stay motivated for my classes," "I am not able to concentrate as well as usual," "I feel confident that I can succeed academically" (reversed), and "I am unable to keep up with my schoolwork"). Participants respond via a four-point Likert-type scale with subscale scores > 2.75 considered "elevated," an indicator of being academically at-risk. The measure had high internal consistency in this sample ($\alpha = .82$ at intake, $.83$ at fifth appointment).

Learning Outcomes and Satisfaction Survey- Academic Outcomes Scale. The Learning Outcomes and Satisfaction Survey (LOS; Winterrowd, et al., 2016) measures client perceptions of counseling outcomes and satisfaction with services. The Academic Outcomes (AO) scale assesses the extent to which clients feel counseling helps their academics, with four items ("Counseling has helped with my academic performance," "Counseling has increased my academic motivation and/or attendance," "Counseling has helped me to focus better on my academics," and "Counseling has helped me stay at school") scored on a five-point Likert-type scale. The scale had high internal consistency in this sample ($\alpha = .83$).

Grade Point Average. Participants' semester grade point averages (GPAs) were collected from the university's Institutional Research Office for the semester prior to counseling (baseline) and the end of the semester in which they received services. Prior-semester GPA—specifically whether students were below the cutoff for academic probation (< 2.0 GPA)—was also used as an indicator of being academically at-risk. Finally, we collected clients' high school GPA, which is commonly used to control for individual differences in academic performance in studies examining GPA (e.g., Lee et al. 2009).

Procedure

Questionnaire data were collected at the counseling center in two stages. All clients completed intake questionnaires (demographics, presenting concerns, CCAPS-62, research informed consent) prior to their first appointment and follow-up measures (CCAPS-34, LOS) at their fifth appointment (defined as intake plus four individual and/or group appointments). This allowed the counseling center to use the questionnaires for clinical purposes in addition to keeping the staff blind to which clients were participating in the study. The fifth appointment was chosen for outcome data collection to maximize both the potential for measurable change and the number of participants (Gallagher, 2011). Questionnaire data and the total number of individual counseling sessions attended during the semester were shared with the researchers for consenting clients only¹. To protect confidentiality, counseling center and institutional data were linked by student identification number so that no client names were used. The study was approved by the university's Institutional Review Board.

¹The counseling staff counted both individual and group counseling sessions toward the total number of appointments for the purpose of collecting outcome data after the fifth appointment. However, these two types of appointments are tracked with different systems, and the counseling center only released data on the number of individual counseling sessions attended by each client. Therefore, all analyses including number of appointments utilize the number of individual counseling sessions.

Results

We examined the impact of measure selection in student affairs assessment research and practice by analyzing differences among three measures of academic outcomes (pre-post self-reports, retrospective reports, and grades) in a single sample of students using mental health services. We compared (1) the representativeness of each subsample (an indicator of the validity of the measure for capturing overall client outcomes) and (2) the conclusions drawn from each measure about academic outcomes of all participants generally, and of participants who were academically at-risk in particular. Measure statistics and correlations are presented in Table 2, and a summary of results by measure is presented in Table 3.

Table 2. *Measure Statistics and Intercorrelations of Baseline Academic Functioning and Academic Outcomes.*

1. Intake CCAPS Academic Distress Scale ^a	-						
2. 5th Appointment CCAPS Academic Distress Scale ^a	.68***	-					
3. Presenting Concerns: School or Grades ^b	.50***	.18*	-				
4. LOS Academic Outcomes Scale ^c	-.01	-.16	.10	-			
5. High School GPA ^d	-.17**	-.29**	-.15*	.05	-		
6. Prior-Semester GPA ^d	-.34***	-.10	-.29***	.10	.29***	-	
7. Current-Semester GPA ^d	-.35***	-.53***	-.26***	.16	.44***	.63***	-
<i>Cronbach's α</i>	.82	.83		.83			
<i>Mean</i>	1.74	1.69	0.46	3.50	3.24	2.79	2.69
<i>Standard Deviation</i>	1.08	1.00	0.50	0.73	0.43	0.82	0.94
<i>N</i>	365	122	368	117	292	240	350

^aPre-post self-reports from Counseling Center Assessment of Psychological Symptoms (CCAPS) ^b"School or Grades" selected from presenting concerns checklist intake: 1 = selected, 0 = not selected ^cRetrospective self-reports from Learning Outcomes and Satisfaction Survey (LOS) at the 5th appointment ^dGrade point averages (GPA) from institutional records * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3. *Comparison of Subsample Representativeness and Counseling Services Impact Across Measures of Academic Outcomes.*

	Pre-Post Self-Reports: CCAPS Academic Distress (AD) ^a	Retrospective Self-Reports: LOS Academic Outcomes (AO) ^b	Grades: Change in Semester GPA	Grades: Relationship Between Number of Appointments and GPA
Description	change in distress scores from the 1 st (intake) to 5 th appointment (average of 4 items on a 4-point Likert-type scale)	students' retrospective perceptions that services helped them academically (average of 4 items on a 5-point Likert-type scale)	change in GPA from the semester prior to services (spring) to the semester in which services were received (fall; on a 4.0 scale)	predicting semester GPA from number of appointments attended, controlling for prior academic performance (i.e., high school GPA; both on a 4.0 scale)
Subsample	students who attended five or more appointments; $n = 121$	students who attended five or more appointments; $n = 117$	sophomore through senior students with continuous enrollment; $n = 226$	students with available high school GPA data; $n = 283$
Criteria for identifying academically at-risk clients	elevated AD scores at intake (> 2.75); $n = 22$ (18%)	listed "school or grades" among their reasons for seeking counseling; $n = 46$ (39%)	academic probation (GPA < 2.0); $n = 35$ (15%)	academic probation (GPA < 2.0); $n = 26$ (14% of those enrolled in college the prior semester)
Under-represented in the subsample	first-year students	first-year students	Students without continuous enrollment (e.g., first-year students, transfer students)	Older students (e.g., non-traditional aged students, veterans, international students, transfer students)
Academic outcomes	on average: no change ($M_{pre} = 1.67$; $M_{post} = 1.68$)	on average: students perceived that services helped academically ($M = 3.50$)	on average: no change ($M_{spring} = 2.81$; $M_{fall} = 2.84$)	on average: no relationship between number of appointments and GPA (each session associated with an increase of 0.03 grade points)
	among academically at-risk students: significant reductions in academic distress ($M_{pre} = 3.19$; $M_{post} = 2.66$)	among academically at-risk students: somewhat stronger perceptions that services helped academically ($M = 3.59$)	among academically at-risk students: significant increases in GPA ($M_{spring} = 1.37$; $M_{fall} = 1.93$)	among academically at-risk students: positive but non-significant relationship (each session associated with an increase of 0.14 grade points)

Note. CCAPS = Counseling Center Assessment of Psychological Symptoms; LOS = Learning Outcomes & Satisfaction Survey; GPA = grade point average. ^aLocke et al. (2011, 2012); ^bWinterrowd et al. (2016)

Representativeness of the Subsamples

We first investigated whether different measures captured academic outcomes for representative subsamples of clients. We began by comparing the subsample of clients that attended five or more appointments and completed outcome measures (i.e., those eligible for analyses of pre-post self-reports and retrospective self-reports: 123 clients, 33% of the total sample²) to the full client sample in terms of gender, race/ethnicity, age, and year in school. First-year students were underrepresented in this subsample (13% vs. 25% of the full sample), $\chi^2(3, N = 122) = 8.55, p = .04$. There were no differences by gender, race/ethnicity, or age, $p > .30$.

We then considered the subsample of clients with available data for two common analyses of semester GPA: change in GPA from the semester prior to counseling (spring) to the semester in which counseling services were received (fall), and the relationship between number of appointments and semester GPA, controlling for high school GPA. Change in GPA was only available if the client was continuously enrolled in credit-bearing courses from spring to fall (226 clients, 61% of the total sample). This excluded 128 clients who were not enrolled in spring (most often because they were first-semester students in fall; $n = 86$, leaving only four first-year/second-semester students), 14 clients who were not enrolled in fall (often because they withdrew from all their courses; $n = 10$), four clients who were not enrolled in either semester, and 46 clients who were missing semester GPA data for other reasons (e.g., taking only noncredit-bearing courses, being a transfer student, or taking the semester off). Despite low representation among first-year students this subsample did not differ from the total sample on gender, race/ethnicity, or age, $p > .10$.

The subsample with available data for examining the relationship between number of appointments and fall GPA, controlling for high school GPA ($n = 283$; 77% of the total sample), excluded 14 clients without fall-semester GPA data and 76 clients for whom the university did not collect high school GPA data (e.g., nontraditionally aged students, veterans, international students). Accordingly, this subsample was younger than the total sample, $t(278) = -9.79, p < .001$. There were no significant differences in gender, race/ethnicity, or year in school, $p > .15$.

In sum, subsamples varied considerably across academic measures, both in size (33–77% of the total sample) and representativeness in terms of age and year in school. First-year students were underrepresented in analyses involving measures collected at the fifth appointment (i.e., pre-post and retrospective self-reports) and systematically excluded from analyses involving change in GPA. Analyses of the relationship between the number of appointments and GPA, with high school GPA as a covariate, underrepresented older students. None of the measures appeared to exclude students of a particular gender or race/ethnicity.

Conclusions Regarding Service Outcomes

Next we investigated whether different measures of academic outcomes would point to the same conclusions about service impacts, both for all clients generally and for clients identified as academically at-risk in particular. For these analyses, we identified clients as academically at-risk using baseline measures that paralleled each outcome measure. For change in CCAPS Academic Distress we used the CCAPS cutoff score for elevated Academic Distress at baseline (> 2.75). For retrospective reports of whether counseling helped academically we used clients' baseline presenting concerns (i.e., whether they listed school and grades among their reasons for seeking counseling). For both analyses of semester GPA (i.e., change in GPA, relationship between number of appointments and GPA), we used clients' academic probation status (prior semester GPA < 2.0).

Pre-post assessments. We first examined academic impact using changes in clients' CCAPS Academic Distress (AD) scores from intake (baseline) to the fifth appointment ($n = 121$ ³). Intake (baseline) scores from the CCAPS Academic Distress (AD) scale revealed

Together our results demonstrate some of the potential consequences of measure selection, highlighting the importance of these choices for best practice in service outcomes assessment.

²Thirteen clients attended five or more appointments but did not complete the outcome measures.

³Two clients who attended five or more appointments and completed the retrospective self-report measure did not complete the pre-post Academic Distress measure.

One finding was consistent across measures: academic outcomes weremost positive for clients identified as academically at-risk at baseline.

These three measures of academic functioning worked differently, even within a single subsample of clients, highlighting the importance of measure selection not just for documenting academic outcomes but also for identifying clients who are most in need of academic support.

low to moderate levels of academic distress overall ($M = 1.67$, $SD = 1.00$); however 18% of these clients ($n = 22$) fell in the elevated range and were therefore identified as academically at-risk. Average AD scores at the fifth appointment ($M = 1.68$, $SD = 0.99$) did not differ from intake, $t(120) = -.09$, $p = .93$, despite improvements in CCAPS scores for overall (nonacademic) well-being, $t(120) = 2.19$, $p = .03$. However, the subset of academically at-risk clients (i.e., those who had elevated academic distress at intake and attended five or more appointments, did show a significant reduction in Academic Distress from intake ($M = 3.19$, $SD = 0.41$) to the fifth appointment ($M = 2.66$, $SD = 0.91$), $t(21) = 3.23$, $p = .004$. Importantly, this indicates an improvement from an average score above the 2.75 cutoff for elevated Academic Distress to an average below the clinical cutoff.

Retrospective self-reports. Learning Outcomes and Satisfaction Survey Academic Outcomes (AO) scores were collected at the fifth counseling session ($n = 117^4$). Just under half of the clients who completed the AO reported that counseling helped increase their academic motivation and/or attendance (43%), academic focus (50%), and academic performance (49%), and helped them stay in school (48%). The resulting AO scale mean of 3.50 ($SD = 0.73$) was significantly higher than the scale's neutral midpoint, $t(116) = 7.31$, $p < .001$. In addition, the clients who listed school or grades among their reasons for seeking counseling (an indicator of being academically at-risk) were especially likely to report that counseling had a positive impact on their academics five sessions later ($n = 46$, $M = 3.59$, $SD = 0.72$).

Semester grades. We then examined academic impact using changes in clients' semester GPAs ($n = 226$). Clients' average prior (spring) GPA was 2.81 ($SD = 0.83$), with 35 clients (15% of this subsample) below the cutoff for academic probation (< 2.0 GPA). On average, current semester (fall) GPAs ($M = 2.84$, $SD = 0.87$) were not significantly higher than the prior semester, $t(225) = 0.46$, $p = .65$. The subset of clients who were on academic probation did make significant improvements in GPA, however ($M_{spring} = 1.37$, $SD = 0.57$; $M_{fall} = 1.93$, $SD = 0.91$), $t(34) = 3.43$, $p = .002$, with 16 clients moving off of academic probation.

We also examined fall GPA as a function of the number of individual counseling appointments attended, controlling for high school GPA ($n = 283$). The relationship between number of individual counseling appointments and fall GPA was positive but small, $b = 0.03$, $SE = 0.02$, $t(280) = 1.25$, $p = .21$. Among clients on academic probation ($n = 26$) each additional individual counseling session was associated with an increase of 0.14 grade points in GPA, an effect that did not reach statistical significance, $b = 0.14$, $SE = 0.08$, $t(23) = 1.62$, $p = .12$, but may be clinically significant for these students.

Ancillary Analyses: Impact of Baseline Measure Selection

We assessed academic outcomes of counseling services using pre-post self-reports, retrospective self-reports, and grades in one CCC sample and replicated the pattern of mixed results found in prior research using disparate samples. Measure selection appeared to influence both the conclusions that could be drawn about counseling impact and the validity of those conclusions (due to the non-representativeness of the samples). However, one finding was consistent across measures: academic outcomes were most positive for clients identified as academically at-risk at baseline. This underscores the importance of selection of baseline measures in addition to outcome measures. Therefore, we conducted ancillary analyses with clients who had data from all three baseline measures (intake CCAPS Academic Distress, presenting concerns, prior-semester GPA) to examine the impact of measure selection on identification of academically at-risk clients.

There were 240 clients (65% of the total sample) with available data on all three baseline measures. Of these, 110 unique clients were identified as academically at-risk by at least one measure; 62 clients were in the dysfunctional range for CCAPS Academic Distress scores, 76 clients listed school and grades among their presenting concerns, and 38 clients were on academic probation (prior-semester GPA < 2.0). However, only 15 clients (13.6% of academically at-risk clients) were identified as struggling by all three measures, and only 49

⁴Six clients who attended five or more appointments and completed the pre-post Academic Distress measure did not complete the LOS-AO

clients (44.5%) were identified as struggling by two measures. In other words, many clients with low GPAs were not distressed about academics, and many clients with higher levels of academic distress or low GPAs did not report school and grades as a primary reason for seeking counseling. These three measures of academic functioning worked differently, even within a single subsample of clients, highlighting the importance of measure selection not just for documenting academic outcomes but also for identifying clients who are most in need of academic support.

Discussion

We compared three types of academic outcome measures (i.e., pre-post self-reports, retrospective self-reports, and grades) within a single counseling center client sample and found that measure selection impacted both the representativeness of the subsample and the conclusions that might be drawn about the effectiveness of services. No one measure captured a fully representative sample on its own: subsamples differed in size (33–77% of the total sample) and in representativeness in terms of age and year in school. Retrospective self-report measures demonstrated positive academic impacts for all clients, on average, and particularly for academically at-risk students. In contrast, pre-post self-report and institutional (GPA) measures showed positive impacts for academically at-risk students only. Interestingly, ancillary analyses revealed that diverse baseline measures resulted in unique groups of students being identified as academically at-risk in the first place. Together our results demonstrate some of the potential consequences of measure selection, highlighting the importance of these choices for best practice in service outcomes assessment.

Representativeness of Measure Subsamples

Inherent in choosing an assessment method is selecting the subsample of students with available data. Consistent with previous literature (e.g., DeStefano et al., 2001), pre-post and retrospective self-reports in this study captured academic outcomes for students who attended a minimum number of appointments (e.g., five) but excluded those who attended fewer (a majority in this study). This demonstrates the dramatic impact of timing of assessment: collecting outcomes at the fifth appointment excluded two-thirds of the students receiving services and also underrepresented first-year student clients. A shorter time frame minimizes attrition (and potentially increases first-year student representation) but longer time frames may maximize opportunities for academic impact.

For grades, using change in semester GPA from before counseling to after (similar to Illovsy, 1997; Wlazelek & Coulter, 1999) may be ideal for students with continuous enrollment but it underrepresents students in their first semester (i.e., first-year and transfer students). These exclusions are particularly problematic because first-year and transfer students may be more likely to struggle academically than their peers (Berger & Malaney, 2003; Lee et al., 2009). Assessing the relationship between number of individual counseling appointments and semester grades, controlling for high school GPA (similar to Lee et al., 2009), provided the largest subsample of students. However, this assessment still underrepresented nontraditional students without high school GPA information (e.g., older students, veterans)—a group with noted differences in academic needs (e.g., Spitzer, 2000). Together, these results call into question the extent to which any single measure can be used to capture academic outcomes representative of the whole client population.

Academic Outcomes of Counseling Services

The existing research on academic outcomes of counseling services is mixed (e.g., Lee et al., 2009; Lockard et al., 2012; Turner & Berry, 2000) and this study provides some insights into why that might be the case. We implemented three common measures of academic outcomes—pre-post self reports, retrospective reports, and grades—and replicated the mixed results of prior research within a single sample, suggesting that the apparent inconsistency in the literature may be due, at least in part, to differences in academic outcome measures. The Lambert and Hawkins (2004) model illuminates some of the important differences among these measures, including variation in content (changes in academic functioning vs. perceptions of being helped vs. course performance), source (client vs. university records),

This demonstrates the dramatic impact of timing of assessment: collecting outcomes at the fifth appointment excluded two-thirds of the students receiving services and also underrepresented first-year student clients.

Together, these results call into question the extent to which any single measure can be used to capture academic outcomes representative of the whole client population.

method (self-report vs. institutional data), and time orientation (self-reports vary day to day, whereas GPA varies semester to semester; Table 1). This is consistent with other research in higher education that highlights variation in outcomes assessment with different sources (e.g., Sexton, 1996) and methods (e.g., Bowman, 2013). Our results suggest that variability in sample characteristics may also contribute to the mixed findings of prior research—a result not explicitly addressed in the Lambert and Hawkins model (2004).

Implications for Best Practice in Service Outcomes Assessment

Our results demonstrate empirically what many practitioners might have guessed intuitively—that differences among pre-post self-reports, retrospective reports, and grades are more profound than simple variations in operationalization; they capture discrete academic outcomes.

This study has important implications for research and practice. Our results suggest that measure selection plays a fundamental role in demonstrating service effectiveness. Specifically, the “best” measure for capturing academic impact appears to depend on which subsample of clients and which aspect of academic functioning researchers or practitioners most want to assess. Practitioners should therefore (1) determine their specific service goals and which aspects of academic functioning they intend to support, (2) identify academic outcome measures consistent with those goals, (3) choose appropriate baseline measures, and (4) determine the best data collection time frame to capture the intended outcomes for the majority of clients and/or targeted client groups.

Our results demonstrate empirically what many practitioners might have guessed intuitively—that differences among pre-post self-reports, retrospective reports, and grades are more profound than simple variations in operationalization; they capture discrete academic outcomes. For example, as depicted in Table 2, responses to retrospective self-reports of service perceptions were unrelated to pre-post symptom questionnaires and GPA. Using multiple measures in combination could therefore help researchers and student affairs practitioners alike to better understand the students’ academic experiences and to document the impact of student services on many different aspects of academic functioning (Astin & Antonio, 2012; Schuh, 2011; Suskie, 2009). In addition, including both self-report and institutional or observational outcomes increases confidence in conclusions about service outcomes (Sexton, 1996) and protects against the risk of using self-reports solely as a proxy for student learning or growth (e.g., Bowman, 2013). We encourage researchers and practitioners to consider a variety of academic outcomes that might be consistent with their service goals, including those examined here as well as others (e.g., academic self-efficacy, engagement, satisfaction; see Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006). Many of these variables have been considered predictors of academic achievement (i.e., grades) but can be important outcomes in and of themselves.

Assessment of academic outcomes continues to be of paramount importance in student affairs—as best practice and as a means of demonstrating each unit’s value in supporting the academic mission of the university.

In this study we analyzed service outcome data statistically. We hope our analyses give practitioners some ideas of ways they can look at their own outcome data. However, we recognize that many counseling centers (and other student affairs services) have small client populations and/or limited staff and resources for statistical analyses. Certainly practitioners could examine their data descriptively. In fact, some assessments (such as the CCAPS assessment we used in this study) include in their user manuals guidance on how to detect and interpret change over time, without statistical analyses. Even if a campus or center is too small to collect meaningful data in any given semester or year, intentional and systematic measure selection will allow for examination of trends in service utilization and outcomes across multiple years or in collaboration with multiple centers (e.g., Winterrowd et al., 2016).

As the Lambert and Hawkins (2004) model emphasizes, it is important that assessments are applicable and practical. Ultimately, outcomes assessments will only lead to service improvement if they are useful to practitioners. Therefore, the “best” measures and methods can and should vary unit to unit and campus to campus. We hope that our study highlights some of the considerations practitioners might take into account when selecting outcome measures and that our suggestions will help student affairs units maximize their opportunities to demonstrate their value and further improve their services.

Limitations

The current study provides a direct comparison of several commonly used measures

of academic outcomes (pre-post self-reports, retrospective reports, and grades, including both change in GPA and the relationship between number of appointments and GPA). However, this study is by no means a comprehensive comparison of all assessment designs and measures. We examined self-reported academic outcomes at the fifth counseling appointment whereas previous research on mental health services has typically examined pre-post self-report outcomes after six appointments (e.g., DeStefano et al., 2001; Lockard et al., 2012; Nafziger et al., 1999) and retrospective reports at the end of the semester (e.g., Winterrowd, et al., 2016; Reynolds & Chris, 2008; Turner & Berry, 2000). Furthermore, we counted both individual and group counseling sessions toward our shorter timeframe for assessment (five appointments) whereas many studies count only individual counseling sessions. For example, Lockard and colleagues (2012) found positive effects of counseling services on the CCAPS Academic Distress scale after six individual counseling sessions. Thus our timeframe for assessment may have limited our ability to detect the academic benefits of counseling.

In terms of institutional measures, we examined changes in semester GPA from the semester prior to counseling to the semester in which counseling services were received as well as the relationship between number of appointments and semester GPA. Other studies have considered cumulative GPA (Lee et al., 2009) or semester GPA from semesters after counseling was received (Illovsy, 1997). In addition, the current study did not examine retention, a variable of interest among many in student affairs. Although retention is argued to be an inappropriate outcome for counseling services (e.g., Heitzmann & Nafziger, 2001; Lockard et al., 2012), it may be more appropriate for other student services and its relationship to diverse measures of academic outcomes should be explored in future research.

Conclusion

Assessment of academic outcomes continues to be of paramount importance in student affairs—as best practice and as a means of demonstrating each unit’s value in supporting the academic mission of the university. However, it can be difficult to determine which academic outcome measures to use to best capture the impact of student services. By considering the characteristics of a given academic outcomes measure, including the subsample of clients who will have available data, student affairs practitioners can select the appropriate measures for the particular population they are trying to serve and evaluate the specific aspects of academic functioning their services are designed to promote. In the end, it may be best to utilize multiple measures in combination in order to fully examine academic outcomes across students.

We hope that our study highlights some of the considerations practitioners might take into account when selecting outcome measures and that our suggestions will help student affairs units maximize their opportunities to demonstrate their value and further improve their services.

References

- American College Personnel Association. (1994). *The student learning imperative: Implications for student affairs*. Washington, D.C.: American College Personnel Association.
- Astin, A. W., & Antonio, A. I. (2012). *Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education* (2nd ed.). New York, NY: Rowman & Littlefield/American Council on Education.
- Berger, J. B., & Malaney, G. D. (2003). Assessing the transition of transfer students from community colleges to a university. *NASPA Journal*, 40(4), 1–23.
- Bowman, N. A. (2013). Understanding and addressing the challenges of assessing college student growth in student affairs. *Research & Practice in Assessment*, 8, 5–14.
- Center for Collegiate Mental Health (2012). *CCAPS 2012 User Manual*. University Park, PA.
- Cholewa, B., & Ramaswami, S. (2015). The effects of counseling on the retention and academic performance of underprepared freshmen. *Journal of College Student Retention*, 17, 204–225. doi: 10.1177/1521025115578233
- DeStefano, T. J., Mellot, R. N., & Petersen, J. D. (2001). A preliminary assessment of the impact of counseling on student adjustment to college. *Journal of College Counseling*, 4, 113–121. doi: 10.1002/j.2161-1882.2001.tb00191.x
- Erwin, T. D., & Sivo, S. A. (2001). Assessing student learning and development in student affairs: A nuts and bolts introduction. In R. B. Winson, Jr., D. G. Creamer, & T. K. Miller (Eds.), *The professional student affairs administrator: Educator, leader, and manager*. (pp. 357–375). New York: Brunner-Routledge.
- Gallagher, R. P. (2011). *National Survey of Counseling Center Directors 2011* (Monograph No. 8T). The International Association of Counseling Services, Inc.
- Heitzmann, D., & Nafziger, K. L. (2001). Assessing counseling services. In J. H. Schuh & M. L. Upcraft and Associates (Eds.), *Assessment practice in student affairs: An applications manual*. (pp. 390–412). San Francisco: Jossey-Bass.
- Illovsky, M. E. (1997). Effects of counseling on grades and retention. *Journal of College Student Psychotherapy*, 12, 29–44. doi:10.1300/J035v12n01_04
- Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). What matters to student success: A review of the literature. Commissioned Report for the National Symposium on Postsecondary Student Success: Spearheading a Dialog on Student Success. Washington, DC: National Postsecondary Education Cooperative. Retrieved from: [https://www.ue.ucsc.edu/sites/default/files/WhatMattersStudentSuccess\(Kuh, July2006\).pdf](https://www.ue.ucsc.edu/sites/default/files/WhatMattersStudentSuccess(Kuh, July2006).pdf)
- Lambert, M. J., & Hawkins, E. J. (2004). Measuring outcome in professional practice: Considerations in selecting and using brief outcome instruments. *Professional Psychology: Research and Practice*, 35, 492–499. doi: 10.1037/0735-7028.35.5.492
- Lee, D., Olson, E. A., Locke, B., Michelson, S. T., & Odes, E. (2009). The effects of college counseling services on academic performance and retention. *Journal of College Student Development*, 50, 305–319. doi: 10.1353/csd.0.0071
- Lockard, A. J., Hayes, J. A., McAleavey, A. A., & Locke, B. D. (2012). Change in academic distress: Examining differences between a clinical and nonclinical sample of college students. *Journal of College Counseling*, 15, 233–246. doi: 10.1002/j.2161-1882.2012.00018.x
- 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. doi: 10.1037/a0021282
- Locke, B. D., McAleavey, A. A., Zhao, Y., Lei, P.-W., Hayes, J. A., Castonguay, L. G., Li, H., Tate, R., & Lin, Y.-C. (2012). Development and initial validation of the counseling center assessment of psychological symptoms-34. *Measurement and Evaluation in Counseling and Development*, 45, 151–169. doi:10.1177/07481756114326242
- Miller, A. & Markman, K. D. (2007). Depression, regulatory focus, and motivation. *Personality and Individual Differences*, 43, 427–436. doi: 10.1016/j.paid.2006.12.006

- Minami, T., Davies, D. R., Tierney, S. C., Bettmann, J. E., McAward, S. M., Averill, L. A., Huebner, L. A., Weitzman, L. M., Benbrook, A. R., Serlin, R. C., & Wampold, B. E. (2009). Preliminary evidence on the effectiveness of psychological treatments delivered at a university counseling center. *Journal of Counseling Psychology*, 56, 309–320. doi: 10.1037/a0015398
- Nafziger, 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.
- Reynolds, A. L., & Chris, S. (2008). Improving practice through outcomes based planning and assessment: A counseling center case study. *Journal of College Student Development*, 49, 374–387. doi: 0.1353/esd.0.0017
- Schuh, J. H. (2011). *Assessment methods for student affairs*. San Francisco, CA: Jossey-Bass, Inc.
- Sexton, T. L. (1996). The relevance of counseling outcome research: Current trends and practical implications. *Journal of Counseling & Development*, 74, 590–600. doi: 10.1002/j.1556-6676.1996.tb02298.x
- Spitzer, T. M. (2000). Predictors of college success: A comparison of traditional and non-traditional age students. *NASPA Journal*, 38, 82–98. doi: 10.2202/1949-6605.1130
- Stallman, H. M. (2010). Psychological distress in university students: A comparison with general population data. *Australian Psychologist*, 45, 249–257. doi: 10.1080/00050067.2010.482109
- Suskie, L. (2009). *Assessing student learning: A common sense guide* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Turner, A. L., & Berry, T. R. (2000). Counseling center contributions to student retention and graduation: A longitudinal assessment. *Journal of College Student Development*, 41, 627–636.
- Upcraft, M. L., & Schuh, J. H. (1996). *Assessment in student affairs: A guide for practitioners*. San Francisco: Jossey-Bass, Inc.
- Varlotta, L. E. (2012). Toward a more data-drive supervision of collegiate counseling centers. *Journal of American College Health*, 60, 336–339. doi: 10.1080/07448481.2012.663843
- Winterrowd, E., Priniski, S. J., Achter, J., & Abhold, J. (2016). Correlates of satisfaction, intrapersonal learning, and academic outcomes at counseling centers in a university system. *College Student Journal*, 50, 288–301.
- Wlazelek, B. G., & Coulter, L. P. (1999). The role of counseling services for students in academic jeopardy: A preliminary study. *Journal of College Counseling*, 2, 33–41.