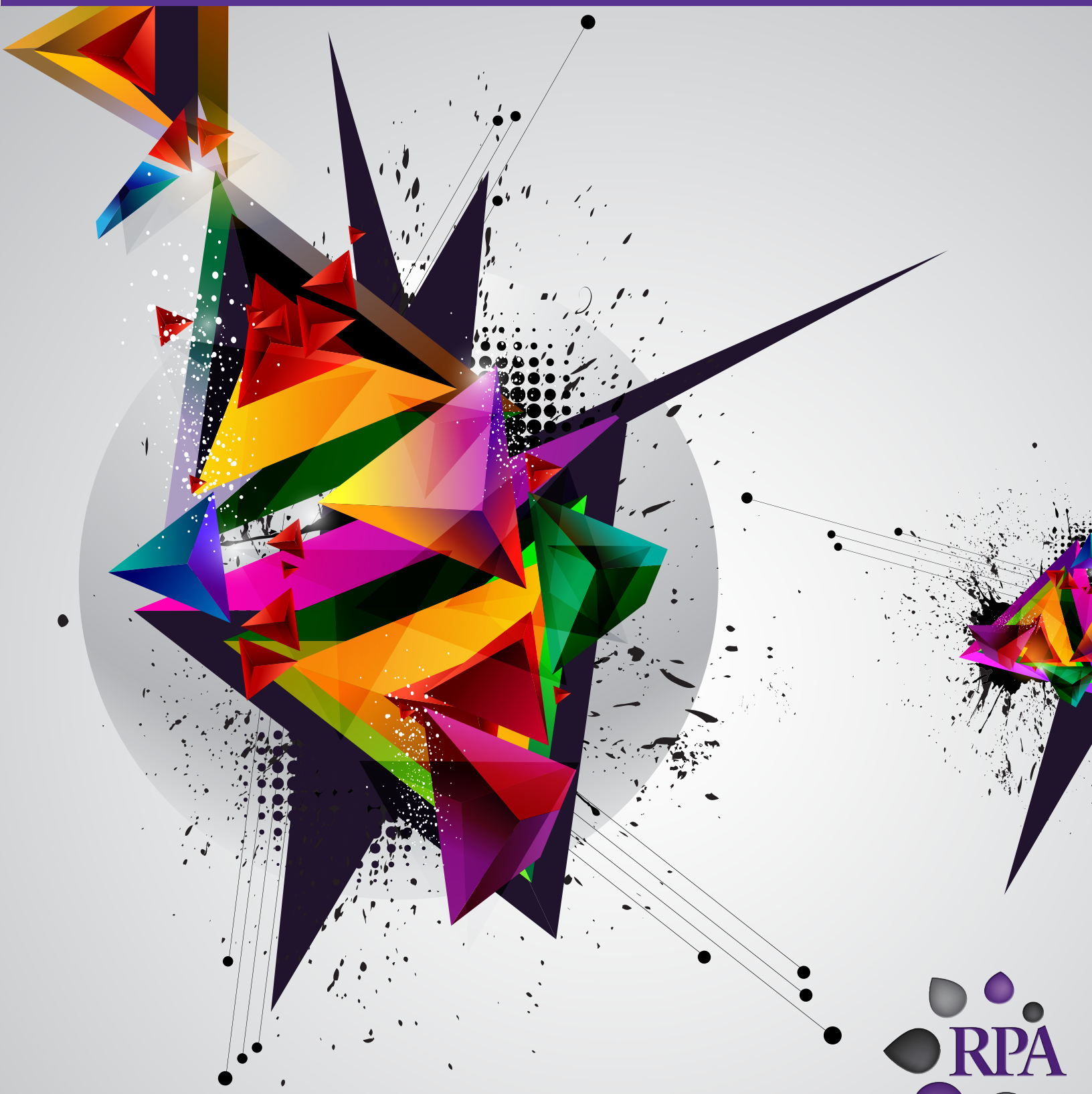
RESEARCH & PRACTICE IN ASSESSMENT

VOLUME FOURTEEN | SUMMER 2019
www.RPAjournal.com
ISSN# 2161-4210



A PUBLICATION OF THE VIRGINIA ASSESSMENT GROUP

RPA RESEARCH & PRACTICE IN ASSESSMENT

CALL FOR PAPERS

Research & Practice in Assessment is currently soliciting articles and reviews for its Winter 2019 issue. Manuscripts submitted to RPA may be related to various higher education assessment themes, and should adopt either an assessment measurement or an assessment policy/foundations framework. Contributions are accepted at any time, but submissions received by October 1 will receive consideration for the winter issue. Manuscripts must comply with the RPA Submission Guidelines and be submitted to our online manuscript submission system found at www.rpajournal.com.

RESEARCH & PRACTICE IN ASSESSMENT

The goal of *Research & Practice in Assessment* is to serve the assessment community as an online journal focusing on higher education assessment. It is dedicated to the advancement of scholarly discussion amongst researchers and practitioners in this evolving field. The journal originated from the Board of the Virginia Assessment Group, one of the oldest continuing professional higher education assessment organizations in the United States. *Research & Practice in Assessment* is a peer-reviewed publication that uses a double-blind review process. Approximately forty percent of submissions are accepted for issues that are published twice annually. *Research & Practice in Assessment* is listed in Cabell's Directory and indexed by EBSCO, ERIC, Gale, and ProQuest.

Published by:
VIRGINIA ASSESSMENT GROUP | www.virginiaassessment.org

Publication Design by Patrice Brown | Copyright © 2019

TABLE OF CONTENTS

FROM THE EDITOR

Growing Together

"*Growth is never by mere chance; it is the result of forces working together*." - James Cash Penney

"I am pleased to introduce myself as the new editor-in-chief of *Research & Practice in Assessment* under such an important theme. During my tenure, it is my hope that we can find new, exciting ways of working together to continue the growth of both the journal and the field. Over the next year, we will be considering a number of changes to the journal designed to facilitate such growth. We are eager to share these ideas with you as we work to make RPA an even better facilitator of the scholarship of assessment. We look forward to your input and feedback as we seek to push the boundaries of our work together.

The Summer 2019 issue of *Research & Practice in Assessment* includes six peer-reviewed articles that highlight the benefits of working together not only with other assessment colleagues, but also with ideas from other disciplines. Pope, Finney, and Bare begin with our learning improvement spotlight article, highlighting the benefits accrued when program theory is included in student affairs assessment. Hart and Robinson then guide us through the process of using charrettes, originally used by architects, to bring together students, faculty, and external stakeholders in assessment. Next, Leaderman and Polychronopolous lean on their background in counselling to propose a model of faculty-assessment consultation and collaboration.

In another application of cross-disciplinary work, Demeter, Robinson, and Frederick apply the sociological framework of triangulation to address the assessment of critical thinking and written communication. Groover, McBrayer, Cleveland, and Riggs then examine both administrative and student affairs units' perceptions of usefulness for different varieties of assessment resources. Finally, Smith and Gordon delve into the intersection of faculty affairs and assessment to make recommendations to support faculty engagement in our work.

I hope this issue of *Research & Practice in Assessment* informs and inspires your efforts to work and grow together.

Regards,

Nicholas Curtis

Marquette University

Abstract

Despite persistent calls by professional organizations and leaders in the field for theory-based programs, it is often difficult for student affairs professionals to articulate why and how their programs should work (i.e., program theory). This lack of program theory influences professionals' ability to use assessment results for program improvement. We, therefore, address two barriers to the articulation of program theory: knowledge of relevant theory and the ability to apply theory to practice. For the latter, we provide a four-step process to assist professionals in developing theory-based programs and assessing their effectiveness. To increase efficiency in assessment practice, we recommend program theory be well-articulated before outcomes assessment data are collected. Importantly, the articulation of program theory should facilitate the realization of the ultimate goal of outcomes assessment: learning improvement.

LEARNING **IMPROVEMENT**

AUTHORS

Andrea M. Pope, M.Ed.
James Madison University

Sara J. Finney, Ph.D.
James Madison University

Aaren K. Bare, M.Ed.
James Madison University

The Essential Role of Program Theory: Fostering Theory-Driven Practice and High-Quality Outcomes Assessment in Student Affairs

"Those student affairs professionals who understand the nature of their profession (e.g., the theories that underlie their work) were able to more effectively engage in outcomes-based assessment and identify how their programs contribute to student learning and development. Without an understanding of theories, others were having difficulty evaluating their programs, even though they had a general understanding of how to implement outcomes-based assessment" (Bresciani, 2010, p.86)

There are many approaches to program development. One manner of characterizing the different approaches is the extent to which programming is theory driven. At one end of the continuum are programs intentionally designed, using theory and research, to address certain problems or achieve particular student learning outcomes (e.g., increase civic engagement, improve grade point average, develop leadership skills). At the other end are programs designed more haphazardly, with little explanation as to why they should "work" or if they are necessary. Most student affairs programs likely fall somewhere in the middle. For these programs, the need for programming may be evident, but why programming should achieve desired student outcomes is unclear. In other words, these programs lack program theory.

Program theory is defined as "the construction of a plausible and sensible model of how a program is supposed to work" (Bickman, 1987, p. 5). Furthermore, it "clarifies the set of cause-and-effect relationships" believed to connect the things students do (i.e., programming) to the outcomes they are expected to achieve (Bickman, 1987, p. 5). Consider

CORRESPONDENCE

Email
popeam@jmu.edu

The need for programming may be evident, but why programming should achieve desired student outcomes is unclear.

a purposefully simple and didactic example: two programs designed to reduce binge drinking on campus. Program A requires all first-year students to read the campus alcohol policy and sign an agreement stating they will abstain from any illegal or irresponsible drinking behavior. This program is depicted as a logic model in Figure 1. The arrows in this logic model represent hypothesized causal relationships. Thus, it appears that developers of Program A believe if students read and sign the alcohol policy, then they will drink less. There is no articulation, however, of why this would be the case. Now consider Program B, a seminar where students are informed about the risks of binge drinking via a lecture. Unlike Program A, the logic model for Program B makes it clear how the program is expected to work: developers believe the lecture will increase knowledge of risks, thereby reducing alcohol consumption.

At this point, it is necessary to distinguish between weak and strong program theory. Weak program theory is often based on hunches, assumptions, or limited personal experiences. Strong program theory, on the other hand, is theory- and/or evidence-based and provides a coherent, theory-based link between program activities and student learning outcomes. For example, imagine if the developers of Program B were asked, "Why should this program result in the intended outcome?" They could state, "We believe the lecture should increase students' knowledge of alcohol-related risks, and their increased knowledge of risks will reduce alcohol consumption." This statement would reflect their program theory. However, without established theory or empirical evidence to support the link between knowledge of risks and alcohol consumption, the program theory would be weak. In practice, we (unfortunately) observe weak program theory daily, which (aligning with the opening quote from Bresciani) prohibits the use of assessment results to improve ineffective programs. In fact, we have witnessed professional workshops that encouraged rapid program development based on hunches/beliefs even when established theory and empirical evidence existed that could guide program development and, in some cases, would be evidence against those hunches/beliefs.

Now assume Program C is based on literature linking alcohol consumption to students' perceptions of how much their peers drink (Prentice, 2008; U.S. Department of Health and Human Services, 2002). The developers could easily explain "why" the program should result in the intended outcomes and, importantly, they could cite theory and research supporting these statements. We say 'importantly', because what student, client, or patient would choose a program that was based on hunches, assumptions, or beliefs when they could engage in a program intentionally designed using research and theory? From this point forward, our use of the term *program theory* refers to strong program theory only. In doing so, we emphasize that program theory is more than a logic model or flow chart that simply lays out program processes without explaining why they work. Program theory involves undergirding each arrow in the model with theory that supports the links (Baldwin, Hutchinson, & Magnuson, 2004).

With program theory defined, we now call back to the decades-old discussion of the importance of theory-based programming found in student affairs literature and professional standards. We then provide a step-by-step process for creating theory-based programs. We close by noting a major barrier to theory-based programming—knowledge of relevant theory—and call on Higher Education Student Affairs (HESA) graduate programs to acknowledge and address this gap. Notably, we intentionally situate outcomes assessment throughout these sections, as it is via assessment that theory-based programs can be evaluated and improved.

Historical and Current Calls for Program Theory

Our call for better articulation of program theory is not novel. We find great comfort in knowing a number of seminal student affairs documents, as well as the most recent professional standards in the field, have articulated the importance of theory and research when creating and assessing program effectiveness. Unfortunately, in our experience, many professionals are unaware of the existence of these documents and standards, much less their actual content. Hence, it is worth revisiting the historical and current emphasis on theory and research when creating and assessing programming in student affairs.

Seminal Historical Documents

In 1949, with the publication of *The Student Personnel Point of View*, the interplay of research and practice was highlighted as "a dominant characteristic of modern [student] personnel work" (Williamson, 1949, p. 12). Nearly fifty years later, the American College Personnel Association (ACPA) codified this sentiment in its seminal document, *The Student Learning Imperative: Implications for Student Affairs*, stating that any student affairs division committed to student learning and personal development should base policies and programs on "promising practices from the research on student learning and institution-specific assessment data" (1994, p. 4). Two years later, ACPA would partner with NASPA and the American Association for Higher Education (AAHE) to release *Powerful Partnerships: A Shared Responsibility for Student Learning*, another publication underscoring the role of theory and research in higher education. More specifically, *Powerful Partnerships* focused on the development of co-curricular programs and shared ten "insights gained through the scholarly study of learning and their implications for pedagogy, curricula, learning environments, and assessment" (AAHE, ACPA, & NASPA, 1998, p.1). In 2004, ACPA and NASPA published *Learning Reconsidered*, which called for professionals to not only be familiar with a wide range of theories and research related to student affairs practice but to develop interventions informed by this research (Keeling, 2004).

Current Professional Standards

Three sets of professional standards also call for theory-based programming (see Table 1). The Council for the Advancement of Standards in Higher Education created the CAS standards to support the development, assessment, and improvement of programming to enhance student learning and development (Council for the Advancement of Standards in Higher Education, 2015). The standards state that a program "must" be intentionally designed using theories of learning and development and that professionals "must" remain current regarding theories that affect their programming.

Two sets of professional standards specifically target the skills of student affairs professionals: *The Assessment Skills and Knowledge (ASK) Content Standards for Student Affairs Practitioners and Scholars* and *The Professional Competency Areas for Student Affairs Educators* (Finney & Horst, 2019). The ASK standards "seek to articulate the areas of content knowledge, skill and dispositions that student affairs professionals need in order to perform as practitioner-scholars to assess the degree to which students are mastering the learning and development outcomes we intend as professionals" (ACPA, 2006, p. 3). *Standard 2: Articulating Learning and Development Outcomes* specifies that professionals must have the ability to employ theory when beginning the assessment process.

A number of seminal student affairs documents, as well as the most recent professional standards in the field, have articulated the importance of theory and research when creating and assessing program effectiveness.

Whereas the ASK Standards focus on professional competency in the domain of assessment, the Professional Competency Areas for Student Affairs Educators are a broader set of ten competencies (American College Personnel Association & National Association of Student Personnel Administrators, 2015). The ability to build theory-based programs is emphasized at multiple points throughout the document. For example, the *Student Learning and Development (SLD) Competency* explicitly calls for professionals to know and use theory to create and improve programs. *The Assessment, Evaluation and Research (AER) Competency* stresses the need for program theory to engage in high-quality outcomes assessment.

The emphasis on theory within the seminal documents is expected given program theory's importance for the development of effective programs. Building programs supported by theory and evidence results in greater confidence that programs should affect students in desired ways. Subsequent outcomes assessment is needed, of course, to formally evaluate program effectiveness. However, program theory allows professionals to intentionally build programs that theoretically should "work" and then use assessment in a confirmatory way to test this hypothesis. This approach can be contrasted with the (often-encountered) conceptualization of program development as a rushed, unsystematic process of cobbling together materials. The theory-based approach is ultimately more efficient with regard to time and resources because the programs generated are more likely to be effective in improving

Table 1
Professional standards and competencies related to knowledge and use of theories

CAS Standards		ACPA-NASPA Professional Competencies		ASK Professional Standards
Program	HESA Graduate Training	Student Learning & Development (SLD) Competency	Assessment, Evaluation & Research (AER) Competency	Standard 2: Articulating Learning & Development Outcomes
• "Programs and services must be guided by theories and knowledge of learning and development." • "Personnel must engage in continuing professional development activities to keep abreast of the research, theories, legislation, policies, and developments that affect their programs and services."	• "The curriculum must include studies of student development theories and research relevant to student learning and personal development." • "Graduates must be able to demonstrate knowledge of how student learning and learning opportunities are influenced by student characteristics and by collegiate environments so that graduates can design and evaluate learning experiences for students."	Foundational Level: • "Articulate theories and models that describe the development of college students and the conditions and practices that facilitate holistic development." • "Identify one's own informal theories of student development ('theories in use') and how they can be informed by formal theories to enhance work with students." • "Assess learning outcomes from programs and services and use theory to improve practice." Intermediate Level: • "Design programs and services to promote student learning and development that are based on current research on student learning and development theories." • "Utilize theory-to-practice models to inform individual or unit practice." • "Justify using learning theory to create learning opportunities." Advanced Level: • "Utilize theory to inform divisional and institutional policy and practice." • "Translate theory to diverse audiences (e.g., colleagues, faculty, students, parents, policy-makers) and use it effectively to enhance understanding of the work of student affairs." • "Analyze and critique prevailing theory for improved unit, division, or campus practice." • "Identify staff members' level of competency regarding the ability to apply learning and development theory to practice, and create professional development opportunities utilizing various learning concepts."	Foundational Level: • "Design program and learning outcomes that are appropriately clear, specific, and measurable, that are informed by theoretical frameworks and that align with organizational outcomes, goals, and values." Intermediate Level: • "Utilize formal student learning and development theories as well as scholarly literature to inform the content and design of individual and program level outcomes as well as assessment tools such as rubrics."	• "Ability to articulate intentional student learning and development goals and their related outcomes. In establishing those goals, the ability to use cognitive and psychosocial development theories germane to the student populations (e.g., traditional age, cultural background, adult education, and so on) as well as an awareness that different subpopulations may have different patterns of development (Love and Guthrie, 1999)." • "Ability to identify the appropriate philosophical or research underpinnings (such as positivist, constructivist, critical theory, and so on) for the articulation of outcomes, dependent on the outcomes themselves."

student learning than theory-less programs. Thus, fewer iterations of the assessment cycle are required to inform changes to the program in order to evidence the desired impact.

Building a Theory-Based Program

Building programs supported by theory and evidence results in greater confidence that programs should affect students in desired ways.

We have attempted to make a strong argument for the articulation of program theory. However, to actually build theory-based programs professionals must (a) know the relevant theories for the student learning and development outcomes (SLOs) they seek to impact, and (b) know how to apply these theories to practice. Given the general nature of theories and the specific nature of programming, the latter can be difficult without proper training. Additionally, the few theory-to-practice models that exist are often too vague to be useful (Reason & Kimball, 2012). As noted by Bloland, Stamatakos, and Rogers (1994), "very little of a practical, nuts-and-bolts nature, is presented for translating theory into campus programs" (p.11). To address this gap in the professional development literature, we articulate a four-step process for building theory-based programming: articulate a feasible and malleable distal outcome; articulate theory-based intermediate (proximal) outcomes; create intentional, theory-based programming; and assess program effectiveness (see Table 2). For each step,

Table 2
Four-step process for building and evaluating a theory-based program

General Model	Most Important Question to Ask	Binge Drinking Example
1. Articulate the Distal Outcome	• What is the problem or distal outcome that needs attention?	• The distal outcome of the program is to significantly reduce the frequency of binge drinking on campus.
2. Articulate Theory-Based Intermediate (Proximal) Student Learning Outcomes (SLOs)	• What is the etiology (i.e., what are the causes) of the distal outcome based on current theory and research?	• Students drink excessively, in large part, due to flawed perceptions of how much their peers drink, coupled with a desire to "fit in" (Prentice, 2008; U.S. Department of Health and Human Services, 2002). • Given this research, the following intermediate SLO was specified: *As a result of participating in the binge drinking program, participants will be able to accurately describe student drinking norms at University X.*
3. Develop Theory-Based Programming to Impact Intermediate SLOs	• What programming affects the intermediate SLOs based on current theory and research?	• Program consists of content and activities that provide students with a realistic perception of their peers' drinking behaviors (e.g., interactive infographic and discussion), as articulated by theory and research in the social normative domain (Prentice, 2008). • Program also consists of content and activities that foster the encoding, integration, and retention of information (e.g., highlighting a small number of important facts, presenting information in multiple formats, engaging students in active discussions), as articulated by theory and research in the domain of cognition and learning (Halpern & Hakel, 2003).
4. Evaluate Outcomes Data to Inform Inferences about Program Effectiveness and Guide Changes in Program for Improvement	• Do assessment results suggest the programming impacts the intermediate SLOs?	• Outcome data were collected for the intermediate SLO to assess students' knowledge of drinking norms (i.e., a multiple choice test was administered before the program, immediately after the program, and 8 weeks post program). • Outcome data were collected for the distal outcome of binge drinking (i.e., number of students who binge drink once or more per week was recorded before the program, immediately after the program, and 8 weeks post program).

Note. Program theory incorporates educational and psychosocial theories that link the proximal intermediate outcomes to the distal outcome (Step 2). Program theory also explicates how program components affect the proximal intermediate outcomes (Step 3).

we highlight several questions to guide professionals through the process. By providing brief examples, we hope to elucidate the process and assist professionals who are committed to building and assessing theory-based programs but may be unsure how to begin.

Step 1: Articulate the Problem or Distal Outcome

Building a theory-based program begins by clearly articulating the problem one is trying to address or, alternatively, the goal one hopes to achieve through programming. Often, this distal outcome (e.g., reduction in binge drinking) will be tied to the mission of the office, department, or institution within which the program is housed. It may also stem from demonstrated student need, staff consensus, or relevant professional standards (e.g., CAS Student Learning and Development Outcomes domains). It is achievement of the distal outcomes that program developers truly care about. However, given the complex nature of these outcomes, they may not be realized due to a single program; hence, the need to specify more proximal intermediate outcomes (see Step 2).

It is achievement of the distal outcomes that program developers truly care about. However… they may not be realized due to a single program.

Two important questions must be asked when articulating distal outcomes. First, "Is it theoretically possible to impact the targeted outcome in a college student population?" In other words, is the outcome a malleable skill or a stable trait? For example, there has been considerable debate about whether creativity can be learned or if it is determined by "inherent neurological and personality traits rather than methodology or practice" (Delistraty, 2014). If creativity is a stable trait, attempting to develop programming to increase creativity would be a waste of resources. Thus, the malleability of the targeted construct must be researched before distal outcomes are specified.

Second, "Is it feasible to impact the distal outcome given time, resources, and other practical constraints often present within college settings?" For example, research may suggest it is theoretically possible to increase empathy in college students. However, if the research also suggests changes in empathy would require expensive programming that spans several years, it may be practically infeasible for an institution to target this outcome.

Relevant Questions:

- What is the desired distal outcome of the program? What problem are you trying to solve?
- Why is the distal outcome important? Is it aligned with department/institution priorities?
- Based on theory and research, is the distal outcome malleable and, if so, is it practically feasible to influence this outcome at your institution?

Step 2: Specify Theory-Based Intermediate Outcomes

Attempting to specify intermediate SLOs without a thorough understanding of the distal outcome (informed by theory and research) is likely to result in SLOs that are misguided.

Once the distal outcome (i.e., the problem/goal) has been articulated, the next step is to consult relevant theory and empirical research to articulate the underlying causes of the problem. In other words, one must understand the etiology of the distal outcome before proceeding (West & Aiken, 1997). With this knowledge, practitioners can specify *intermediate* student learning outcomes. Programming can then be developed (in Step 3) to influence these more *proximal* SLOs (Timm, Davis Barham, McKinney, & Knerr, 2013).

Attempting to specify intermediate SLOs without a thorough understanding of the distal outcome (informed by theory and research) is likely to result in SLOs that are misguided. Specifically, there may be a disconnect between achievement of the more proximal, intermediate outcomes of a program and achievement of the distal outcome. For example, consider again Program B from the binge drinking example in Figure 1. The desired distal outcome is to reduce binge drinking and programmers have specified the following intermediate SLO: *As a result of participating in Program B, students will report increased knowledge of alcohol-related risks*. Thus, the program is hypothesized to work (i.e., reduce alcohol consumption) via the intermediate process of increasing students' knowledge of risks. However, if the consensus among researchers is that knowledge of alcohol-related risks actually has no impact on drinking behavior (U.S. Department of Health and Human Services, 2002), then even if Program B does an excellent job of teaching the risks, the desired distal outcome is not likely to be achieved. Had developers researched the causes of binge drinking they would have discovered that college students binge drink in large part due to flawed perceptions of how much their peers drink, coupled with a desire to fit in (Prentice, 2008; U.S. Department of Health and Human Services, 2002). Thus, more appropriate SLOs might focus on accurate perceptions or managing the desire to fit in. In sum, theory and research helps determine what specific knowledge, skills, behaviors, and/or attitudes (i.e., intermediate SLOs) should be cultivated through programming to achieve desired distal outcomes.

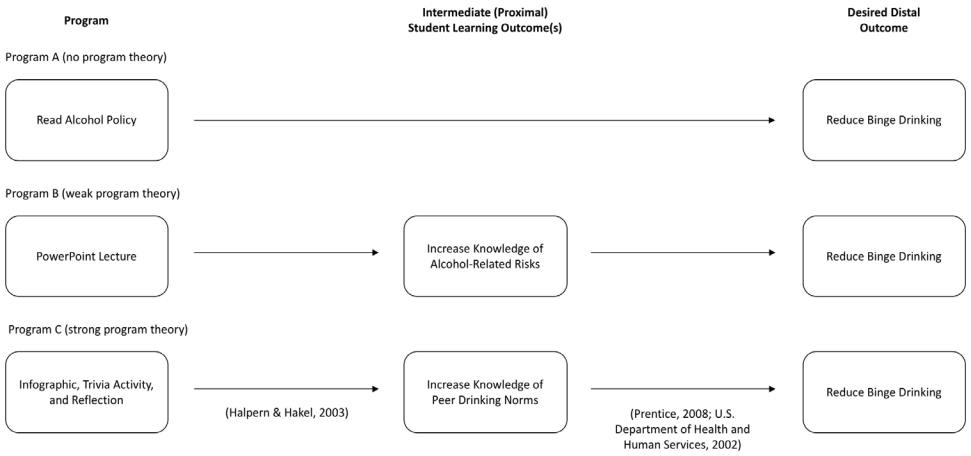


Figure 1. Three logic models depicting the difference between a program with no program theory (links between programming and desired distal outcome not specified), a program with weak program theory (links between programming and desired distal outcome specified, but not theory-based), and a program with strong program theory (links between programming and desired distal outcome specified and theory-based).

Relevant Questions:

- What is the etiology of the distal outcome? What are the underlying causes of the problem the program is designed to address?
- What knowledge, skills, attitudes, or behaviors influence the desired distal outcome? What are the intermediate SLOs?
- For each arrow linking an intermediate SLO to the distal outcome, what theories or research support the link?

Step 3: Develop Theory-Based Programming

Once the distal outcome and intermediate SLOs are specified the next step is to determine how to achieve the intermediate SLOs through programming. Programming encompasses content (e.g., specific activities) and delivery (e.g., pedagogical techniques).

Just as theory was used to articulate intermediate SLOs and link them to the distal outcome, theory should also inform the programming and the link between programming and intermediate SLOs. Thus, there are two types of theories that underlie the two sets of arrows in a logic model (Baldwin, et al., 2004). One type of theory articulates the etiology of the distal outcome and justifies the link between the distal outcome and intermediate SLOs (Step 2). The other type of theory justifies the link between programming and intermediate SLOs (Step 3). The articulation of both types of theories results in strong program theory (i.e., an explicit, theory-based explanation of how programming affects intermediate outcomes and how intermediate outcomes affect the distal outcome) as illustrated in Figure 2.

A good starting point for building theory-based program components/activities is to look for intervention studies or empirical research that evaluates the effectiveness of theory-based programs. It can also be helpful to consult research review articles on how students learn (e.g., Halpern & Hakel, 2003) and/or how attitudes and behaviors are changed (e.g., Funnell & Rogers, 2011; Yeager & Walton, 2011) when building program content, structure, and delivery. Additionally, there is extensive literature on evidence-based pedagogical techniques (e.g., Fink, 2013).

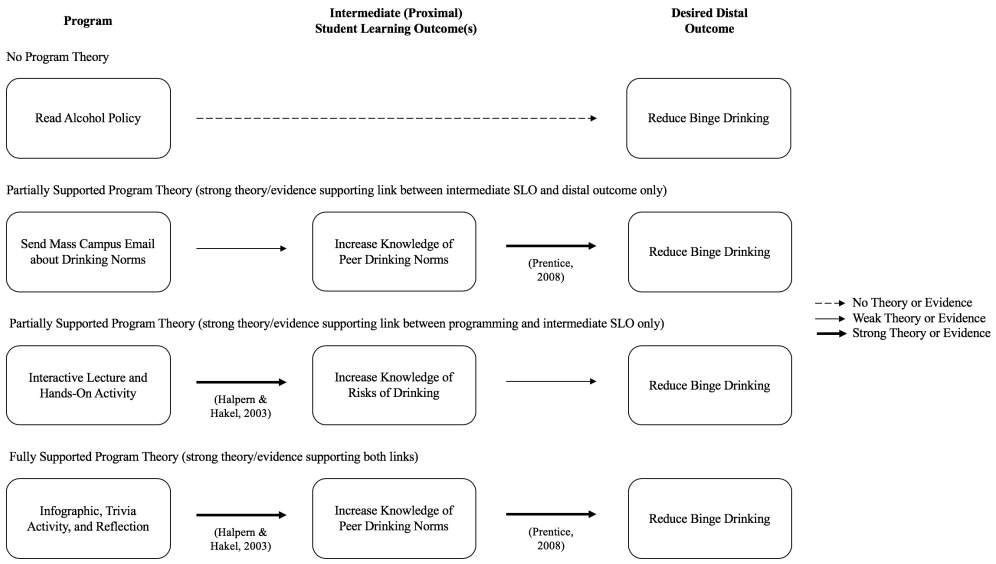


Figure 2. A visual representation of the two types of theories/research needed to articulate strong program theory. The logic models illustrate the difference between a program with no program theory, a program with weak theory/evidence supporting the link between programming and intermediate SLOs, a program with weak theory/evidence supporting the link between intermediate SLOs and the distal outcome, and a program with strong theory/evidence supporting both links.

Although program theory allows one to state that a program should work, success is not guaranteed.

To illustrate this step, consider Program C in Figure 1. This program has specified the following theory-based intermediate SLO: *As a result of participating in Program C, participants will be able to accurately describe student drinking norms at University X*. Given the knowledge focus of this SLO (students must know campus drinking norms to describe them), program developers consult research on cognition and learning to build a program that facilitates long-term retention of information. From this research, program developers focus on three learning principles: active learning is more effective for long-term retention than passive learning; the more information students are given in a short period of time, the less information they are likely to remember; and presenting information in multiple formats (e.g., visually and verbally) can improve retention (Halpern & Hakel, 2013). Using these theory-based principles as a guide, the developers decide that each student will be given an interactive, electronic infographic containing three statistics and corresponding narratives about their peers' normative alcohol consumption behaviors. Then, for each statistic, a trivia-style question will be posed and the correct answer will be revealed after the student responds. Afterwards, a facilitator will lead an interactive lecture about drinking norms at University X. During this activity, students will recall and explain the information to check their knowledge and receive immediate feedback, resulting in more accurate and stronger encoding of the information. Finally, the facilitator will ask students for their reactions to the information and lead a discussion about any discrepancies between their perceptions and reality.

Relevant Questions:

- How will your program achieve the intermediate SLOs? What specific strategies, activities, and approaches will be employed?
- What theory or evidence is there to support the arrows linking these strategies, activities, and approaches to the intermediate SLOs?
- Is the theory-based programming feasible from a resource perspective?

Step 4:
Assess Program Effectiveness and Use Results to Improve Programming

Although program theory allows one to state that a program should work, success is not guaranteed. Once a theory-based program has been built and implemented the final step is to assess its effectiveness. Knowledge of theory facilitates the collection of outcomes data, interpretation of assessment results, and use of results for improvement (Bresciani, 2010).

To understand why this may be the case, imagine that you implement a program, assess it, and find the SLOs were not achieved. How should you use this information? Should you scrap the entire program or modify parts of it? If you decide to modify the program, what parts should you change? Alternatively, what if the program was effective and student learning did improve? Could you determine which parts of the program were crucial to its success? Under what conditions could you replicate the program and still be relatively confident in its efficacy? Without program theory, it would be difficult to answer these questions. However, by clearly articulating how the different components of the program should (theoretically) result in achievement of intermediate SLOs and progress toward the distal outcome you would be able to collect the more nuanced outcomes data needed to make correct inferences about program effectiveness. This is echoed by Rogers (2000), who notes:

We have observed a serious barrier when professionals work through the steps of articulating program theory: lack of knowledge of relevant theories.

> "If a program achieves its intended outcomes, program theory can help to identify the elements of a program which are understood to be essential for its widespread replication and can then analyze whether these elements are plausibly and empirically associated with success. It should also be able to identify whether program success has been achieved despite (or perhaps because of) failure to implement the program as designed. If a program does not achieve its intended outcomes, a program theory evaluation may be able to identify whether this is due to implementation failure (the program wasn't implemented as intended, which might, in itself, explain the lack of outcomes); unsuitable context (the program was implemented in a context in which the necessary mechanisms did not operate); or theory failure (the program was implemented as intended, in a suitable context and evaluated with a powerful design and measures which would probably have detected important effects if they had been present)." (pp. 210-211)

To illustrate how program theory can be used to generate and test several hypotheses about why a program did not work, consider the following examples. Imagine an alcohol intervention program implemented at a women's college was based on research showing that students' binge drinking (i.e., distal outcome) was due to social pressure (i.e., intermediate outcome). Upon further investigation, however, it is discovered that this research was conducted on male students only, and that for female students, social pressure is not a major cause of alcohol consumption. In this case of *inappropriate theory selection*, even if the intermediate SLO of resisting peer pressure was achieved, the distal outcome of reducing binge drinking would not be achieved. This result is understandable because the link between the intermediate SLO and the distal outcome was not theoretically/empirically supported in the target population. Thus, the assessment results would indicate the need to change the programming to align with a more appropriate intermediate SLO based on a more applicable theory. More important, if the developers only collected data on the distal outcome (i.e., rate of binge drinking) it would have been impossible to know why the programming was ineffective. Instead, the clearly articulated program theory in this example led the program developers to collect data on the intermediate SLO as well. This additional data was then used to identify the broken link between the intermediate SLO and distal outcome.

Alternatively, it could be that the intermediate SLOs were appropriate for the population but the theory-based programming used to achieve them was insufficient or ineffective due to *inappropriate theory application*. That is, the way in which the underlying theories/research were translated into actual program features may have been faulty. For example, imagine assessment results showed that students in Program C (from Figure 1) did not meet the intermediate SLO of increased knowledge of campus drinking norms. Although

research might support the use of active learning strategies to increase long-term knowledge retention (see Step 3), it may be that the specific activities described in Step 3 did not actively engage students. For instance, it could be that the facilitator-to-student ratio was too small, enabling many students to complete the program without actively participating or engaging with the material. This result would indicate the program needs to be tweaked rather than completely overhauled.

Finally, it could be that one or more of the theories on which the program is based is simply wrong. In other words, the assumptions made (causal relations specified in the logic model) by the selected theories are untenable. This is most likely to occur with new theories that have not been thoroughly tested. In this case of *theory failure*, the results would indicate a need to redesign the entire program based on more strongly supported theory and research.

Relevant Questions:

- Did student learning/development improve as a function of the program? Were the intermediate SLOs and distal outcome achieved?
- If the intermediate SLOs and/or distal outcome was not met, where did the breakdown occur? Was the theory underlying the program wrong or inappropriate? Alternatively, was the theory implemented poorly?

Need for Additional Training in Relevant Theories and their Application

We provided the four-step process above to support the creation and assessment of theory-based programs. With that said, we have observed a serious barrier when professionals work through the steps of articulating program theory: lack of knowledge of relevant theories. Recall, theory is necessary to specify the link between intermediate SLOs and distal outcomes. Additionally, theory is needed to specify the link between program components and intermediate SLOs. Although professionals agree that application of student learning and development theories is essential to intentionally plan, assess, and improve programs (Bresciani, 2010; Blimling &Whitt, 1999; Kuh, Kinzie, Schuh, & Whitt, 2005; Love & Estanek, 2004; Mentkowski, 2000; Pascarella, 2006), we have observed a tendency in the field to overlook other relevant bodies of literature.

If professionals embrace the responsibility to keep abreast of relevant research, theory-based programs will be less difficult to create, assess, and improve.

Consider the textbooks frequently used to train student affairs professionals on theory (e.g., Evans, Forney, Guido, Patton, & Renn, 2010; Schuh, Jones, & Torres, 2017). These texts tend to focus on a small cannon of broad student development theories considered foundational to student affairs practice (e.g., Chickering's Theory of Identity Development, Baxter-Magolda's Model of Epistemological Reflection, Perry's Theory of Intellectual and Ethical Development). Although these student development theories can be helpful for describing where students are (and where they should be) developmentally, they are much less useful for prescribing how to create programming to move students from one developmental stage to another. Additionally, these theories describe only a fraction of the outcomes of interest in higher education. Theories related to noncognitive outcomes such as civic engagement, inter-cultural competence, and self-regulation (e.g., self-efficacy, goal orientation, growth mindset) are largely overlooked, even though these constructs are often the focus of programming and rich bodies of research articulate their associated intermediate outcomes and interventions. Likewise, with respect to outcomes related to knowledge acquisition, there is little to no mention of current research on how students learn (i.e., information transfer and long-term retention) or best practices for instruction (however, see Bresciani, 2016).

As a result, burgeoning professionals may believe knowledge of a handful of student development theories is sufficient to guide practice. Evans et al. (2010) speak to this danger:

> "…many student affairs educators have inappropriately elevated student development theory to something resembling icon status. If this has happened or is happening in the student affairs profession, the act deserves to be challenged. No single resource stands alone as the foundation for professional practices. Student development theory, for example, is one of several knowledge bases that can inform student affairs practice." (p.39)

In sum, given student affairs professionals are perceived as educators providing high-impact curricular programming (Keeling, 2004, 2006), they have a responsibility to be familiar with research relevant to students' learning and development (Barber, 2006; Hatfield & Wise, 2015). Likewise, according to the CAS Standards for Master's-Level Student Affairs Preparation Programs (2013), HESA programs have a responsibility to "foster an appreciation of intellectual inquiry in faculty members and students, as evidenced by active involvement in producing and using research, evaluation, and assessment information" (p. 8). If professionals embrace the responsibility to keep abreast of relevant research, theory-based programs will be less difficult to create, assess, and improve.

References

American Association for Higher Education, American College Personnel Association, & National Association of Student Personnel Administrators. (1998). *Powerful partnerships: A shared responsibility for learning*. Washington, DC: Author.

American College Personnel Association. (1994). *The student learning imperative: Implications for student affairs*. Washington DC: Author.

American College Personnel Association (2006). *ASK standards: Assessment skills and knowledge content standards for student affairs practitioners and scholars*. Washington, DC: Author.

American College Personnel Association & National Association of Student Personnel Administrators (2015). *ACPA/NASPA professional competency areas for student affairs practitioners*. Washington, DC: Authors.

Baldwin, C. K., Hutchinson, S. L., & Magnuson, D. R. (2004). Program theory: A framework for theory-driven programming and evaluation. *Therapeutic Recreation Journal, 38*, 16.

Barber, J. P. (2006). Tenet two: Commit to student learning as a primary focus. In R. P. Bingham, D. Bureau, & A. G. Duncan (Eds.), *Leading assessment for student success* (pp. 22-39). Sterling, VA: Stylus Publishing.

Bickman, L. (1987). The functions of program theory. *New Directions for Program Evaluation, 33*, 5–18.

Blimling, G. S., & Whitt, E. J. (1999). *Good practice in student affairs: Principles to foster student learning*. San Francisco, CA: Jossey-Bass, Inc.

Bloland, P., Stamatakos, L., & Rogers, R. (1994). *Reform in student affairs: A critique of student development*. Greensboro, NC: ERIC Counseling and Student Services Clearinghouse.

Bresciani, M. J. (2010). Understanding barriers to student affairs professionals' engagement in outcomes-based assessment of student learning and development. *Journal of Student Affairs, 14*, 81–90.

Bresciani, M. J. (2016). *The neuroscience of learning and development: Enhancing creativity, compassion, critical thinking, and peace in higher education*. Sterling, VA: Stylus Publishing.

Council for the Advancement of Standards in Higher Education (2015). *CAS professional standards for higher education* (9th ed.). Washington, DC: Author.

Council for the Advancement of Standards in Higher Education (2013). *CAS self-assessment guide for master's level student affairs preparation programs*. Washington, DC: Author.

Delistraty, C. C. (2014, July 6). Can creativity be learned? *The Atlantic*. Retrieved from https://www.theatlantic.com/health/archive/2014/07/can-creativity-be-learned/372605/.

Evans, N. J., Forney, D. S., Guido, F. M., Patton, L. D., & Renn, K. A. (2009). *Student development in college: Theory, research, and practice*. San Francisco, CA: John Wiley & Sons.

Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses*. San Francisco, CA: John Wiley & Sons.

Finney, S. J. & Horst, S. J. (2019). Standards, standards, standards: Mapping professional standards for outcomes assessment to assessment practice. *Journal of Student Affairs Research and Practice, 56*, 310-325.

Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: Effective use of theories of change and logic models*. San Francisco, CA: John Wiley & Sons.

Halpern, D. F., & Hakel, M. D. (2003). Applying the science of learning to the university and beyond: Teaching for long-term retention and transfer. *Change: The Magazine of Higher Learning, 35*, 36–41.

Hatfield, L. J., & Wise, V. L. (2015). *A guide to becoming a scholarly practitioner in student affairs*. Sterling, VA: Stylus Publishing.

Keeling, R. P. (2004). *Learning reconsidered: A campus-wide focus on the student experience*. Washington, DC: National Association of Student Personnel Administrators.

Keeling, R. P. (2006). *Learning reconsidered 2: Implementing a campus-wide focus on the student experience*. Washington, DC: National Association of Student Personnel Administrators.

Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2005). *Assessing conditions to enhance educational effectiveness*. San Francisco: Jossey-Bass.

Love, P., & Estanek, S. (2004). *Rethinking student affairs practice*. San Francisco: Jossey-Bass.

Mentkowski, M., Rogers, G., Doherty, A., Loacker, G., Hart, J. R., Rickards, W., ... & Diez, M. (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. San Francisco: Jossey-Bass.

Pascarella, E. T. (2006). How college affects students: Ten directions for future research. *Journal of college student development, 47*, 508–520.

Prentice, D. (2008). Mobilizing and weakening peer influence as mechanisms for changing behavior: Implications for alcohol intervention programs. In M. Prinstein & K. Dodge (Eds.), *Understanding Peer Influence in Children and Adolescents* (pp. 161–180). New York, NY: Guilford.

Reason, R. D., & Kimball, E. W. (2012). A new theory-to-practice model for student affairs: Integrating scholarship, context, and reflection. *Journal of Student Affairs Research and Practice, 49*, 359–376.

Rogers, P. (2000). Program theory: Not whether programs work but how they work. In D. Stufflebeam, G. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on Educational and Human Services Evaluation* (pp. 209–232). Dordrecht, Netherlands: Kluwer Academic Publishers.

Schuh, J. H., Jones, S. R., & Torres, V. (Eds.). (2017). *Student services: A handbook for the profession*. San Francisco, CA: John Wiley & Sons.

Timm, D. M., Davis Barham, J., McKinney, K., & Knerr, A. R. (2013). *Assessment in practice: A companion guide to the ASK Standards*. Retrieved from https://thekeep.eiu.edu/csd_fac/33.

U.S. Department of Health and Human Services. (2002). *How to reduce high-risk college drinking: Use proven strategies, fill research gaps*. Retrieved from https://www.collegedrinkingprevention.gov/media/finalpanel2.pdf.

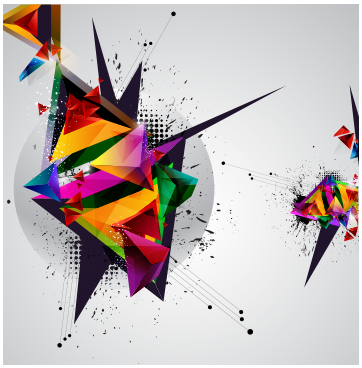
West, S., & Aiken, L. (1997). Toward understanding individual effects in multicomponent prevention programs: Design and analysis strategies. In K. Bryant, M. Windle, & S. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 167–209). Washington, DC: American Psychological Association.

Williamson, E. G. (1949). *The student personnel point of view*. Washington, DC: American Council on Education.

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267–301.

AUTHORS NOTE

Andrea M. Pope, Assessment and Measurement PhD Program, James Madison University; Sara J. Finney, Center for Assessment and Research Studies, James Madison University; Aaren K. Bare, College Student Personnel Administration Master's Program, James Madison University.



Abstract

Today, higher education faces increased accountability to assess and measure student learning. However, traditional metrics (like standardized tests and graduation rates) fail to provide a comprehensive understanding of exactly what students learn in a contextualized program of study. To address this gap, the researchers piloted a charrette model to bring together students, faculty, and external stakeholders. Current and former students were asked to metacognitively reflect on their learning experiences and share them in a storytelling format. Stakeholders then provided critical feedback to faculty on what they perceived students had learned. Data results showed that external stakeholders were able to assess student learning and to provide actionable feedback on next steps, establishing trust and enhanced understanding between all groups. This model may serve as a viable approach to provide much-needed context to stakeholders on student learning outcomes in higher education.

AUTHORS

Laura C. Hart, Ed.D.
UNC Charlotte

Christine Robinson, Ed.D.
UNC Charlotte

"From the Mouths of Babes": Using a Charrette Model to Assess Student Learning and Engage External Stakeholders

Across the country, the drive for educational accountability – well-documented in the P-12 school sector – is rapidly moving into higher education (Ewell, 2009; Kelchen, 2018; Leveille, 2013). Typical methods of assessment have included quantifiable metrics (such as standardized test scores or graduation rates) to document whether or not learning has occurred. While these methods have had questionable success in accurately gauging student achievement in public education (Popham, 1999, 2016; Ravich, 2013), employing these kinds of metrics in higher education becomes even more problematic. Using traditional quantitative measures as policy to assess the wide variety of learning that occurs in higher education can be difficult (Association of American Colleges and Universities, 2016; deBoer, 2016). Accrediting agencies are beginning to implement policies that closely examine proxy measures to assess institutional effectiveness, such as loan default rates and graduation/retention rates (Council of Regional Accrediting Commissions, 2018; Kreighbaum, 2016). State policymakers and higher education systems are adopting these measures as part of their strategic plans and setting targets for higher education institutions to achieve (Carlson, 2017; State Council of Higher Education for Virginia, 2015). Although these data should be reviewed and considered in the context of the individual institution, they do not tell the whole story. External entities create challenges by holding institutions accountable with assessment measures over which institutions have little control, and which may not align with the specific program learning objectives. Although metrics like standardized test scores, default loan rates, retention rates, and time to degree are universally understandable to stakeholders, the story higher education institutions tell of program-specific student learning outcomes (and the methods to assess these explicit outcomes) are unfamiliar to stakeholders in terms of the language used and practices employed. In a multidisciplinary institution with varying levels of expectations and expertise, clearly communicating exactly what students have learned (outcomes) as compared to what they should be learning (objectives) is more complicated than it sounds.

CORRESPONDENCE

Email
lchart2869@gmail.com

This process becomes more challenging when the various contexts of program-specific assessment are added to the mix. Best practice for data collection mandates using triangulated data sources to allow for the clearest interpretation of results; using multiple measures from both formative and summative assessments provides the most salient information about student performance (Black & Wiliam, 2018; Jones, Carr, & Ataya, 2007). Higher education program assessments may also include periodic reviews about relevance, sustainability, and impact. However, all these dimensions use distinctive metrics with content-specific vocabulary and concepts, making it tricky to assess the effectiveness of the institution overall. For example, many higher education programs note the need for college graduates to demonstrate "critical thinking," rightly noted by Suskie (2016) as an umbrella term that tends to encompass a wide lexicon of broader thinking skills. The problem is that these skills are specific to various disciplines—in fact, context and subject-specific knowledge is essential to developing effective critical thinking habits (Bean, 2011; Jones, 2007; Santos, 2016). How a history program opts to measure critical thinking may involve asking students to review primary source documents to analyze societal change. In comparison, measuring critical thinking in mathematics could include requiring students to solve complex statistical problems; measuring critical thinking for teacher education programs might assess how well students implement and adapt lesson plans to the needs of P-12 classroom learners. One size does not fit all.

The story higher education institutions tell of program-specific student learning outcomes (and the methods to assess these explicit outcomes) are unfamiliar to stakeholders in terms of the language used and practices employed.

Even if programs have content-appropriate measures in place to assess essential skills and knowledge, communicating with stakeholders about how these various assessments are parceled and used by faculty can devolve into a futile exercise. Higher education programs struggle to find a clear, cohesive way to demonstrate program effectiveness to stakeholders. As Suskie (2016) pointed out, "American higher education has failed to tell you … what we do and how we do it …. We have to figure out a way to tell our very complex story in short, simple ways that busy people can digest quickly." Institutions must develop better ways of sharing the impact of higher education on student learning with stakeholders.

One possible way to address this need may be to move away from using indirect proxy measures of student learning and institutional quality (like gainful employment and loan default rates, which require little-to-no effort by the students to compile). Instead, programs could engage students in a metacognitive model, one in which students are actively and authentically involved in their own reflective assessment of what they have learned in their program of study. Tom Angelo and Keston Fulcher, both respected experts in the field, indicated that one challenge of higher education assessment is the need to engage current students and graduates about their learning experiences at universities and colleges (personal communication, October 17, 2016). Providing structured opportunities for students to explain or demonstrate their learned subject-specific knowledge and skills to stakeholders would allow for a clearer assessment of learning outcomes. When triangulated with measures already in place, these kinds of "real time," authentic assessments could better educate stakeholders effectively on what students actually know and can do upon program completion (Baer, 2015; Braskamp & Engberg, 2014).

These kinds of interactions also present an opportunity for an additional dimension of program improvement by creating a two-directional relationship with external stakeholders to encourage their feedback. Engaging students in sharing what they have learned directly with stakeholders can provide much-needed clarity to the data and provide a richer understanding of exactly what kind of "education" students are paying for. Asking students to articulate and evaluate their own learning builds critical thinking capacity. In addition, structuring opportunities for external stakeholders to then provide feedback about what they see/hear can build trust between all parties, especially if faculty use the stakeholder feedback as another data source to initiate programmatic change. This would send a clear message that institutions are eager to embrace authentic and viable recommendations. Creating these kinds of direct interactions between external stakeholders, students, and faculty may also combat the notion that institutions are not fulfilling their mission (as assessed on a survey or standardized test). Finally, previous research has established that when assessment is motivated by internal improvement purposes (versus accountability

purposes) the assessment results are actually more likely to be used by faculty (Herman & Hilton, 2017). Giving both students and external stakeholders a voice in the assessment cycle could result in increased buy-in from all parties.

The purpose of this paper is to describe how one university/college designed and piloted an event for students, faculty, and external stakeholders based on a charrette model. *Charrette* is a term whose origins are rooted in architectural design but has evolved to define a process intended to integrate purposeful opportunities for stakeholder feedback on the presented product. Our charrette allowed external stakeholders to interact directly with graduates and current students for the purposes of (a) assessing what students know and are able to do; and (b) providing authentic feedback to program faculty for future improvements. The event was deliberately designed to engage students as not only presenters of knowledge but also as reflective practitioners. Stakeholders invited to the event included employers from the regional community familiar with the demands of the profession and therefore able to provide viable feedback. It is the intention of the researchers to present a model that may be replicated by other institutions for future assessment purposes.

Method: Piloting a Charrette Model

The research was conducted at a large university in the southeastern United States. Student enrollment has steadily increased over the last several years with approximately 29,000 students attending in 2017. The College of Education within the university includes approximately 1,300 students enrolled in teacher preparation programs and 1,000 additional students enrolled in graduate-level master's and doctoral programs. The institution is accredited by the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC); in addition, the college's teacher preparation programs are also accredited by the Council for the Accreditation of Educator Preparation (CAEP, formerly NCATE).

Providing structured opportunities for students to explain or demonstrate their learned subject-specific knowledge and skills to stakeholders would allow for a clearer assessment of learning outcomes.

The idea was to pilot a charrette feedback model with stakeholders—a collaborative approach incorporating multiple two-way communication loops—versus a traditional presentation that typically only features one-way communication (like a performer with an audience). The model begins with a presentation of a product to a group of knowledgeable individuals to receive feedback for improvement purposes. The feedback generated via a charrette is expected to be critical as well as laudatory; once feedback is received the product is redesigned ostensibly incorporating the stakeholder feedback results into the revamped product, making it better. Subsequent feedback loops would continue with updated versions of the product as needed. In the context of program assessment a presentation modeled on a charrette design would, by definition, expose the selected program of study to increased critique—which may push some participants outside their comfort zones. Figure 1 illustrates the charrette feedback framework.

Starting the Work

The university Executive Director (EXD) approached the College of Education assessment director (CAD) about piloting the charrette assessment model for several reasons; first, College of Education (COED) faculty were very familiar with an accountability mindset, from their ongoing work with P-12 school partners. Second, COED faculty were experienced with various assessment models and were open to the concept. Faculty are more likely to engage with program improvement if they perceive the assessment is meaningful and valuable (Emil & Cress, 2014). By asking faculty to be part of the charrette process the EXD and CAD theorized that faculty would be amenable to student and stakeholder feedback. Finally, the COED Dean had already initiated an internal data review with faculty and was supportive in involving external stakeholders as a logical next step. Given that the logistics of this event would require resource allocation, the COED Dean and the Senior Associate Provost worked to provide funds for the event. Anticipated costs included logistical monies for food/event space and stipends for student participants.

The planning committee identified two primary objectives for the project. Objective 1 was *external stakeholders will use candidates' stories/work to assess program effectiveness*. Objective 1 was designed to actively and purposefully incorporate students (both current

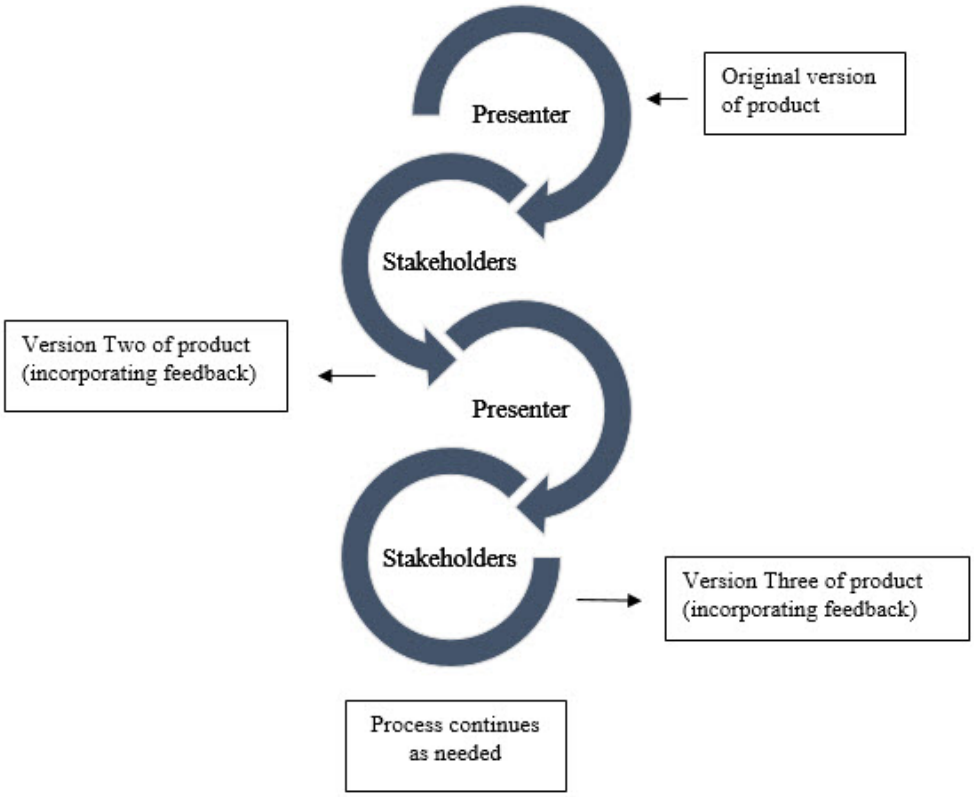


Figure 1. Charrette Feedback Model

students and recently graduated alumni) into the assessment process. This objective addressed the need identified by Angelo and others (personal communication, October 17, 2016) for students to have greater voice in their own assessment. Students were asked to engage in metacognition by reflecting on their own learning and presenting stories about that learning to stakeholders. Stakeholders would then use those stories to assess whether the program had been effective in training teachers for P-12 classrooms. Objective 2 was *external stakeholders will provide actionable feedback to COED faculty for future improvements*. Objective 2 was designed to focus on the next phase of program development. Existing research makes it clear that soliciting stakeholder feedback—especially from employers—can provide a robust data source for university programs (Morgan, 2008; Ulewicz, 2017). Feedback solicited from external stakeholders would identify not only what the COED was doing well but also what the COED needed to improve related to both curriculum and assessment. Both objectives aligned to university assessment and accreditation goals.

Objective 1 was designed to actively and purposefully incorporate students (both current students and recently graduated alumni) into the assessment process.

The committee structured the COED charrette as a one-day event; the morning events would feature students telling their stories about what they had learned during their time in the COED. Afternoon events would provide external stakeholders opportunities to interact with students and with each other, to ask questions, to give critique/feedback, and to assist in developing action strategies for program improvements. The final participant list included 16 current students and eight alumni presenters; 36 external stakeholders representing seven local area school districts and two philanthropic foundations; and 31 College of Education faculty. The external stakeholders were primarily principals and school district leaders who were routinely involved in the clinical placement and hiring of College of Education graduates for employment purposes. They were familiar with the demands of the teaching profession and could speak knowledgeably about the skills teachers need to be successful in their first year.

Engaging Students in Telling their Stories

The morning sessions focused on students telling their stories about what they had learned… In the afternoon the focus shifted to a critical lens.

Two groups of students were asked to participate: current students and recent graduates. The faculty committee debated on whether to share student work samples representing a wider range of student capabilities but finally opted to select students who both produced high-quality work and could be relied upon to complete the work on time. The faculty reasoned that asking less-than-stellar students to participate in this project could create additional logistical issues; in addition, if stakeholder feedback indicated concerns based on the work of our best students, then the committee surmised that faculty would be more likely to take the feedback seriously when considering improvements. Both current students and recent graduates represented a variety of subject areas and grade levels in their respective groups, including birth-kindergarten, elementary, middle grades (6-9), secondary (9-12), and special education. Students were invited to participate and received a small stipend for their work. Designated faculty agreed to serve as mentors for students to answer questions and assist with projects as needed.

Current students. The current students included 16 candidates from a range of points across the program of study. These students created poster presentations designed to showcase a variety of program features from the students' perspectives. A faculty subgroup from the charrette committee worked to develop the presentation framework for consistency but students determined the content of the presentation. The four-part framework for poster presentations provided to current students included the following guiding questions:

A. *What is your story? Think about areas that you feel are strengths for you related to teaching, things that you feel comfortable talking about. What specifically have you learned in your program that has improved your skills and knowledge in one or more of the following:*
    - *"how to teach _____" (i.e., "how to teach fractions", "how to teach reading", "how to teach science inquiry to middle school students," etc.); incorporating research into practice; teaching diverse populations; working with families; building global competency and awareness; written communication*

B. *How / when have you had opportunities to put this knowledge into practice? Describe these experiences in the context of one of the following:*
    1) In courses through class projects/presentations;
    2) In clinicals prior to student teaching;
    3) In student teaching internships.

C. *What data do you have that you are using evidence-based practices in practicing this knowledge/skills? Based on the idea(s) you selected to discuss in Part A above, explain how this is a strength for you. Why is it your strength? How do you know? What evidence to do you have to show that you are becoming more knowledgeablein this area? What opportunities to practice these skills have you had so you know you have improved?*

D. *What is your plan going forward? How do you plan to use the success you have experienced?*

From this framework, a wide array of student presentations were developed. For example, candidates reflected on their clinical experiences in contrast to their university classroom activities; others explored action research projects completed as course assignments; some shared specific things they had learned in their classes, such as why one instructional approach was better than another. Current students received a $200 stipend for their work and time.

Recent graduates. The second student group included eight recent graduates who had completed the program within the last two years and were now employed as classroom teachers. These individuals were involved more deeply in the charrette than current students, as they could speak to both the student and practitioner experience. Instead of poster presentations, the alumni group developed a five minute "TED-like" Talk with the following prompt: *Consider your teacher preparation at [this institution]. What aspects of your program were done well? Reflect on an experience, give personal examples, tell your story. Describe your experience in the context of application in clinicals and now in your classroom. Limit to your talk to 5 minutes*. Faculty mentors provided guidance as needed, but the individual presenters determined the content. The recent graduates also served on an alumni panel to answer unscripted questions about their learning experiences directly from the external stakeholders. Finally, the alumni participated in roundtable stakeholder feedback groups. Alumni received a $500 stipend for their assistance for the entire day.

Structuring the Charrette

Morning sessions: What we do well. The morning sessions focused on students telling their stories about what they had learned. These were primarily positive stories that highlighted good things happening in the COED. The morning session was divided into two rounds of presentations: half (eight) of the poster presentations by current students were shared. During this time, external stakeholders rotated among the presentations, heard the students speak, and asked questions directly to individual students about what they saw/heard. The groups were then asked to turn their collective attention to the stage, where half (four) of the alumni presented their five-minute TED Talks in succession. Stakeholders were provided with Post-it Notes on the tables to record any questions or comments they had about what they heard. These Post-it Notes were collected and incorporated later as questions for the alumni panel in the afternoon. After a brief break, the second round of presentations began, with the rest of the students/alumni presenting their posters/TED Talks. Including the welcome and transitions, the morning session lasted about two-and-a-half hours.

Afternoon sessions: What we need to improve. In the afternoon the focus shifted to a critical lens. Although the positive vibe established in the morning was still present, the afternoon activities specifically queried external stakeholders to identify ways the COED needed to improve their teacher training practices. To ease this transition, the first activity was an alumni panel. Questions were brainstormed in advance but also generated by stakeholders during TED Talks (via the Post-it Note comments) and from the floor. The alumni panel lasted about 45 minutes.

After the panel, roundtable groups were convened with the primary purpose of generating problem-solving conversations between external stakeholders and university faculty, based on everything seen/heard so far. Participants were assigned to groups, ensuring a diverse mix of perspectives at each roundtable. The groups had structured questions to guide the conversation, generated by the planning committee in advance with input from the dean and at-large faculty. The roundtable questions posed to stakeholders were as follows: *1) How can P-12 and Higher Education work together to address the needs of both groups?; 2) What has impressed you about what you've heard today?; 3) How can we facilitate/further support P12 partners learning more about edTPA (a key candidate assessment required for licensure in our state)?; 4) How can we help principals feel more comfortable about hosting/mentoring teacher candidates in their schools?; and 5) How do we resolve the tension between candidate quality and candidate demand?* Each roundtable group briefly shared with the larger group and the individual roundtable data were collected. The entire process took approximately 65 minutes.

External stakeholders were also able to articulate their impressions of the overall quality of the candidates, and by extension, of the program.

As the last activity of the day, all participants completed an *exit ticket* before leaving the event. The purpose of the exit ticket was to gather stakeholder reaction to what they had heard/seen and collect recommendations for future program improvements. The exit ticket asked three open-ended qualitative response questions, followed by four quantitative questions asking participants to respond to a statement on a Likert scale from 1 to 6, with

1 being *strongly disagree* to 6 being *strongly agree*. The three open-ended responses were: *1) What did you hear or see today that impressed you?; 2) What did you not see or hear but wish you had?; and 3) What did you see or hear that worried you, and why? The four quantitative questions with the 1 to 6 rating scale were: 4) I am glad that I attended this event; 5) I felt that the event was well organized and well planned; 6) Activities were helpful in understanding the education programs; 7) The College of Education (COED) is effective in preparing good teachers.* A final open-ended response space allowed for any additional comments participants wished to provide.

External Feedback for Program Improvement

After the charrette, all the data sources were reviewed to determine if the charrette project objectives where met. Written responses (like the roundtable notes and the exit tickets) were compiled and coded thematically. Most of the data was qualitative in the form of responses to open-ended questions but there were some quantitative data collected from questions 4–7 on the exit tickets. Coding was completed by the CAD and the COED Associate Dean independently and then reviewed for accuracy.

Stakeholders found the activities to be impactful on the overall assessment process and helpful in building cohesive relationships between the COED and the stakeholder community.

In determining whether Objective 1 was met *(external stakeholders will use candidates' stories/work to assess program effectiveness)*, data results indicated that stakeholders were able to (a) ascertain some of what candidates were learning/had learned during the program of study; and (b) gain an overall impression of the quality of candidates being produced by the program. Participants' responses included specific information that stakeholders had gleaned from candidates' stories and presentations. In response to roundtable and exit ticket questions asking *What did you hear or see today that impressed you?*, one participant said, "[I] really enjoyed the poster presentations that showed a pre-assessment, intervention, and post-assessment with one or a group of students … specific and concrete examples." Another said she was impressed by, "The focus on relationships and culture as critical components of effective classroom practice." A third stakeholder noted that it was impressive to see, "Alumni and candidates knowing how to use assessment data to inform instruction." A fourth added, "Pleased to hear about the number of early opportunities that teachers [candidates] have to be in the schools." In addition to these specifics, external stakeholders were also able to articulate their impressions of the overall quality of the candidates, and by extension, of the program. Many of the written responses included complimentary comments about the quality of the candidates and their stories. "I really appreciated the candidates' stories in the TED Talks," wrote one stakeholder. "I was extremely impressed with hearing about other student's experiences and learning about the different issues from several different perspectives," added another. A third participant commented, "So impressed with stories I heard from alumni. Excited for future collaboration opportunities."

Not all of the external stakeholder feedback was laudatory. For example, stakeholders noted that, based on what they observed during the charrette, classroom management was an area where candidates may need additional support. "What is [the college] doing to prepare teacher candidates for better classroom management?" asked one participant. "Student [poster presenter] said relationships were only important for impoverished students," noted another. Technology was another area specifically noted for possible improvement. "More on technology integration," stated one participant, while another noted, "I'd like to hear more about what alumni think of integrating technology in the classroom." They also noted the need to engage candidates in opportunities to practice their knowledge prior to entering the classroom. "I worry a little bit that candidates are better at writing lesson plans than they are at teaching practices," one participant said. Another perceived that students might not be getting appropriate feedback from faculty on their performance: "I asked a couple of student teachers what specifically they were working on to improve their practice and none of them could name one or two things specifically (they said 'everything'). Wondered what kind of targeted feedback they are getting."

Responses also indicated that stakeholders found the activities to be impactful on the overall assessment process and helpful in building cohesive relationships between the COED and the stakeholder community. One respondent said, "I appreciate the transparency.

It was a risk to have an alumni panel. It shows me that [the college] wants honest feedback to improve." Another said it was impressive to see, "COED's commitment to growth and improvements. [They are] not complacent in success." A third noted, "Relationships matter. [I am impressed by] the willingness of [the college] to elevate, risk getting feedback." Finally, stakeholders were asked to rate their level of agreement with the following statement: "These activities were helpful in understanding COED programs (exit ticket question 6)." On a 6-point scale, 94 percent of participants (51 of 54) rated the item as either a 5 or 6, with 6 being the highest level of agreement.

In evaluating whether Objective 2 *(external stakeholders will provide actionable feedback to COED faculty for future improvements)* was met, data results indicated that external stakeholders were able to provide specific feedback to faculty that could be used in the next phase of program changes. In addition to the areas already noted, two additional themes emerged from the analysis: the first was categorized as *P-12 Partner Involvement/ Collaboration*, and the second as *qw*. Each of these themes provided possible avenues for improvements in different ways. The first identified theme, P-12 Partner Involvement/ Collaborations, provided information on how to better engage school stakeholders, particularly cooperating teachers who work directly student teaching candidates. Stakeholders noted that hearing from cooperating teachers about what candidates know and are able to do would be impactful. "Would have loved for CTs [cooperating teachers] to have participated and also been present," noted one participant. "I wish I had heard more about coaching professional development and support to coordinating teachers/district support staff, to provide beneficial reflection feedback to candidates." Other responses coded to this theme indicated a desire to deepen and extend collaborations between P-12 partners and the COED. "We need to continue to develop tight alignment between P-12 and university partners," one participant wrote. "I'd like to see more specifics with how districts can be better partners with university," stated another.

As current students and recent graduates came together with external stakeholders and faculty to tell their stories, a collective narrative emerged that was informative and purposeful.

The second improvement theme identified, Concerns About the Profession, reflected participants' concerns about the obstacles they perceive are in place to hinder teacher development and recruitment. For example, several of the alumni had commented that during their early clinical experiences as candidates at least one or more practicing teachers had encouraged them to consider another profession. External stakeholders seemed to view this as the result of policy decisions, including an inadequate salary scale and increased accountability demands. One participant wrote, "It's concerning that students considering entering the teaching field are being deferred or deciding not to enter the teaching field because of negative experiences. There are a lot of great teachers being lost at the early stages." Another added, "I am concerned with college/graduate students getting the message from the outside world that teaching is not a place to go." These comments were framed in the context of the current teacher shortages in the state: "[This is] reaffirmation that there is a very real shortage of future teachers." Several participants noted that next steps should include policymakers: "Seems like a bigger picture charrette to include policy and decision makers would be a good next step."

Discussion

How do we measure the impact of a college education? Universities and colleges have long wrestled with this question. The concept of what it means to be well-educated is highly subjective and content specific. The easiest measures to collect, such as graduation and loan default rates, do not necessarily reflect the complex nuances inherent in the teaching and learning process. One way to address this gap is to allow students and external stakeholders to engage one another in discussions of what students actually have learned and are able to demonstrate. Further, by using a charrette model—a model purposefully designed to solicit feedback for improvement purposes—another layer is added to the student-faculty-stakeholder interactions. What could begin as a show-and-tell experience can evolve into a multi-faceted conversation, one that celebrates the good aspects of a program and asks for authentic feedback about program improvement from external stakeholders.

The charrette alone would be ineffective in assessing all students, but the charrette could be effectively utilized as one data set in a larger assessment model for program evaluation

This was indeed the outcome of our pilot charrette focusing on student stories as a method of assessing College of Education academic programs. As current students and recent graduates came together with external stakeholders and faculty to tell their stories, a collective narrative emerged that was informative and purposeful. Feedback regarding specific programs and the program/college as a whole was provided. Much of the feedback was positive; however, the nature of the data collection does raise questions about the authenticity of the feedback. Would the feedback have been less positive if the participants' perceptions had been collected in some other fashion? This is difficult to determine. The COED also presented work from some of our best students, which also muddies this question. This is why the structure and setup of the charrette are so important. By taking the initiative and asking for critical feedback from stakeholders, faculty send a clear message that they seek to improve. The charrette identified areas of celebration as well as areas for improvement. In fact, because the afternoon sessions specifically requested a critical lens, the shared information allowed external stakeholders to witness first-hand the commitment of faculty in making authentic, data-based decisions. Students also benefitted from participation in these events. By using a metacognitive approach that allowed them to reflect on what they had learned, students gained valuable experience with critical thinking and with presenting their thought processes to others.

Data collected and analyzed indicated that both objectives for the project were met; external stakeholders were able to assess the current academic program and they were able to provide actionable feedback to faculty for future improvements. The impact of the narratives was powerful; framing the narrative with the charrette model communicated to stakeholders the eagerness and sincerity of faculty to engage them as partners in the improvement cycle. The COED faculty used the feedback provided in this pilot (along with other sources) to write a grant funding a collaborative summer institute for faculty and the teachers who mentor our candidates. This summer institute (now in its second year) focuses on breaking down silos between faculty and P-12 partners to strengthen and collaborate on teacher preparation practices. The charrette pilot was a first step in this process. In addition, while the COED has not replicated the charrette event in its entirety, we have seized on the momentum with our P-12 partners to conduct annual data days with stakeholders, where we replicate parts of the charrette model by sharing COED candidate assessment data and asking for critical feedback. These have been highly successful and strengthened our relationships with stakeholders in the region.

Suskie (2016) and others point out the need for higher education to simplify the complexity of program assessment for stakeholders. As funding for higher education becomes more anchored in accountability, this need is a practical as well as a moral one. The Spellings Report on the Future of Higher Education (U. S. Department of Education, 2006) noted the need for higher education to better document student outcomes, ensuring that taxpayers and the students themselves are securing a return on their investment of education dollars. Student learning outcomes are rapidly being pushed to the background as quantifiable metrics take center stage. In this context, creating a model of assessment where: student outcomes are the focal point; a variety of student work samples and interactive sessions are presented; and external stakeholders can engage directly with students and receive real-time feedback on what students know and can do; then, would seem to embody high-quality outcome-based assessment.

Astin and Antonio (2012) advocate for a talent development approach to assessment, one where both students and faculty improve their work product as a result of direct, actionable feedback. Taking the additional step of engaging with external stakeholders in a charrette to solicit actionable (and perhaps critical) feedback to improve the quality of the program can also build trust among university faculty and external stakeholders. There are several possible avenues to explore in considering how to build upon this pilot project moving forward. While our charrette included only external stakeholders who were also employers (e.g., principals, school district leaders, etc.), a next logical step might be to include policymakers as an additional external stakeholder group for future charrettes. It would be interesting to learn if the different external groups (policymakers versus employers) would

arrive at similar conclusions. The COED faculty have discussed another charrette in the future that does engage policymakers as well as P-12 school district stakeholders but some faculty have expressed reluctance to pursue this, due to the changing political landscape related to teacher preparation and accountability. If offered, resources would need to be set aside specifically for this work, including funds to cover logistical costs (food, event space, posterboard printing) and stipends for student participants. We have agreed to revisit the issue moving forward.

Limitations of the charrette pilot

Some limitations do exist with this model as we piloted it. A limited sample of students were directly assessed in the charrette; however, in considering the goals for our project, our focus was on engaging external stakeholders through students' stories for program evaluation and feedback. We were not attempting to assess individual student proficiency. The charrette alone would be ineffective in assessing all students, but the charrette could be effectively utilized as one data set in a larger assessment model for program evaluation. The students who presented at the charrette may not be representative of all students in a program, and their experiences may not reflect the viewpoints of other students in the college/program. We also selected top-tier students to participate, those considered by faculty to be capable of producing good work in a timely fashion. This was a deliberate choice on our part; first, we wanted to ensure we had good work to present to stakeholders and we were on a timeline—we needed students who could be relied upon to assist. This was a realistic logistical consideration; the faculty work group did consider inviting less-proficient students to participate, but reasoned that by showcasing our best student work and then asking for critique, the feedback would still be quite informative; any perceived gaps among our best-and-brightest would be taken more seriously by faculty. This choice did create the possibility, however, that there are weaknesses in our programs that were never exposed for charrette participants. While our pilot project objectives would still be met, additional or different areas of candidate growth may have been identified if we had included less academically proficient students in the charrette. This is one additional reason why triangulation of program data outcomes is a critical part of program evaluation work; the charrette can serve as one source of information, but multiple sources should be considered in making decisions about program improvement.

As the accountability movement continues its march into the world of higher education, questions regarding how to best assess learning for a college graduate (and effectiveness of college programs) will remain.

The charrette primarily produced qualitative data, which could reflect a biased perspective from some stakeholders. The data produced are valid only to the extent that the participants truthfully responded to the questions. Much of the feedback received was actionable for the COED faculty because our programs follow a similar course trajectory. However, some participants did provide program-specific feedback but neglected to note which areas of improvement from the roundtable questions and exit tickets best applied to specific programs. It would be helpful if the question formats in any future projects were revised to require stakeholders to indicate exactly which programs they were talking about with their feedback. Although COED faculty in this project were receptive to external stakeholder feedback, not all faculty may be as welcoming to outsiders commenting on their programs. Some additional conversations or professional development may be needed to generate faculty buy-in prior to implementation. The results generated by the charrette are not part of the established metrics typically used to assess higher education; some additional context might be needed to frame the results cleanly for external stakeholders.

Conclusion

As the accountability movement continues its march into the world of higher education, questions regarding how to best assess learning for a college graduate (and effectiveness of college programs) will remain. Engaging with current students and recent graduates as they share their own stories of what they have learned (or not learned) can inform external stakeholders as to exactly what occurs in program coursework and where potential gaps/areas of improvement persist. This kind of feedback is context and program specific, and provides more actionable information for faculty than other broad measures, such as the results of a standardized test or graduation rates. Faculty and stakeholder buy-in for this kind of story-telling assessment narrative may also be greater—and the information more valuable—than with other forms of assessment. As institutions face increased pressures to better assess and improve their programs, engaging stakeholders with students in the charrette model may be a viable and much-needed option for higher education assessment.

References

Association of American Colleges and Universities. (2016). *Trends in learning outcomes assessment: Key findings from a survey among administrators at AAC&U member institutions*. Retrieved from https://www.aacu.org/sites/default/files/files/LEAP/ 2015_Survey_Report3.pdf.

Astin, A. W. & Antonio, A. L. (2012). *Assessment for excellence: The philosophy practice of assessment and evaluation in higher education* (2nd ed.). Lanham, MD: Rowman & Littlefield, Inc.

Baer, L. (2015). *Quality, accountability and action* [White paper]. Retrieved from the Accrediting Council for Independent Colleges and Schools website: http://www.acics.org/ contact/content.aspx?id=6382.

Bean, J. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. San Francisco, CA: Jossey Bass.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in education: Principles, policy, & practice*. Advanced online publication.

Braskamp, L. A., & Engberg, M. E. (2014). *Guidelines for judging the effectiveness of assessing student learning*. Loyola University Chicago: Chicago, IL. Retrieved from http://www.learningoutcomesassessment.org/documents/BraskampGuidelines.pdf.

Carlson, K. (2017, December 17). State approves new performance-based funding measures for higher ed [Online article]. *The Missourian*.

Council of Regional Accrediting Commissions. (2018). *A one-year review of the Council of Regional Accrediting Commissions' graduation rate information project*. Retrieved from http://download.hlcommission.org/C-RAC_Grad_Rate_Study_2018-02.pdf.

deBoer, F. (2016). *Standardized assessments of college learning: Past and future*. Retrieved from New America website: https://static.newamerica.org/attachments/12843-standardized-assessments-of-college-learning/Assessment-White-Paper. 72b9bbb2cf0447b4aacaabe6e6ea1824.pdf.

Emil, S., & Cress, C. (2014). Faculty perspectives on programmed curricular assessment: individual and institutional characteristics that influence participation engagement. *Assessment & Evaluation in Higher Education, 39*(5), 531–552.

Ewell, P. T. (2009). *Assessment, accountability, and improvement: Revisiting the tension* (Occasional Paper No. 1). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Herman, J., & Hilton, M. (Eds.). (2017). *Supporting students' college success: The role of assessment of intrapersonal and interpersonal competencies. A consensus study report of the National Academies of Sciences, Engineering, and Medicine*. Washington, DC: National Academies Press.

Jones, A. (2007). Multiplicities or manna from heaven? Critical thinking and the disciplinary context. *Australian Journal of Education, 51*(1), 84–103.

Jones, P., Carr, J. F., & Ataya, R. (2007). *A pig don't get fatter the more you weigh it: Classroom assessments that work*. New York, NY: Columbia Teachers Press.

Kelchen, R. (2018). Higher Education Accountability. Baltimore, MD: Johns Hopkins University Press.

Kreighbaum, A. (2016, September 21). Tougher scrutiny for colleges with low graduation rates [Online article]. Inside Higher Ed. Retrieved from https://www.insidehighered.com/news/2016/09/21/regional-accreditors-refocus-institutions-low-grad-rates.

Leveille, D. E. (2013). Accountability in postsecondary education revisited. UC Berkeley Research and Occasional Papers Series. Retrieved https://escholarship.org/uc/item/ 61h1b3p3

Morgan, H. (2008). How place managers can harness the power of stakeholder feedback to improve satisfaction. *Journal of Place Management and Development, 1*(3), 307–314.

Popham, W. J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership, 56*(6), 8–15.

Popham, W. J. (2016). Standardized tests: Purpose is the point. *Educational Leadership, 73*(7), 44–49.

Ravich, D. (2013). *Reign of error: The hoax of the privatization movement and the danger to America's schools*. New York, NY: Alfred A. Knopf, Inc.

Santos, P. (2016). *Teaching critical thinking skills in a disciplinary context*. Retrieved from Teachers of English as a Second Language Association of Ontario website: http://contact.teslontario.org/teaching-critical-thinking-skills-in-a-disciplinary-context/.

State Council of Higher Education for Virginia. (2015). *Measures, targets, and related indicators (statewide strategic plan for higher education 2014-10-28)*. Retrieved from http://www.schev.edu/index/statewide-strategic-plan/measures.

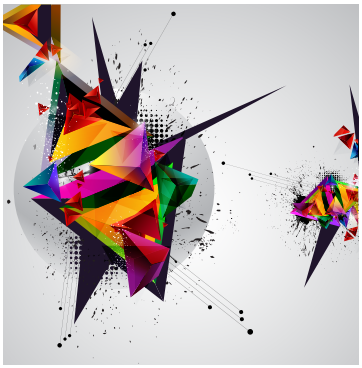
Suskie, L. (2016, April 10). A response to Bob Shireman on "inane" SLOs [Web blog post]. Retrieved from http://www.lindasuskie.com/apps/blog/show/43910127-a-response-to-bob-shireman-on-inane-slos

Ulewicz, R. (2017). The role of stakeholders in quality assurance in higher education. *Human Resources Management & Ergonomics*, XL, 93–107. Retrieved from https://www.researchgate.net/publication/317301877.

U.S. Department of Education. (2006). *A test of leadership: Charting the future of U. S. higher education*. Washington, DC: USDOE. Retrieved from https://www2.ed.gov/about/ bdscomm/list/hiedfuture/reports/final-report.pdf.

AUTHORS NOTE

Laura C. Hart, Office of Assessment and Accreditation, Cato College of Education, UNC Charlotte; Christine Robinson, Academic Affairs Assessment Office, UNC Charlotte.



Abstract

This paper introduces a conceptual framework for overcoming common assessment challenges and supporting a positive assessment culture in higher education through fostering collaborative relationships with faculty and staff. By using a lens that integrates concepts from person-centered and solution-focused counseling, positive psychology, and motivational interviewing, assessment practitioners can better understand what guides the cultivation of inclusive and participatory relationships in assessment. The RARE model provides a common set of strategies for implementing principles of effective assessment practice, developed by two assessment professionals from universities located in different accrediting regions: WASC (Western Senior College and University Commission) and SACSCOC (Southern Association of Colleges and Schools Commission on Colleges). In calling attention to the influence of their practitioner training and background, this model also highlights the benefit of exploring the disciplinary diversity that exists within the assessment field. Through exploration of this reflexive, strengths-based approach to assessment practice, the authors contribute to the discourse about professional identity in higher education assessment.

AUTHORS

Emilie Clucas Leaderman, Ed.D.
Santa Clara University

Gina B. Polychronopoulos, Ph.D.
Christopher Newport University

Humanizing the Assessment Process: How the RARE Model Informs Best Practices

Higher education assessment is a complex professional identity, as some practitioners are recruited as a faculty or staff member to their role while others transition from outside of academia. Our understanding of disciplinary identity in higher education assessment is emerging, with assessment professionals entering their positions from a broad range of academic fields (Suskie, 2009). Data from Nicholas and Slotnik (2018) confirm significant disciplinary diversity, with the majority holding their highest degrees in social sciences (30%) and education (44%). Social science respondents in this survey ranged from psychology, sociology, history, and organizational leadership disciplines; while those in education included higher education, administration, leadership, educational psychology, assessment and measurement, and curriculum and design. A closer look at the years of experience among this same group of professionals reveals that 75% have moved into the profession within the last seven to 10 years (Nicholas & Slotnik, 2018), which suggests that the assessment field is evolving. While the disciplinary paths of assessment practitioners are identifiable, their industry and career experiences, as well as implications for assessment practice, remain to be seen.

To complicate identity matters, our profession has a reputation of not being well liked among its counterparts in academia. Following recent media editorials from faculty criticizing the aims of higher education assessment (Worthen, 2018) and articles calling for our community of practitioners to better define standards of practice for themselves (Eubanks, 2017), we recognize the need to articulate how we develop interpersonal relationships with faculty and staff. One recent survey points to assessment professionals' desire for learning strategies for successfully overcoming unique or common challenges at their institution (Combs & Rose, 2016), and faculty resistance has been considered one of the main barriers to successful measurement of student learning outcomes (Katz, 2010). Although collaboration and relationship-building have been identified as best practices toward meaningful engagement in the assessment process (Kinzie, Jankowski, & Provezis, 2014), it has also been suggested that inclusive practices are necessary to establish

CORRESPONDENCE

Email
eclucasleaderman@scu.edu
gina.polychronopoulos@cnu.edu

assessment as a sustainable process (Hutchings, Ewell, & Banta, 2012). In undertaking inclusive practices, assessment professionals can consider how they engage and involve staff and faculty.

As the public and those inside of the academy continue to scrutinize motivation within our field, we also view self-reflection as essential for navigating its future. The field of educational development, also known as academic or faculty development, has engaged in a similar dialogue at both the international (Green & Little, 2016) and national levels (Green & Little, 2013; Little, 2014; Sorcinelli, Austin, Eddy, & Beach, 2006) to gain a better understanding of its complex professional identity as a community and the implications for research and practice. As the assessment profession continues to define itself, some common interpersonal roles and their respective tasks have already been identified through research. Several terms used to describe the relational nature of our work include (a) facilitator/guide: mentoring individuals, assisting people in the assessment process, and collaborating across disciplines; (b) political navigator: emphasizing people skills, sensitivity to culture, collaboration, and framing sensitive results (Jankowski & Slotnik, 2015); and (c) change agent: "responding to weaknesses; designing change; reflection; redesign; using results; making a difference; and closing the loop" (Ariovich, Bral, Gregg, Gulliford, & Morrow, 2018). Although both of the previous studies suggest that these necessary interpersonal roles intertwine, it is unclear how assessment professionals as a collective group prepare themselves for this change-oriented work.

In seeking to better understand how our own background and training in counseling inform our interpersonal work in higher education assessment, we reflected on the philosophical underpinnings that inspire our individual approaches to working with faculty and staff in the assessment process. We formalized our thinking into a model of reflexive, strengths-based assessment practice. The following questions guided the creation of our model:

1. How do our own professional identities influence our work with staff and faculty in assessment? Which theories have shaped our current practices in assessment?
2. How do we as assessment professionals cultivate participatory relationships with faculty and staff? What strategies have we used to develop healthy assessment cultures that allow for inclusive best practices to occur?

Although examples of successful interpersonal strategies have been well documented, the theoretical mindset and processes that guide practitioners in inclusive assessment work have yet to be explored. Kinzie and colleagues note that integrating this type of generative assessment as effective practice continues to be a challenge for institutions (Kinzie et al., 2014). Moreover, the Watermark study suggests that the roles and competencies of assessment practitioners should be further explored and defined (Ariovich et al., 2018). Considering all of these recommendations, it is beneficial to examine collaborative practice—including building strong relationships with faculty and staff—as a key aspect of successful engagement in assessment and enhancing institutional assessment culture.

As the public and those inside of the academy continue to scrutinize motivation within our field, we also view self-reflection as essential for navigating its future

Purpose Statement

The multitude of disciplinary backgrounds within the assessment profession lends itself to ambiguity in defining clear links to developing effective assessment practice. However, we see this diversity of disciplines as a strength within our field. This paper aims to describe a model for a strengths-based approach to assessment practice (which was informed by our shared background in professional counseling) as well as how the underlying theories link to best practices in assessment. Our purpose in developing the RARE model is to demonstrate how we use our disciplinary lens to create participatory and inclusive relationships with faculty and staff in the assessment process.

We propose our approach as a set of strategies for developing inclusive partnerships with faculty and staff in good assessment work. Derived from our experiences as counseling

professionals and from several counseling theories, the following model demonstrates strategies from person-centered (Rogers, 1950) and solution-focused therapies (de Shazer, 1985), motivational interviewing (Miller & Rollnick, 1991, 2002), and positive psychology (Seligman, 2011) to conceptualize how these particular theories inform our practice by describing them in action. Please note that although these strategies are inspired by counseling skills we do not intend for them to be employed as such by assessment professionals; doing so would be unethical and outside the scope of assessment practice. However, it is our hope that our initial discussion of the influence of our disciplinary backgrounds, as well as engaging in reflexive, strengths-based assessment practice, will spark future discourse within the field and contribute to the development of our collective professional identity.

RARE Model

The RARE model approach is informed by both humanistic and postmodern counseling theories.

The RARE model approach is informed by both humanistic and postmodern counseling theories (R for Relate, A for Acknowledge, R for Reflect, and E for Empower), representing four groups of strategies inherent in these theories. Humanistic theory emphasizes the importance of an egalitarian relationship between the client and counselor (Hansen, 2006). That is, the counselor is not considered the "expert" in the relationship; rather, he or she partners with the client in an effort to understand his or her experience. There is often a distinct power differential between assessment professionals and faculty/staff members in higher education, which may contribute to a sense of cautiousness about engaging in the assessment process. Although the assessment practitioner may indeed have expertise, the person-centered approach of the RARE model seeks to minimize the power differential in an effort to strengthen relationships and promote a collaborative and inclusive culture of assessment. Relatedly, the RARE model also has postmodern theoretical influences in that collaboration and co-construction of meaning are the primary tenets; the counselor and client create the knowledge together through multiple perspectives on the problem (Sanders, 2011). In this manner, the assessment practitioner does not impose his or her knowledge upon the faculty/staff member; instead, they work together to construct the assessment process and interpret meaning.

The RARE model emphasizes four components of effective assessment practice, as well as strategies within each component. The following is a brief description of each strategy, including their specific theoretical underpinnings.

1) **R–Relate:** Effective assessment practice includes building relationships with others. The foundation of this model stems from the person-centered approach to counseling developed by Carl Rogers (1950), using three core conditions necessary for establishing a trusting relationship and working effectively toward goals: genuineness, unconditional positive regard, and empathy. Some of the humanistic strategies employed in this non-directive approach include active listening, reflection, and clarification.

2) **A–Acknowledge:** To foster collaborative relationships, it is helpful to recognize and highlight the strengths of others. This postmodern component builds upon humanistic counseling theory by recognizing and acknowledging the strengths and resources that faculty/staff bring to the assessment process. Solution-focused in theory, the intent is to guide faculty in setting their own goals. Some strategies employed in this category include: supporting a collaborative relationship, building upon strengths, and changing the "doing and viewing" of the problem (Murphy, 2008).

3) **R–Reflect:** When working toward change, noticing and embracing resistance will help to redefine it and promote growth. Motivational interviewing (MI) focuses on the power of the individual in creating change by meeting faculty members at their current level of assessment practice. MI strategies employed in promoting assessment growth include: collaboration (vs. confrontation), rolling with resistance, and developing discrepancy (Miller & Rollnick, 1991, 2002).

4) **E–Empower:** Linking all of the components together, we strive to empower faculty and staff to engage in meaningful and autonomous assessment practice. Core elements of positive psychology are used to better understand faculty needs and facilitate their readiness to act using guidance and support. In practice this translates into: identifying the great assessment work faculty are already doing, often referred to as self-efficacy; and supporting faculty as they take the next new step. Positive psychology strategies employed include: focusing on what is right/going well, building upon what is enabling success to help them flourish, and PERMA (positive emotion, engagement, relationships, meaning, and accomplishment (Seligman, 2011)).

Figure 1. RARE Model

As demonstrated in Figure 1, each of the four components is represented by a link. The links are joined together to demonstrate an interconnected method of applying the strategies of the RARE model into assessment work. Each component linking together also symbolizes an equitable partnership in working with faculty and staff. Overlap is inherent in the strategies throughout the components of the RARE model as they derive from conjunctive theories (i.e., humanistic and postmodern counseling theories). Accordingly, the RARE model is an integrative application of counseling strategies that facilitate trust, motivation, and change within the practice of educational assessment. In the following section we will discuss each component of the model in greater depth and offer examples of how one might use the strategies of this approach. The examples highlight some of the challenges that we may face as assessment practitioners, such as resistance from colleagues, making changes for improvement, finding meaningful data, and promoting faculty autonomy and positive assessment culture. The RARE model applies to assessment work with not only academic faculty members but also administrative and student affairs staff; therefore, discussions throughout this manuscript will include references to these individuals interchangeably.

The RARE model is an integrative application of counseling strategies that facilitate trust, motivation, and change within the practice of educational assessment.

R–Relate

Building strong relationships with faculty and staff members relies on gaining their trust. Relate involves meeting with faculty and staff individually to establish oneself as a supportive colleague. In the counseling relationship a crucial aspect of building trust includes active listening, a term developed by Carl Rogers (1951) that describes a way of being that facilitates rapport and understanding. Translating this concept to assessment work, an intentional focus on learning about an individual's experiences and about his or her assessment challenges communicates the belief that the faculty or staff member is the expert in the relationship. Another strategy of this component includes conveying empathy by recognizing individuals' experiences as valid. Listening actively involves repeating back

what you heard from an individual by paraphrasing what that person said to you in your own words. Clarify that you have heard their assessment concerns correctly by asking questions to check for understanding. Inquire about potential areas where you are less certain of their meaning. Communicate a genuine interest in interpreting individuals' experiences as accurately as possible. Finally, validate the feelings they express with unconditional positive regard.

Building strong relationships with faculty and staff members relies on gaining their trust.

Although this technique can feel awkward and robotic at first, it will feel more natural and genuine over time as it is repeatedly practiced. During initial meetings with faculty or staff set an agenda that focuses on learning from them. Build trust with individuals by asking about their curriculum, programs, services, research, and students. Invest time in developing a shared understanding of their perspectives of assessment (including current challenges and previous frustrations regarding what has not been effective or meaningful), and acknowledge small victories when possible. As they identify and describe problems with assessment, create tools and adapt resources that meet their specific needs. Convey a tone in reports, e-mails, and face-to-face meetings that communicates a willingness to see the problems as they do. The strategies within this component are essential in helping the faculty or staff member perceive the assessment professional as an ally and it will enhance their ability to overcome challenges collaboratively in the future.

A–Acknowledge

Once the assessment practitioner has developed positive relationships with faculty and staff members the next step is to recognize the individual strengths that they bring to the assessment process. Rooted in solution-focused theory (de Shazer, 1985), this postmodern approach builds upon humanistic counseling theory by acknowledging the independence of the individual with whom you are collaborating, helping to guide, versus lead, the individual in establishing meaningful goals. There are several solution-focused strategies that one may use in a solution-focused approach to educational assessment: supporting a collaborative relationship, building upon resources (i.e., strengths) and exceptions, and changing the "doing and viewing" of the problem (Murphy, 2008). For assessment professionals who are trying to strengthen their alliances with faculty and staff members on assessment practices, this component is essential, as it lets them know that you recognize the work they are already doing and will continue to work alongside them in a partnership capacity.

Supporting a collaborative relationship. Faculty and staff are the experts in their discipline or professional areas, and they are undoubtedly engaging in some form of assessment in their regular practices. For instance, they may be constructing exams and assignments to evaluate how well students have mastered specific learning outcomes, collecting data about what services students utilized most frequently on campus, or distributing student opinion surveys. In supporting collaborative relationships it is important to highlight the individual strengths and contributions that one brings to the assessment table, especially before offering any suggestions to a faculty or staff member. Murphy (2008) suggests the following strategies for establishing a collaborative and change-focused relationship: approach others with humility and a desire to learn from their perspectives, use language consistent with change, and solicit feedback on the collaborative process. An example of employing collaborative strategies in assessment can be seen in how we approach faculty and staff when working together. We can minimize the existing power differential by adjusting how we intervene. If you are reviewing an assessment plan with a faculty member and notice that their data collection methods are not aligned with the learning outcomes, ask for information with a focus on learning more from them, which then invites a conversation. For example, "I'm wondering about this particular assignment and what it looks like," or "I'm not sure I understand how the questions on this test relate to this outcome. Is it possible that students might be learning about this outcome in another area of your program?" or "Could there be another way to word this learning outcome that might reflect what you hoped they would learn? What do you think?" By shifting our own perspective to one of curiosity, adjusting our language to be less absolute, and inviting others to share their perspectives and reactions, we are fostering collaboration and promoting autonomy of faculty and staff in the assessment process.

Building upon faculty/staff strengths. While continuing to nurture positive and collaborative relationships with faculty and staff, the assessment practitioner has a good understanding of what strengths and resources they bring to the process. The next step is to highlight what has been more or less helpful, acknowledge what has worked in the past (not simply what did not work), and discover places where those problems did not exist (i.e., exceptions). An example of this strategy is learning what type of information (i.e. data) faculty or staff are already gathering in their regular practices (e.g., exams, meetings with students, outreach programs, capstone projects). Then, the assessment professional could explore (a) what has worked in the past with the goal of doing more of what works (building upon resources); (b) what has proven to be a challenge or obstacle; and, most importantly, (c) where an identified obstacle does not manifest itself in their assessment process (i.e., exception). For example, if a faculty member describes how challenging it can be to engage fellow faculty members in submitting assessment data, the assessment practitioner might inquire about those faculty who typically have been good about submitting assessment data, discussing how that process was different. Employing this strategy helps to lessen resistance because one is not imposing additional burden upon them or their existing processes. Building upon exceptions also allows faculty and staff to feel more positive about the process, while assessment practitioners focus on what has worked and where obstacles did not exist.

Changing the doing and viewing of the problem. Assessment professionals have long practiced the mantra of "doing more of what works, and if it doesn't work, then do something different," which is also prevalent in solution-focused interventions. However, one challenge in assessment occurs when faculty or staff get stuck in a pattern of maintaining the status quo through continuing to conduct assessment tasks that are easy to do but have consistently lacked value for the department or program, and have not led to usable results. This type of resistance can make it difficult to implement change. Changing the doing and viewing of the problem involves facilitating a shift in perspective for faculty and staff. Supporting collaborative relationships, while highlighting strengths, will facilitate this reframing process. As a result, the assessment professional is in a good position to help faculty and staff overcome their resistance in trying something different. Faculty and staff are more likely to receive suggestions when the assessment professional has already communicated a genuine, caring interest (by employing strategies from the R and A components). Changing the doing and viewing of the problem involves going with the flow, inviting criticism and feedback, and frequently requesting client input (Murphy, 2008). For example, when a faculty member shows concern that their data collection methods are not providing useful findings, the assessment practitioner can facilitate an idea based on a pride point already expressed by fellow colleagues in the department. For instance, "several faculty seem pleased with the quality of the capstone projects, what about the thesis essays that students write during their senior seminar? How might the essays offer insight as to how well students are performing on those learning outcomes?" By helping the staff member to shift perspective from what is not useful to what could be useful, the assessment professional is acknowledging existing resources in the process, helping the staff member to focus on positive and productive actions rather than the problems with assessment.

The goal of building motivation for change is future oriented: plant a seed that will bloom later.

R–Reflect

Assessment professionals are often responsible for facilitating decisions supported by data, but we also understand that changes happen slowly in higher education. The Reflect component uses MI strategies to notice resistance and work together with the individual faculty or staff member to redefine it. The mindset involved in these strategies involves gentle persuasion and unconditional support that focuses on enhancing readiness for change. Originally developed for addictions counseling, MI strategies have been notably effective in resolving ambivalence, which is often a barrier in taking the next step. In acquiring this stance, it is important to have "a strong sense of purpose, clear strategies and skills for pursuing that purpose, and a sense of timing to intervene in particular ways at incisive moments" (Miller & Rollnick, 1991, pp. 51–52). In assessment work, this component can be most useful when facing resistance from faculty or staff and understanding when they are ready to take action independently, if at all. Especially if departmental dynamics appear challenging, these strategies can be incredibly helpful as they encourage the individual to

think deeply about what would be meaningful in their assessment process. Much of the work in this category involves helping the individual to see and articulate this meaning, rather than the assessment professional prescribing an assessment task.

Enhancing readiness to change relies on a strong relationship between the practitioner and the individual faculty/staff member as well as the autonomy of the individual. By first gaining a thorough understanding of previous experiences faculty and staff have had with assessment, it will be easier to notice when they are contemplating their next step toward action. Instead of assertively advocating a new idea or suggestion to a colleague, the assessment professional recognizes this person as the expert of their program, curriculum, or course and asks them to identify which option they view as most beneficial, based on how they perceive possible consequences. This tactic invites the faculty or staff partner to weigh both the short-term and long-term advantages and disadvantages of a particular option that they believe will address their assessment needs. When their colleague is hesitant to adopt a particular action, the assessment professional accepts that this individual may not be ready and welcomes the opportunity for discussion to see the problem from their perspective, also known as rolling with resistance in MI (Miller & Rollnick, 2002). At this point, it is helpful to ask follow-up questions without any judgmental tone, to better understand their reasons for ambivalence, while also being careful not to persuade them toward a particular direction or solution.

It can be helpful to remind faculty and staff that assessment is a learning process for everyone involved; perfection is not our goal, and each project furthers collective learning.

The goal of building motivation for change is future oriented: plant a seed that will bloom later. One way of planting a seed is to highlight the consequences one is currently having that conflict with his or her individual values, a skill known as "developing discrepancy" (Miller & Rollnick, 2002). This strategy is consistent with academic freedom principles ("American Association of University Professors," n.d.), as it involves frequent reminders of all choices that are available to the individual. An example of this sounds like, "This option is considered a best practice, but what do you think would fit best for your program/department/discipline needs?" Align assessment strategies as solutions for the problems that faculty or staff are already concerned about, while also validating their feelings and ideas. An example of this could be, "It sounds like you are frustrated with the results because they point to challenges with students in your colleague's courses. How can we communicate the data in a way that meets your department's needs?" When using this style it is important for the assessment professional to find out: (a) what are faculty and staff most concerned about as it relates to student learning, (b) what options have they already considered, (c) which choice(s) seems most plausible based on their identified costs and benefits, and finally, (d) what seems to be getting in the way of trying that option.

This way of connecting in the counseling literature is referred to as holding up a mirror. The practitioner is actively listening without reacting, while reflecting back the problem and potential solutions as they hear them with an empathetic mindset (Rogers, 1950, 1951, 1952, 1954). Reflecting in this way helps the client effectively understand their challenges and weigh their choices realistically and nonjudgmentally. Arnold (2014) suggests that this process facilitates a more focused awareness of available options and eventually leads to tangible efforts, yet the practitioner must be careful to not impose his or her needs on the individual. By entirely focusing on the faculty or staff member and understanding their experiences, the assessment practitioner is tapping into the individual's autonomy and motivation, allowing for change to occur.

E–Empower

The Empower component in the RARE model represents a paramount goal in higher education assessment: to cultivate a positive and inclusive assessment culture on campus by empowering faculty and staff members to feel confident in their assessment practices. Rooted in positive psychology theory, the focus is to make the process of assessment more meaningful, or "fulfilling" in positive psychology terms, and shift our focus from strictly what needs to be fixed or changed (Seligman, 2011). This is not to say that assessment professionals should ignore aspects of the process that are going completely wrong; rather, the goal is to supplement the practice of identifying a hitch in the system with intentional

optimism and empowerment of faculty to build upon resources that will enable them to navigate a more meaningful assessment process.

The core elements of positive psychology align particularly well with the theories inherent in the first three components of the RARE model, i.e., humanistic, solution-focused, and motivational interviewing, and these elements can be applied to help assessment professionals better understand faculty needs. Some of the positive psychology strategies that translate well to assessment practice include the following: (a) focusing on what is going well, (b) building upon what is enabling success to help them flourish, and (c) PERMA (positive emotion, engagement, relationships, meaning, and accomplishment; Seligman, 2011). This is done by offering a balance of supportive language with encouragement that guides faculty and staff to feel more confident experimenting with assessment. It can be helpful to remind them that assessment is a learning process for everyone involved; perfection is not our goal, and each project furthers collective learning. When departments are overly critical in reporting their annual assessment results, the assessment practitioner can help reframe their narrative in the feedback they provide. For example, reframing feedback by recognizing strong methodology, complimenting well-written program learning outcomes, and gently raising questions that promote inquiry about student learning, can encourage departments to pursue their own questions in future assessment processes and yield more meaningful results. By employing the strategies of empowerment, nurturing positive relationships, acknowledging strengths and existing resources, and reflecting upon readiness to change, assessment practitioners can offer faculty and staff the opportunity to take ownership of their assessment process, which ultimately increases faculty buy-in and promotes their engagement in meaningful assessment practices.

Putting it All Together: Case Example

We understand relationship building and collaboration to be central tenets to inclusive assessment and have developed this model as one response to the gap between principles and effective practice. In having academic and professional training that integrates the theories embedded in our model, we also recognize the need to directly apply and further explain the discipline-based language and concepts embedded in the mindset we are proposing in our approach. Below is a case example designed to exemplify the value and benefits of using the RARE model through a common assessment scenario with an academic department.

Dr. Smith is an associate professor of sociology who has recently been tasked with coordinating the assessment process in his department. He is already overwhelmed with his teaching load and research projects and now he has been appointed to lead his colleagues in annual program-level assessment. Assessment duties in his department are turned over to a new person every year or two, and Dr. Smith has inherited the current assessment plan from previous faculty in his department. Although the student learning outcomes are well articulated and align with the overarching goals of his department, Dr. Smith is frustrated because he thinks the assessment process has been a waste of time and resources. In general, he and his colleagues feel that programmatic assessment is "yet another thing we have to do for accreditation, so let's just get it over with." Dr. Smith does not agree with some of the measures that they have been using to assess their goals. The department is using a standardized, content-based test and students are achieving above expectations on it; however, the students are not performing well on their capstone projects. Dr. Smith reached out to the Office of Assessment after attending a meeting facilitated for each department's assessment representative.

Employing the RARE Model

Our primary task as assessment professionals is to ensure that the sociology annual assessment plan is effective in measuring their program learning outcomes. The ultimate goal is to help Dr. Smith and his colleagues construct a more meaningful assessment process that will provide them with valuable information about what their students are learning, not simply fulfill their regional accreditation requirements. Through the process, our hope is to support a positive shift in the assessment culture within the department. Below is an

illustrative example where an assessment professional could employ the strategies within the four components of the RARE model.

Through the process, our hope is to support a positive shift in the assessment culture within the department.

R–Relate. Dr. Smith is new to the assessment process and you are in the early stages of working with him. Building the relationship is essential, as you notice that he already seems hesitant to engage in the assessment process. Picture yourself in Dr. Smith's position and imagine what would be frustrating about the assessment process as he experiences it. While you might not agree with his perspective you can still seek to understand his challenges and verbally empathize, which will support a collaborative relationship. Validate his perspective and validate his frustrations with a process that has not previously helped his department or students taking courses in the major.

A–Acknowledge. Build upon the resources of Dr. Smith and the sociology department. Start by forming allies/advocates within the department, including the faculty who work closely with Dr. Smith and already use assessment strategies in their courses; they may, for example, speak to the benefit of using an authentic assignment, or they can share their perspective of how students are performing in key areas. Become familiar with their program learning outcomes and potential measures that can be used, or are already being used effectively. By identifying the strengths of their current practices and their existing resources, you are helping Dr. Smith to see assessment through a different lens. Encourage his ideas for strategies and solutions. Ask questions to help him identify those resources, such as: What strategies are you already using that work well in your department? How can they be applied to this situation? Also, because you have developed a strong collaborative relationship, while helping to reframe his perspective of the perceived problem, Dr. Smith may be more willing to receive suggestions or ideas from you, the assessment professional.

R–Reflect. When you meet with Dr. Smith ask him about his most pressing concerns as a faculty member. Validate his frustrations and seek to understand his views without being negative. Identify areas where he and his colleagues have taken risks and seen the benefit with student learning (i.e., developed a new course with an innovative teaching approach that has increased enrollment in the major, etc.). Assess where and how the department seems ready to take steps towards action. Show the value of assessment for things that Dr. Smith has identified as a concern (i.e., in getting new sociology courses approved for general education requirements). Meet Dr. Smith and his colleagues where they are in the assessment process, facilitate identifying possible alternative measures (i.e., using rubrics to score key assignments), and assist in weighing the implications associated with each option. Seek to understand how assessment can help with problems they are already interested in and motivated to solve.

E–Empower. After establishing a positive and collaborative relationship, acknowledging strengths, and reflecting on readiness to change, the assessment practitioner can then continue to foster optimism in the assessment process. Recognizing productive changes, no matter how minor, as positive feedback can provide encouragement. Celebrate successes in the process and avoid focusing solely on what is not working. At your next assessment meeting with faculty members as a group, ask Dr. Smith to share his successes in reconstructing a meaningful assessment process for his department. Highlight faculty and staff accomplishments and improvements in their assessment processes to serve as an example of how assessment can be useful and meaningful. These strategies can also promote faculty empowerment and support the cultural shift on campus.

Concluding Thoughts

Our hope is that the RARE model can serve as a guide for assessment practitioners as they encounter challenges in their work with a faculty or staff member when trying something new, improving a plan or process, or changing direction entirely. We found the process of coming together to discuss our common disciplinary background in counseling and ways of grounding our approach in theory and research to be refreshing and professionally validating. It can benefit fellow assessment professionals from other disciplinary paths to similarly share the impact of their own professional backgrounds, particularly as their specific knowledge

and training connects to implementing principles of effective assessment practice. We hope that by sharing our stories we can encourage the assessment community at large to uncover the mindset and processes embedded in their individual professional identities to further improve how we understand the collective identity of assessment professionals.

References

American Association of University Professors. (n.d.). The 1940 Statement of Principles on Academic Freedom and Tenure. Retrieved from https://www.aaup.org/report/1940-statement-principles-academic-freedom-and-tenure.

Ariovich, L., Bral, C., Gregg, P., Gulliford, M., & Morrow, J.A. (2018, May). The Assessment Profession in Higher Education: Addressing the Varied Professional Development Needs of Practitioners. Lexington, KY: Association for the Assessment of Learning in Higher Education and Watermark. Retrieved from https://go.watermarkinsights.com/rs/178-ONE-865/images/AALHE-Watermark-2018-Report-Final.pdf.

Arnold, K. (2014). Behind the mirror: Reflective listening and its tain in the work of Carl Rogers. *Humanistic Psychology, 42*, 354–369. American Psychological Association.

Combs, B., & Rose, T. (2016). Status of assessment staff in higher education: An overview. Paper presented at the annual Association for the Assessment in Higher Learning Conference, Milwaukee, WI.

de Shazer, S. (1985). *Keys to Solutions in Brief Therapy*. New York, NY: W. W. Norton.

Eubanks, D. (2017). *A guide for the perplexed. Intersection* (Fall), 4–12. Retrieved from https://c.ymcdn.com/sites/www.aalhe.org/resource/resmgr/docs/Int/AAHLE_Fall_2017_Intersection.pdf.

Green, D.A., & Little, D. (2016). Family portrait: A profile of educational developers around the world. *International Journal for Academic Development, 2*1(2), 135–150.

Green, D. A., & Little, D. (2013). Academic development on the margins. *Studies in Higher Education, 38*, 523–537.

Hansen, J. T. (2006). Counseling theories within a postmodernist epistemology: New roles for theories in counseling practice. *Journal of Counseling & Development, 84*, 291–297.

Hutchings, P., Ewell, P., & Banta, T. (2012). *AALHE Principles of Good Practice: Aging Nicely*. Retrieved from: www.learningoutcomesassessment.org/PrinciplesofAssessment.html.

Jankowski, N. A., & Slotnick, R. C. (2015). The five essential roles of assessment practitioners. *Journal of Assessment and Institutional Effectiveness, 5*(1), 78–100.

Katz, S. N. (2010). Beyond crude measurement and consumerism. *Academe, 96*(5). Retrieved from https://www.aaup.org/article/beyond-crude-measurement-and-consumerism#.XLuPsehKg2w.

Kinzie, J., Jankowski, N., & Provezis, S. (2014). Do good assessment practices measure up to the principles of assessment? *Assessment Update, 26*(3), 1–2, 14–16.

Little, D. (2014). Reflections on the state of the scholarship of educational development. *To Improve the Academy, 33*, 1–13.

Miller, W. R., & Rollnick, S. (1991). *Motivational interviewing: Preparing people to change addictive behavior*. New York, NY: Guilford Press.

Miller W. R., & Rollnick S. (2002) *Motivational interviewing: Preparing people for change* (2nd ed.). New York, NY: Guilford Press.

Murphy, J. J. (2008). *Solution-focused counseling in schools* (2nd ed.). Alexandria, VA: American Counseling Association.

Nicholas, M. C., & Slotnick, R. C. (2018, April). *A portrait of the assessment professional in the United States: Results from a national survey*. (Occasional Paper 34). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA). Retrieved from: http://learningoutcomesassessment.org/occasionalpaperthirtyfour.html.

Rogers, C. R. (1950). A current formulation of client-centered therapy. *Social Service Review, 24*(1), 442–449.

Rogers, C. R. (1951). *Client-centered therapy: Its current practice, implications and theory*. Boston, MA: Houghton Mifflin.

Rogers, C. R. (1952). A personal formulation of client-centered therapy. *Marriage and Family Living, 14*, 341–361.

Rogers, C. R. (1954). Becoming a person. Two lectures delivered on the Nellie Heldt lecture fund. *Oberlin College Nellie Heldt Lecture Series*. Oberlin, OH: Oberlin Printing Co.

Sanders, C. J. (2011). An exploration of knowledge and power in narrative, collaborative-based, postmodern therapies: A commentary. *The Professional Counselor, 1*(3), 201–207.

Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being*. New York, NY: Free Press.

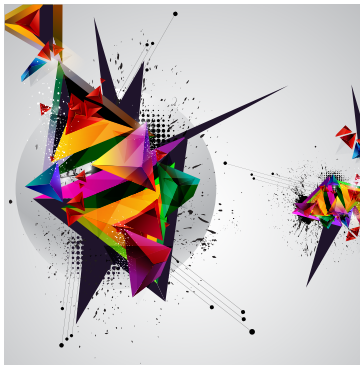
Sorcinelli, M. D., Austin, A. E., Eddy, P. L., & Beach, A. L. (2006). *Creating the future of faculty development: Learning from the past, understanding the present*. Bolton, MA: Anker.

Suskie, L. (2009). *Assessing student learning: A common sense approach*. (2nd ed.). San Francisco, CA: Jossey-Bass.

Worthen, M. (2018, Feb 23). *The misguided drive to measure learning outcomes*. New York Times. Retrieved from https://www.nytimes.com/2018/02/23/opinion/sunday/colleges-measure-learning-outcomes.html.

Abstract

Students at many higher education institutions are expected to acquire effective written communication skills and strong critical thinking skills. Although these student learning outcomes are common, these skills are complex. Assessing students' attainment of these outcomes at an institutional level can be difficult as many approaches to measuring student learning have inherent flaws or biases. This study uses direct and indirect measures of student learning to triangulate freshmen and seniors' attainment of written communication and critical thinking outcomes. The data collected through these complementary methods allowed institutional stakeholders to examine student learning through different lenses and to create a detailed and nuanced portrait of student learning within the institution.



AUTHORS

Elise Demeter, Ph.D.
UNC Charlotte

Christine Robinson, Ph.D.
UNC Charlotte

John G. Frederick, Ph.D.
Central Piedmont Community College

Holistically Assessing Critical Thinking and Written Communication Learning Outcomes with Direct and Indirect Measures

The ability to think critically and to write clearly are highly desirable skills in the workplace (Association of American Colleges & Universities (AAC&U), 2013). Higher education institutions also value these two domains, with 84% of institutions affiliated with the AAC&U reporting these skills are important outcomes for their undergraduates (AAC&U, 2016). However, in spite of this national recognition that critical thinking and effective written communication are important student learning outcomes, considerable diversity exists among higher education institutions in how students' attainment of these outcomes are assessed (e.g., Condon & Kelly-Riley, 2004; Douglass, Thomson & Zhao, 2012; Ennis, 1993; Haswell, 2000; Lui, 2011; Mazer, Hunt, & Kuznekoff, 2007; Moore, O'Neill, & Huot, 2009; Peach, Mukherjee, & Hornyak, 2007).

Differing cultures, needs, and student populations, along with the complexity of these higher-order skills, contribute to the lack of consensus about how to best assess critical thinking and written communication in higher education settings. For instance, some schools of thought have stressed the interconnection and positive relationship between students' writing skills and their critical thinking capabilities (e.g., National Center for Education Statistics, 1993, 1994, 1995), whereas other research points to the independence of these constructs (e.g., Condon & Kelly-Riley, 2004; Haswell, 1991). The range of assessment methods and measures may also make it difficult for stakeholders to understand how successful institutions are effectively teaching students critical thinking and written communication skills (e.g., Arum & Roksa, 2011; Council for Aid to Education, 2013).

Institutions are often motivated to use standardized assessments in order to provide employers and other external stakeholders with a snapshot of students' ability to critically think and to effectively communicate in writing. Standardized assessments are also useful because they allow for potential comparisons among institutions, and are typically rigorously vetted for validity and reliability issues. The pervasiveness of standardized testing in American

CORRESPONDENCE

Email

edemeter@uncc.edu

educational systems mean their results are easily communicated to and consumable by wide and lay audiences. Indeed, over a third to a half of AAC&U member institutions who assess critical thinking as a general education outcome report using a standardized exam (AAC&U, 2015; see also, Council for Aid to Education, 2016).

Although institutions may use standardized tests to assess students' critical thinking and communication abilities, this type of assessment also has many limitations (Banta, 2006). A primary critique of many standardized exams is that the test scope may or may not fit with an institution's definition of effective critical thinking or written communication. The format of a particular test may also privilege certain disciplines, populations of students, and ways of knowing or quantifying knowledge. Large-scale assessment strategies are also often low-stakes for students, meaning there is no significant reward or consequence for how well or poorly a student performs. This can lead to low-motivated test takers and underestimations of students' abilities (Lui, Bridgeman, & Adler, 2012). Faculty may also be resistant to the notion of reducing student learning to a single score.

Differing cultures, needs, and student populations, along with the complexity of these higher-order skills, contribute to the lack of consensus about how to best assess critical thinking and written communication.

The reductionist approach of using a standardized assessment to measure student learning is compounded by institutional assessment strategies that involve the collection of many disparate pieces of data but lack a comprehensive lens for synthesizing different measures of student learning. According to the Wabash National Study (Pascarella & Blaich, 2013), a comprehensive, multi-institution longitudinal study of the impact of academic and nonacademic collegiate experiences on liberal arts outcomes, the institutions studied typically collected good-quality data on student outcomes using multiple measures. However, these institutions often lacked a synthesis across various data collected, and lacked a focused communication strategy designed to deeply engage relevant constituencies in conversations about how to act upon data in order to improve student outcomes (Blaich & Wise, 2011). An advantage of the triangulation approach presented here is that it is less reductionist and allows for a sophisticated discussion of student attainment, which may help assessment professionals develop more holistic, and perhaps more actionable, narratives about student outcomes than narratives produced by a single measure.

Triangulation is a way of increasing confidence in the conclusions about assessment data by using multiple data sources, measures, perspectives or methodological approaches (Denzin, 1973; see also Coats & Stevenson (2006) for an alternative approach to triangulation in higher education settings). Often triangulation is used to overcome validity issues and to confirm results by limiting the biases that come from using a single source (Thomas, Lightcap, & Rosencranz, 2005; Ghrayeb, Damodaran, & Vohra, 2011). For instance, surveys can produce results that are not representative of the whole population due to sampling errors and issues with which individuals choose to respond to a given survey (Fowler, 2013). Responses on surveys may also be biased, or systematically different from the true scores, due to respondents' misremembering information or underreporting on certain issues. Although it is possible to mitigate some of these limitations when administering surveys, these limitations remain inherent to survey methodologies (Fowler, 2013). However, other methods may not share these specific intrinsic biases. Hence, if different methods all point to the same conclusion we can increase our confidence in that conclusion. Assessment experts advocate for the use of multiple methods (Banta 2002; Maki, 2002; Nelson, 2010; Springfield, Gwozdek, Peet, & Kerschbaum, 2012; Suskie, 2000); however, assessment practitioners may hesitate to utilize a triangulation approach because it is complex and time consuming (e.g., Guion, 2002).

This study used a direct measure (a standardized exam) and indirect measures (results from two student surveys) to gain a more holistic picture of student attainment of critical thinking and written communication learning outcomes. Direct measures are primary observations or examinations of student knowledge or skills, including results from exams, quizzes, and written assignments. Indirect measures are secondary observations or examination of student knowledge or skills, such as survey results of students' self-perceptions of their learning (e.g., Rogers, 2006). Our choice of exam and surveys was largely driven by what data was available on our campus. These measures were not only used in a confirmatory way to show agreement or disagreement with each other but also in parallel with each other to better inform what the results mean within this institutional context. To preview our results, we

found evidence for student attainment of written communication outcomes, but our evidence on students' attainment of critical thinking outcomes was mixed. However, the complementary information provided by the different methodologies and data sources used provided insight for how our institution may benefit from a focused plan to modify instructional strategies for teaching particular critical thinking subskills.

Methods

General Procedures

Our undergraduate participants were freshmen and seniors enrolled at a large, public four-year university in the southern United States. The focus of our analyses was on students from a single college (referred to throughout text as 'College A') located within the university. College A was chosen as a focus for this institutional assessment investigation by design so that the study findings had a clear constituency. All measures were administered in the same spring semester. Except where specifically indicated otherwise, participants assessed were independent samples on each measure.

This study used a direct measure (a standardized exam) and indirect measures (results from two student surveys) to gain a more holistic picture.

For the ETS HEIghten standardized exams used, participants were drawn from core courses within College A. Freshmen and senior participants were randomly assigned to complete one of the two ETS HEIghten exams used for this investigation. A description of the exams used and the participants is provided in the "Measures and Participants" section. Participants completed their assigned exam during a typical course meeting time in a campus computer lab. Faculty were involved in the planning and logistics of administering these assessments. The ETS HEIghten exams were proctored and administered via a secure ETS testing browser. The ETS HEIghten exams were administered as a low-stakes assessment and our student participants did not receive course credit for their participation. A random selection of <15% of participants were provided with a gift card of a nominal monetary value after completing the HEIghten testing. Students completing the survey measures came from across the university, with students in College A identified based on their selected major or pre-major at the time the study was conducted.

For correlational analyses, Pearson's *r* correlational tests were conducted when both of the variables analyzed were continuous and normally distributed; Spearman's nonparametric *r* (sometimes denoted as *rho*) tests were applied to analyses using survey response data on a Likert scale. For independent-samples t-tests, equal variances were not assumed if the Levene's Test for equality of variances was significant. Degrees of freedom are rounded to the nearest integer for readability. Effect sizes are provided as Hedges' *g*, a variant of Cohen's *d* that corrects for unequal group sample sizes.

Defining effective written communication and critical thinking

Effective written communication and critical thinking are included among our institution's student learning outcomes, as well as in student learning outcomes specific to the academic programs within College A. Our institutional definition of written communication includes: the effective development of written content, presented in a structured exposition that conveys and creates meaning consistent with the conventions appropriate for a given communicative context. Our associated institution-level student learning outcome states: Students will be able to demonstrate the ability to adapt and apply a variety of writing strategies (invention, research, analyses, organization, and revision) to communicate effectively with a target audience. Institutionally, we define critical thinking as including creative thinking, innovation, inquiry and analysis, evaluation and syntheses of information. Our associated outcome states: Students will be able to demonstrate critical thinking skills through a process of inquiry that explores evidence for developing innovative and creative solutions to make informed decisions and evaluations. Programs within College A include student learning outcomes that align to these institutional outcomes, as well as critical thinking-related outcomes pertinent to the College's discipline such as using numerical analysis skills to draw appropriate conclusions.

For this investigation, we operationally defined the subskills that comprise effective written communication as follows: knowledge of rhetorical conventions in different contexts; awareness of audience and purpose of writing; appropriate content development and organization; knowledge of language use and conventions including appropriate spelling, grammar, tone, style; and knowledge of the writing process (i.e., drafting, revising). Our operational definition of critical thinking was a process that encompassed some or all of the following steps: a process of inquiry and hypothesis-generation; gathering information and data of good quality; evaluation of information's credibility, validity, reliability and logical strength; analysis of quantitative and qualitative information and data; interpretation of information's significance and meaning; drawing inferences from evidence, assessing alternatives, determining sufficiency of evidence, and evidenced-based decision-making; and communication of one's thought process and conclusions to others (see also Facione, 1990; Rhodes, 2010). For this study we did not seek assessment measures that would encompass all aspects of our definitions. Instead, our measures typically aligned to specific aspects of effective written communication or critical thinking, and we constrained our conclusions drawn accordingly.

Measures & Participants

ETS HEIghten Written Communication Assessment. Freshmen (46 women, 69 men, age *M* ± S.E.: 19.4 ± 0.1 years) and seniors (46 women, 52 men, age 22.6 ± 1.4 years) sat for the ETS HEIghten Written Communication Assessment (Rios, Sparks, Zhang, & Liu, 2017; Sparks, Song, Brantley, & Liu, 2014). This is a two-part test consisting of a constructed persuasive essay and multiple-choice questions based on presented reading passages. This exam assesses students' knowledge of rhetorical conventions in different contexts and ability to identify writing for certain purposes and audiences; students' knowledge of conceptual strategies including how to develop ideas in an organized, logical and coherent sequence; knowledge of language use and conventions; and knowledge of the writing process. The test is scored on a scale of 150 to 180.

ETS HEIghten Critical Thinking Assessment. Freshmen (38 women, 65 men, *M* ± S.E. age: 19.8 ± 0.1 years) and seniors (53 women, 51 men, age 22.9 ± 0.2 years) sat for the ETS HEIghten Critical Thinking Assessment (Liu, Frankel, & Roohr, 2014; Liu, Mao, Frankel & Xu, 2016). This multiple-choice exam is a test of logical and deductive reasoning, and it addresses evaluation, interpretation, and inference skills. For example, it provides students with lists of evidence and asks questions about whether or not pieces of evidence support a given conclusion, and how strongly. It also asks test takers to identify assumptions in provided written arguments and to solve logic word problems. It is scored on a scale of 150 to 180.

We define critical thinking as including creative thinking, innovation, inquiry and analysis, evaluation and syntheses of information.

National Survey of Student Engagement (NSSE). To complement the standardized test data generated by the ETS HEIghten Exams, we examined students' perceptions of their writing and critical thinking skills and experience using items from the National Survey of Student Engagement (NSSE; http://nsse.indiana.edu/). The NSSE is a large-scale, multi-institutional survey typically administered by participating institutions to first-year and senior bachelor's degree-seeking students. It has undergone extensive psychometric testing for validity and reliability (psychometric profile available at http://nsse.indiana.edu/html/psychometric_portfolio.cfm). Broadly, the survey asks students about their engagement in educationally purposeful activities. At our institution, this survey was electronically administered using stratified random sampling to freshmen (19% response rate, freshmen survey population: 332 women, 235 men, *M* ± S.E. age: 18.4 ± 0.1 years; College A subpopulation: 20 women, 53 men, age: 18.3 ± 0.1 years) and seniors (20% response rate, senior survey population: 447 women, 261 men, *M* ± S.E. age: 24.8 ± 0.3 years; College A subpopulation: 53 women, 40 men, age: 23.7 ± 0.6 years). Survey respondents received no incentives. Specific items used are detailed in Table 1. For analyses using the NSSE data students in College A were generally compared to their peers at our institution who are not enrolled in College A.

Senior Survey. The Senior Survey is electronically distributed to seniors with their graduation application materials. This survey was developed for collective use by the public universities within our institution's state system and has been routinely used by our institution

for several years. Participation is voluntary and not incentivized, but generally 60% of our graduating seniors respond to this survey. Data from 490 College A seniors (210 women, 280 men, M ± S.E. age: 23.8 ± 0.2 years) and 2,070 seniors (1136 women, 934 men, age 24.9 ± 0.1 years) affiliated with colleges at our university other than College A were used for these analyses. This survey asks students questions about their satisfaction with their instructors and their educational experiences as well as their opinions about how much the institution has contributed to their knowledge, skills, and development in various areas including written communication and critical thinking.

Results

Evidence for attainment of effective written communication capabilities

The evidence examined suggests students in College A are developing effective written communication skills, although they may self-perceive less gains in writing than their peers in other colleges at our institution. As the ETS HEIghten exam suite is relatively new to market and this represents our institution's first use of the exams, as a basic validity check we first examined how scores on the HEIghten Written Communication exam correlated with students' cumulative grade point average (GPA) on a 4.0 scale. Across all participants who sat for the exam, we found that exam scores showed a small but significant positive correlation with students' GPA, with students with higher exam scores also having higher GPAs (Pearson's r correlational analysis, $r(211) = .22$, $p = .001$). This finding held for freshmen ($r(113) = .36$, $p < .001$), but was not significant for seniors alone ($r(96) = .17$, $p = .07$). We then examined how seniors' scores on the ETS HEIghten exam compared to those of freshmen. Seniors scored significantly higher on the exam than freshmen (Figure 1; mean scores ± standard error for freshmen: 162.5 ± 0.5, seniors: 164.1 ± 0.6; independent samples t-test: $t(211) = 2.11$, $p = .04$, *Hedges' g* = 0.3), suggesting students in College A are making gains in their writing abilities and in their knowledge of the writing process.

[Students] are developing effective written communication skills.

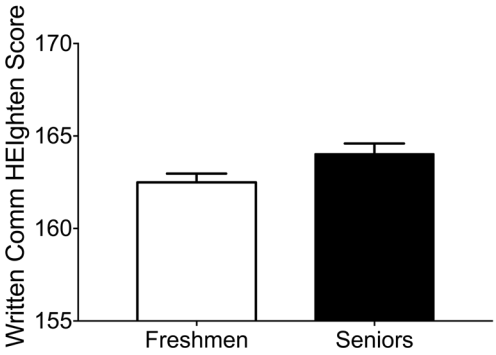


Figure 1. Seniors score higher on the HEIghten Written Communications exam than freshman. Bars represent the mean score on the ETS HEIghten Written Communications exam for freshman (white bar) and seniors (black bar). Error bars represent the standard error around the mean.

However, our survey data suggest students in College A self-perceive less gains in writing than their peers and that they may be producing fewer pages of academic writing during their senior year than their peers. College A seniors and seniors affiliated with other colleges within the university both report on average that their college education positively contributed to their written communication skills (Senior Survey item, 'To what extent do you think your college education contributed to your knowledge, skills, and personal development in writing effectively?'; response scale: very much (4), somewhat (3), very little (2) not at all (1)). However, students in College A give lower ratings than their peers (mean rating ± standard error for College A seniors: 3.25 ± 0.03; seniors from other colleges: 3.46 ± 0.02; independent-samples t-test $t(2537) = 6.13$, $p < .001$, *Hedges' g* = 0.2). One possibility of why students in College A may differ in their self-perceptions about writing may be because these students are assigned fewer pages of academic writing than their peers. According to

the NSSE survey data on how many pages of writing students were assigned in the current school year, seniors in College A reported they were assigned about 18 fewer pages of writing than their peers affiliated with other colleges (College A: 62.1 ± 6.1 pages, other colleges: 80.8 ± 4.2 pages; independent samples *t*-test, $t(144) = 2.53$, $p = .01$, *Hedges' g* = 0.2). This difference in pages produced does not mean College A students were assigned fewer writing assignments than students in other colleges; but, it is possible students who produce more pages of course writing are also more willing to attribute self-perceived gains in writing to their experiences in those courses.

Mixed evidence for attainment of critical thinking capabilities

The examined evidence was mixed on whether College A students are gaining strong critical thinking skills during their undergraduate education. As with the Written Communication exam, we first examined whether students' scores on the ETS HEIghten Critical Thinking exam correlated with students' GPAs. Students' Critical Thinking exam scores indeed positively correlated with their cumulative GPA, with students with higher exam scores also exhibiting higher GPAs (Pearson's *r* correlational analysis, $r(207) = .27$, $p < .001$; freshmen: $r(101) = .23$, $p = .02$, seniors: $r(102) = 0.31$, $p < .01$). Next, we examined how seniors' scores compared to those of freshmen. We found seniors' scores increased but not significantly compared to freshmen (mean scores ± standard error for freshmen: 162.3 ± 0.6, seniors: 163.2 ± 0.6; independent samples *t*-test, $t(206) = 1.07$, $p = .31$, *Hedges' g* = .1).

[students] may differ in their self-perceptions about writing…because these students are assigned fewer pages of academic writing than their peers.

The Senior Survey and NSSE data proved useful for both interpreting the ETS HEIghten results and for more holistically examining the subskills that comprised our definition of critical thinking. In response to the Senior Survey item, 'To what extent do you think your college education contributed to your knowledge, skills, and personal development in using critical thinking skills?' (response scale: very much (4), somewhat (3); very little (2) not at all (1)), although all students rated this item fairly high, seniors in College A gave lower ratings than seniors affiliated with other colleges at our university (mean rating ± standard error for College A seniors: 3.53 ± 0.03; seniors from other colleges: 3.66 ± 0.01; independent-samples *t*-test, equal variances not assumed, $t(703) = 4.05$, $p < 0.001$, *Hedges' g* = 0.4). Compared to students affiliated with other colleges at our institution, students in College A reported that their courses placed more emphasis on lower-level skills like memorization and less emphasis on higher-level skills like analyzing an idea and forming new ideas (NSSE items in Table 1; independent-samples *t*-tests, all $t > 2.07$, $p < .04$, *Hedges' g* > .2). These results also held if only data from seniors was examined. From a triangulation perspective the lower ratings for the higher-order skills of analyzing and forming ideas are logically consistent with the findings from the ETS HEIghten Critical Thinking exam, given the exam's focus on interpretation and inference skills.

The large sample size of the Senior Survey enabled us to identify a subgroup of seniors (n = 80) who sat for the ETS HEIghten Critical Thinking exam and who completed the Senior Survey. We conducted exploratory analyses on this subgroup to see if there were any correlational relationships between students' exam scores and students' responses to the Senior Survey items. Students' exam scores did not correlate with the broad item on the Senior Survey about the extent to which their college education contributed to their using critical thinking skills (Spearman's nonparametric r, r(79) = .002, p = .99). However, students' exam scores did positively correlate with three survey items related to effective teaching practices: 'Overall satisfaction with instructors in my major department's:' (a) 'Ability to motivate me to do my best', (b) 'How quickly they provide feedback on my work', and (c) 'Effectiveness in using instructional technology' (Figure 2; response scale for all items: very satisfied (5), satisfied (4), neither (3), dissatisfied (2), very dissatisfied (1); Spearman's nonparametric r, all r > .23, p < .05). For each of these items, students with higher levels of satisfaction also tended to have higher Critical Thinking exam scores. Students' exam scores did not correlate with any other item on the Senior Survey.

The survey data also provided insight into critical thinking subskills not covered by the ETS HEIghten Critical Thinking exam. Specifically, our definition of critical thinking included subskills related to the analysis of information and data. According to the NSSE data, students in College A perceive greater gains in analyzing data than their peers in other colleges (Table 1; independent samples *t*-test, equal variances not assumed, $t(170) = 3.00$, $p = .003$, Hedges' $g = .3$). Similarly, students in College A also reported engaging in quantitative reasoning skills more frequently than their peers (Table 1; independent samples *t*-tests, all $t > 3.02$, $p < .01$, *Hedges' g* > 0.2). Again, these results held if only data from seniors was examined. These data suggest College A students self-perceived gains in their ability and experience with analysis and quantitative reasoning skills.

Table 1
National Survey of Student Engagement (NSSE) items used for critical thinking analyses.

Item Stem, Item Text, and Response Scale	College A Students (*M* ± *SEM*)	Other Students[a] (*M* ± *SEM*)
During the current school year, how much has your coursework emphasized the following:		
Memorizing course material	3.17 ± 0.06	2.93 ± 0.03
Analyzing an idea, experience or line of reasoning in depth by examining its parts	2.90 ± 0.07	3.04 ± 0.02
Forming a new idea or understanding from various pieces of information	2.74 ± 0.08	2.92 ± 0.03
Very often (4), Often (3), Sometimes (2), Never (1)		
How much has your experience at this institution contributed to your knowledge, skills, and personal development in the following areas?		
Analyzing numerical and statistical information	2.96 ± 0.08	2.71 ± 0.03
Very much (4), Quite a bit (3), Some (2), Very little (1)		
During the current school year, about how often have you done the following?		
Reached conclusions based on your own analysis of numerical information (numbers, graphs, statistics etc.)	2.74 ± 0.07	2.56 ± 0.03
Used numerical information to examine a real-world problem or issue (unemployment, climate change, public health, etc.)	2.53 ± 0.07	2.29 ± 0.03
Evaluated what others have concluded from numerical information	2.48 ± 0.07	2.28 ± 0.03
Very often (4), Often (3), Sometimes (2), Never (1)		

Note. M = mean; SEM = standard error around the mean

[a]Other students are students affiliated with colleges at the university other than College A.

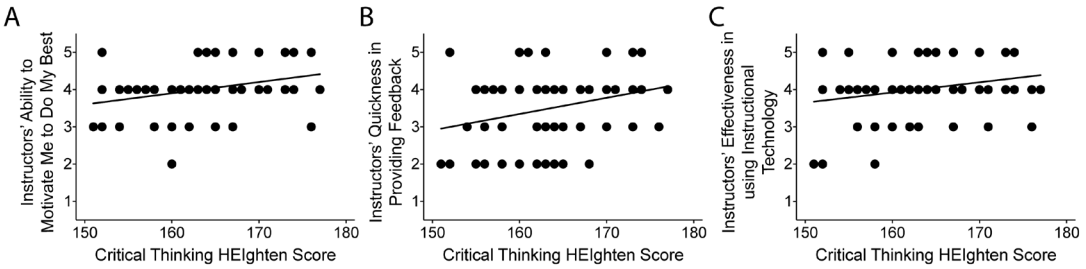


Figure 2. HEIghten Critical Thinking scores correlate with students' satisfaction with instructors' use of effective teaching practices. Each dot on the scatterplots indicates an individual College A senior's score on the HEIghten Critical Thinking exam versus the student's response to an item on the Senior Survey about students' level of satisfaction with (a) instructors' ability to motivate the student to do their best, (b) instructors' quickness in providing feedback, and (c) instructors' effectiveness in using instructional technology. Students rated the survey items on a scale of very satisfied (5) to very dissatisfied (1). Line on scatterplots represents the linear regression line for each plot. For each of these three survey items on effective teaching practices, students' level of satisfaction with their instructors' use of these practices positively correlated with students' scores on the critical thinking exam. These findings provide indirect evidence to support the idea that effective teaching practices can foster the development of stronger critical thinking skills.

Discussion

This study brought together data from direct measures (standardized exams) as well as indirect measures (student surveys) to assess students' attainment of written communication and critical thinking outcomes, with a specific focus on how seniors compared to freshmen. The findings from this assessment work were written up in an executive summary and were orally presented to the faculty and senior leadership in the college studied.

Students with higher levels of satisfaction also tended to have higher Critical Thinking exam scores.

In the domain of written communication we found evidence that seniors showed more advanced writing abilities and knowledge than freshmen, an encouraging result for the college studied. Seniors also self-reported gains in writing abilities and knowledge, although to a slightly lower degree than seniors affiliated with other colleges within the institution. Students in the college studied also reported being assigned fewer pages of academic writing during their senior year than our other seniors within the institution. When these data were presented to faculty and leadership within the college, this data point sparked a discussion about how different disciplines may favor assignments of different lengths (e.g., public policy memos stress concise communication). As a follow-up to this study, it would be interesting to gain further details about how often seniors are writing for academic purposes and to better understand what disciplinary writing may 'count' in students' opinions as academic writing.

For critical thinking, students in the college studied self-reported their coursework placed more emphasis on memorization and less emphasis on skills that may contribute to critical thinking, such as logical reasoning, analyzing ideas based on examining its parts, and forming new ideas and understandings based on various pieces of information. This finding was particularly interesting considering it logically supported the conclusion drawn from the standardized exam data. On our standardized exam measure of logical reasoning, evaluation, and inference skills, our seniors did not score significantly higher than freshmen. These findings spurred intense faculty discussion over instructional practices, including how existing course activities and assessments might be revised to allow students to engage in evaluation and inference skills. More broadly, the college and the institution recently infused additional critical thinking teaching and learning opportunities in the undergraduate curriculum. Core courses within the college's curriculum have been rigorously revamped to focus on analysis and evaluation skills, and our institution has added an inquiry-based freshmen critical thinking seminar to our general education curriculum. The seniors tested here have not experienced the sum of these curricular changes, but future assessment findings may reveal evidence indicating current and future students are reaping the benefits of these changes.

The NSSE and Senior Survey data examined highlighted a strength for students in the college studied in the area of numerical analysis skills. The college studied is one that devotes considerable instructional time and attention in its curriculum to fostering students' quantitative analytic skills, and students in this college self-perceive higher gains in analyzing numerical information than students affiliated with other colleges. Quantitative reasoning and numerical analysis skills are outside the scope of the critical thinking standardized exam used here but do fit within our institutional definition of critical thinking. In this case using different data sources and measures was an asset because it allowed us to see a broader picture of how students perceive they are attaining different critical thinking subskills.

The correlational findings between students' critical thinking standardized test scores and those students' satisfaction rankings regarding certain instructional practices are intriguing because they suggest pedagogical strategies and instructional conditions that may foster students' critical thinking development (Tsui, 2016). For example, motivation is an important precursor for engagement of higher-order, cognitively demanding processes like engaging in critical thinking (Miel & Wigfield, 2014). Instructors adept at motivating students may also engender the best conditions for fostering growth in critical thinking. Timely, accurate feedback on performance is a key characteristic of deliberate practice, and deliberate practice of critical thinking skills is important for developing mastery (van Gelder, 2005). The interpretation of the correlation between satisfaction with instructors' effectiveness in using instructional technology and critical thinking skills is less straightforward. However, it may be that 'effective use of instructional technology' is a proxy for how effective instructors are at engaging students. Student engagement is likely critical for motivating students and creating conditions where students will be encouraged to stretch and grow their cognitive abilities, including critical thinking. These findings point to potentially fruitful areas for future empirical research on how pedagogical structures and instructional conditions can enhance students' critical thinking skills.

The triangulation approach was particularly useful for the discussion with faculty and senior leadership about this study's findings as it allowed us to evaluate students' attainment of outcomes through various lenses. Although standardized exams are commonly used to apply for admission into our university, our faculty largely saw standardized exams for assessing student learning as tools to be used to satisfy accountability purposes more than as tools to provide insightful data that could be used to improve teaching and learning. Bringing in students' perspectives, as well as discussing how the data presented did or did not match with the college's internal data on student performance, helped to engage the faculty in a discussion about how closely our students' outcomes are matching our objectives for their learning.

As the ETS HEIghten exams are more widely adopted we foresee additional possibilities for comparing exam results to results from other peer or aspirational peer institutions. At the time of our test administration the institutional comparisons available on the particular exams used here were rather limited. In some ways this limitation can be a benefit, as it strongly encourages institutions to create internal comparison groups of interest in order to contextualize the results. Though large-scale assessments may appease external stakeholders, our experiences also reinforced the best practice of starting these types of investigations with a clear question about student learning that the measures used will address. Similarly, institution-level assessment plans may have some appeal but we have found results are more actionable when situated within a particular college or academic program (Robinson, Sanders, Hobbs, Demeter, Singer-Freeman, 2019). Although the ETS HEIghten exams were too new to market at the time of our administration to conduct a longitudinal study, future options may also include tracking individual students' growth between freshmen and senior years.

Our experiences also reinforced the best practice of starting these types of investigations with a clear question about student learning.

Effective written communication and strong critical thinking skills will continue to be highly valued both within academia and within the workplace. Triangulating data from different methods and sources allowed us to develop a more holistic assessment of students' attainment of these two important outcomes. Although using multiple measures brings a complexity to assessment work, it also creates the opportunity to create more sophisticated narratives of student achievement.

References

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.

Association of American Colleges and Universities. (2013). *It takes more than a major: Employer priorities for college learning and student success*. Washington, DC: Association of American Colleges and Universities and Hart Research Associates.

Association of American Colleges and Universities. (2016). *Recent trends in general education design, learning outcomes, and teaching approaches*. Washington, DC: Association of American Colleges and Universities and Hart Research Associates.

Banta, T. (2006). Reliving the history of large-scale assessment in higher education. *Assessment Update, 18*(4), 3–4, 15.

Banta, T. (2002). *Building a scholarship of assessment*. San Francisco: Jossey Bass, pp. 262–3.

Blaich, C.F., & Wise, K.S. (2011, January). *From gathering to using assessment results: Lessons from the Wabash National Study*. (NILOA Occasional Paper No. 8). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Council for Aid to Education. (2013). *Does college matter? Measuring critical-thinking outcomes using the CLA*. Retrieved from https://cae.org/images/uploads/pdf/Does_College_Matter.pdf.

Council for Aid to Education. (2016). *CLA+ National Results, 2015–2016*. Retrieved from https://cae.org/images/uploads/pdf/CLA_National_Results_2015-16.pdf.

Coats, M., & Stevenson, A. (2006). Towards outcomes-based assessment: An unfinished story of triangulation and transformation. Paper presented at the Association for the Study of Evaluation and Assessment in Education in South Africa Conference, Johannesburg, South Africa.

Condon, W., & Kelly-Riley, D. (2004) Assessing and teaching what we value: The relationship between college-level writing and critical thinking abilities. *Assessing Writing, 9*(1), 56–75.

Denzin, Norman K. (1973). *The research act: A theoretical introduction to sociological methods*. Piscataway, NJ: Transaction Publishers.

Douglass, J.A., Thomson, G., & Zhao, CM. (2012). The learning outcomes race: the value of self-reported gains in large research universities. *Higher Education, 64*(3), 317–335.

Ennis, R.H. (1993). Critical thinking assessment. *Theory Into Practice, 32*(3), 179–186.

Facione (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction—The Delphi report*. Millbrae, CA: California Academic Press.

Fowler, F. J. (2013). *Survey Research Methods* (5th ed.) Thousand Oaks, CA: SAGE Publications.

Ghrayeb, O., Damodaran, P., & Vohra, P. (2011). Art of triangulation: an effective assessment validation strategy. Global *Journal of Engineering Education, 13*(3), 96–101.

Guion, L. A. (2002). Triangulation: Establishing the Validity of Qualitative Studies. *Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences*, FCS6014, 1–3.

Haswell, R. H. (2000). Documenting improvement in college writing: A longitudinal approach. *Written communication, 17*(3), 307–352.

Haswell, R. H. (1991). *Gaining ground in college writing: Tales of development and interpretation*. Dallas, TX: Southern Methodist University Press.

Liu, O.L. (2011). Value-added assessment in higher education: a comparison of two methods. *Higher Education, 61*(4), 445–461.

Liu, O.L., Bridgeman, B., & Adler, R.M. (2012). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher, 41*(9), 352–362.

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing critical thinking in higher education: current state and directions for next generation assessment. *ETS Research Reports Series*, 1–23.

Liu, O.L., Mao, L., Frankel, L., & Xu, J. (2016). Assessing critical thinking in higher education: the HEIghten™ approach and preliminary validity evidence. *Assessment & Evaluation in Higher Education, 41*(5), 677–694.

Maki, P. (2002). Using multiple assessment methods to explore student learning and development inside and outside of the classroom. *NASPA's NetResults*. Retrieved from https://www.apu.edu/live_data/files/333/multiple_assessment_methods_to_explore_student_learning_and_deve.pdf.

Mazer, J. P. & Hunt, S. K., & Kuznekoff, J. H. (2007). Revising general education: Assessing a critical thinking instructional model in the basic communication course. *The Journal of General Education 56*(3), 173–199.

Miele, D., & Wigfield, A. (2014). Quantitative and qualitative relations between motivation and critical-analytic thinking. *Educational Psychology Review, 26*(4), 519–541.

Moore, C., O'Neill, P., & Huot, B. (2009). Creating a culture of assessment in writing programs and beyond. *College Composition and Communication, 61*(1), 107–132.

National Center for Education Statistics. (1993). *National assessment of college student learning: Getting started*. Washington, DC: OERI.

National Center for Education Statistics. (1994). *National assessment of college student learning: Identification of the skill to be taught, learned and assessed*. Washington, DC: OERI.

National Center for Education Statistics. (1995). *National assessment of college student learning: Identifying college graduates' essential skills in writing, speech and listening, and critical thinking*. Washington, DC: OERI.

Nelson, K.H. (2010). Triangulation in assessment of student learning outcomes. Paper presented at the Southern Association of Colleges and Schools Commission on Colleges Annual Meeting, Louisville, KY.

Peach, B.E., Mukherjee, A., & Hornyak, M. (2007). Assessing critical thinking: A college's journey and lessons learned. *Journal of Education for Business, 82*(6), 313–320.

Pascarella, E.T., & Blaich, C. (2013). Lessons from the Wabash National Study of liberal arts education. *Change: The Magazine of Higher Learning, 45*(2), 6–15.

Rhodes, T. (2010). Assessing outcomes and improving achievement: Tips and tools for using rubrics. Washington, DC: Association of American Colleges and Universities.

Robinson, C., Sanders, C. Hobbs, H., Demeter, E. & Singer-Freeman, K. (2019). University of North Carolina at Charlotte: Using the EIA application as a catalyst for intentional improvement. *Assessment Update, 31*(1), 2-16.

Rogers, G. (2006). Direct and indirect assessment. Assessment 101: Assessment Tips. Retrieved from https://www.abet.org/wp-content/uploads/2015/04/direct-and-indirect-assessment.pdf.

Rios, J. A., Sparks, J. R., Zhang, M., and Liu, O. L. (2017). Development and validation of the written communication assessment of the *HEIghten*® Outcomes Assessment Suite. ETS Research Report Series, 1–16.

Sparks, J. R., Song, Y., Brantley, W., and Liu, O. L. (2014). Assessing written communication in higher education: Review and recommendations for next generation assessment. *ETS Research Report Series*, 1–52.

Springfield, E., Gwozdek, A.E., Peet, M., & Kerschbaum, W.E. (2012). Using multiple methods to assess learning and outcomes in an online degree-granting dental hygiene program. *Journal of Dental Education, 76*(4), 414–26.

Suskie, L. (2000). Fair assessment practices: Giving students equitable opportunities to demonstrate learning. *AAHE Bulletin, 52*(9), 7–9.

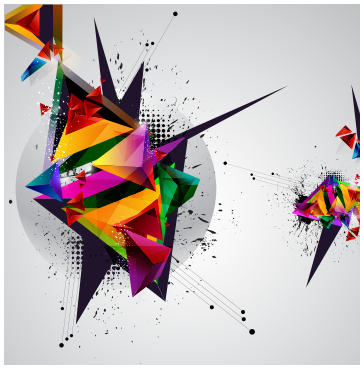
Thomas, B., Lightcap, T., & Rosencranz, L. (2005). Taming the hydra: A triangulation approach to assessing an interdisciplinary core curriculum. [Online]

Tsui, L. (2016). Fostering critical thinking through effective pedagogy: Evidence from four institutional case studies. *Journal of Higher Education, 73*(6), 740–763.

van Gelder, T. (2005). Teaching critical thinking: Some lessons From cognitive science, *College Teaching, 53*(1), 41–48.

Abstract

Many institutions implement assessment teams as resources to develop faculty and staff knowledge of and confidence in programmatic assessment processes. Additional resources may include rubrics or peer review and feedback, but effectiveness of these resources is rarely evaluated. Programmatic assessment allows institutions to examine the impact of multiple resources to determine which positively impact assessment outcomes. This quantitative study examined administrative and student affairs units' perceptions of assessment resources supported by the Office of Institutional Effectiveness (OIE) at one large public, southeastern university. The findings corroborated the positive impact of rubrics and peer review and feedback, providing a basis for continued support of many of the institution's existing resources. Although institutions cannot control the utilization of resources available, they can be more certain that the resources provided are beneficial to those who seek them.



AUTHORS

Cynthia D. Groover, Ed.D
Georgia Southern University

Juliann Sergi McBrayer, Ed.D
Georgia Southern University

Richard Cleveland, Ph.D.
Georgia Southern University

Amy Jo Riggs, Ph.D.
Georgia Southern University

Utilization And Perceived Utility of Institutional Administrative And Student Affairs Assessment Resources

Regional accreditation is the mechanism through which many institutions account for the quality of the education provided to their students as well as the quality of the environment within which this education is provided. With the growing move toward accountability in higher education (Martin, Goulet, Martin, & Owens, 2015), institutions have found themselves facing more rigorous assessment demands from their regional accreditors (Eaton, 2013). Without regional accreditation, institutions are unable to offer federal financial aid, the primary funding source for many students. Nor can their students graduate with degrees from programs holding disciplinary accreditation, which is a must for many employers. Given the trend toward increased accountability both during and after students' time on campus, investigating the quality of a comprehensive institutional assessment process is vital to both student and institutional success. This challenge can only be met by institutions being actively and effectively engaged in the assessment process.

Considering the breadth of assessment being conducted across institutions, effectively promoting and sustaining institutional assessment processes can be overwhelming for those officially charged with the tasks. The number of faculty and staff in need of training and support in this critical institutional function is often disproportionately large compared to the number of assessment professionals available. In response, many assessment offices have implemented assessment teams to assist both faculty and administrative and student affairs units across campuses in promoting and sustaining effective assessment processes (Fishman 2017; Krzykowski & Kinser, 2014; Slager & Oaks, 2013).

This study sought to better understand participants' perceptions of their own knowledge of and confidence in the assessment process. Specifically, this study examined how those perceptions are impacted by the peer review process facilitated by an Institutional Effectiveness (IE) Review Team and by other specific resources supported by

CORRESPONDENCE

Email

cgroover@georgiasouthern.edu.

an Office of Institutional Effectiveness (OIE), such as consultation and website materials. The research questions that guided this study were as follows: (a) What are the perceived strengths and weaknesses of the resources in place to develop knowledge of and confidence in the assessment process? (b) How does perceived utility differ among divisions of the institution? (c) How do participants perceive their own knowledge of and confidence in the assessment process? and (d) What is the relationship between knowledge of and confidence in the assessment process and the utility of specific resources in place?

Because research in the field of assessment has been lacking in terms of data-driven processes to assess the effectiveness of institutional assessment practices, particularly related to administrative and student affairs units, a gap in the literature exists and thus, further research was warranted.

Background

Common resources used to promote effective institutional assessment processes include the use of rubrics and peer review and feedback (Fulcher, Coleman, & Sundre, 2016; Jonsson, 2013; Kahlon, Delgado-Angulo, & Bernabé, 2015; Panadero & Romero, 2014). Assessment teams often apply institutional rubrics to annual assessment reports to supplement quantitative evaluation with qualitative feedback. Apart from this annual process, assessment offices may provide additional resources, such as consultation opportunities or website materials. However, any relationship between these resources, assessment teams, and successful assessment processes, "is only speculative until systematically evaluated" (Fulcher & Bashkov, 2012, p. 7). Assessment offices and the review teams devote significant effort in applying rubrics, providing feedback, and developing support materials. Impact of these efforts is difficult to gauge, but programmatic evaluation allows institutions to look at the impact of a multitude of practices to determine if they have the most appropriate resources in place to positively impact assessment processes across campus (Fink, 2013). Any programmatic assessment process "should continue to undergo evaluation where it can be modified to ensure that every element contributes to the program's outcomes" (Shutt, Garrett, Lynch, & Dean, 2012, p. 78).

Investigating the quality of a comprehensive institutional assessment process is vital to both student and institutional success.

This focus on specific assessment resources is important because often institutions focus their assessment on participant satisfaction instead of the impact of specific resources on assessment outcomes (Chalmers & Gardiner, 2015). For example, Meyer and Murrell (2014) examined how a variety of institutions evaluated their faculty development programs in online learning and found that 95% of responding institutions focused outcome measures on faculty satisfaction with the training, and 90% focused outcome measures on faculty perception of the usefulness of the training. A more effective approach may be to collect data addressing the frequency with which participants consult specific resources provided and apply the skills learned, as well as their reasons for not using the specific resources provided or applying the skills taught (Yarber et al., 2015). Collecting data specific to the utility and application of specific resources could allow program developers to address more systematically any weaknesses or shortcomings participants reveal.

Methods

Research Design

The purpose of this nonexperimental quantitative study was two-fold. First, the researchers sought to better understand participants' perceptions of their own knowledge of and confidence in the assessment process. Second, this study identified perceived strengths and weaknesses of existing resources to determine their utility. In doing so, the researchers intended to go beyond anecdotal findings and examine a model being implemented at one large public, southeastern university. Specifically, this study examined how participants' perceptions are impacted by the peer review process facilitated by the IE Review Team and by other specific resources supported by the OIE. This study examined the "process of interaction" between IE Review Team members and administrative and student affairs units, relying on the participants' views of the process to construct a clearer picture of perceived strengths and weaknesses of the resources in place (Creswell, 2014, p. 8).

Participants

Researchers used saturation sampling to survey all administrative and student affairs unit administrators, assessment coordinators, and staff who were responsible for or had contributed to the preparation of their units' annual assessment reports or plans during any of the six previous assessment cycles. Total study population was 85, and of the 85 surveyed, 61 participants provided data, yielding a response rate of 72%.

Data Collection

This study relied on data collected by the OIE through an anonymous electronic survey, modified, with permission from the original authors, and adapted to accurately reflect the resources specific to the research university (Rodgers, Grays, Fulcher, & Jurich, 2013). The complete survey instrument is included in Appendix A. Administered at the conclusion of a yearly assessment cycle, the survey addressed two main areas, *Use of Assessment Resources* and *Assessment Environment*. Each item in the Use of Assessment Resources section described a unique resource available to administrative and student affairs units, such as face-to-face feedback from an IE Review Team member or general information on the OIE website. Likert-scaled responses included: *I did not know about this resource*; *I knew about this resource but did not use it*; *This resource was not at all helpful*; *This resource was a little helpful*; *This resource was quite helpful*; and *This resource was very helpful*. Each item in the Assessment Environment section addressed participants' confidence in their understanding of good assessment processes, their ability to conduct assessment activities, and their ability to successfully report assessment activities. Likert-scaled responses for all questions included: *Very Untrue*, *Somewhat Untrue*, *Neither True nor Untrue*, *Somewhat True*, and *Very True*.

Creswell (2014) stated that "[when] one modifies an instrument…the original validity and reliability may not hold for the new instrument, and it becomes important to reestablish validity and reliability during data analysis" (p. 160). To establish validity and reliability, the OIE pilot tested the survey with the Associate Vice President for Institutional Effectiveness, and all seven members of the IE Review Team provided feedback regarding item clarity and arrangement of scale items. Gay, Airasian, and Mills (2009) stated that "if numbers are used to represent the response choices," as with the series of Likert–scaled items that make up the research instrument for this study, "analysis for internal consistency can be accomplished using Cronbach's alpha" (p. 161). Reliability of the instrument was assessed using Cronbach's Alpha, and results showed moderate reliability for utility of individual practices ($\alpha = .64$) and high reliability for knowledge of and confidence in assessment ($\alpha = .92$).

Data Analysis

This study examined the "process of interaction" between IE Review Team members and administrative and student affairs units.

Researchers used descriptive statistical measures to evaluate perceived knowledge of and confidence in the assessment process and utility of specific resources. Mean scores were calculated both in the aggregate and by division to determine any variance in utility amongst the divisions represented. These data addressing the first three research questions provided the OIE with a better understanding of participants' knowledge of and confidence in the assessment process, as well as the perceived strengths and weaknesses of the resources the OIE supports.

Treating the impact of specific resources as an independent variable, the researchers applied regression and correlation methods to determine if relationships existed between each independent variable and a constructed dependent variable, the knowledge of and confidence in assessment composite score (KCC score). Researchers constructed individual KCC scores by calculating an average of each participant's responses to the three questions in the Assessment Environment section of the survey. Regression coefficients provided the means of estimating the extent to which one variable impacted another, while correlation coefficients provided a way to assess the accuracy of those estimates (de Vaus, 2014). This

provided an appropriate means of examining the effects of specific resources supported, such as face-to-face feedback and written feedback, and knowledge of and confidence in the assessment process. Correlation matrices were compiled to display and review the results of these analyses.

Results and Discussion

Findings are presented in two primary categories. The first category addresses perceived utility of specific resources and participants' perception of their knowledge of and confidence in the assessment process. The second category addresses the relationship between perceived utility of specific resources and participants' perception of their knowledge of and confidence in the assessment process.

Individual Practices and Processes

Participants rated the utility of each specific resource using a six-point Likert scale, with 1 indicating *I did not know about this resource*, 2 indicating *I knew about this resource but did not use it*, and three through six indicating levels of utility, ranging from *This resource was not at all helpful* (3) to *This resource was very helpful* (6). Individual items addressed the utility of *General information about assessment from OIE's website* (OIE Website), *General information about assessment from sources other than the OIE website, such as assessment books or conference workshops* (External Individual Resources), *Face-to Face* (F2F), *feedback from IE Review Team Members during the annual review*, *Electronic feedback from OIE and IE Review Team Members outside the annual review* (Electronic Feedback), *Consultation with IE Review Team Members outside the annual review* (RT Off Cycle), *Consultation with OIE staff outside the annual review* (OIE Off Cycle), *Administrative, Academic*, and *Student Support Services Rubric* (OIE Rubric), and the *Rubric and example specific to each division* (Divisional Example). Table 1 highlights descriptive statistics for the specific resources while Table 2 presents inter-item correlations.

Table 1
Descriptive Statistics for Utility of Specific Resources

	OIE Website	External Resources	F2F	Electronic Feedback	RT Off Cycle	OIE Off Cycle	OIE Rubric	Divisional Example
Mean	3.21	3.00	5.11	4.92	4.05	4.21	3.54	3.70
Median	4.00	3.00	5.00	5.00	4.00	5.00	4.00	4.00
Mode	1.00	1.00	6.00	5.00	6.00	6.00	4.00	5.00
Std. Dev.	1.77	1.81	0.92	1.01	1.72	1.77	1.76	1.80
Variance	3.14	3.27	0.84	1.01	2.95	3.14	3.09	3.25
Skewness	-0.02	0.07	-0.91	-1.16	-0.43	-0.54	-0.27	-0.39
Kurtosis	-1.52	-1.66	0.78	2.45	-1.23	-1.20	-1.19	-1.24
Range	4.00	5.00	5.00	4.00	5.00	5.00	5.00	5.00

Note. $n = 61$

In the aggregate, participants reported the least useful resources to be the OIE Website and External Resources that participants seek or experience outside their interaction with the OIE. Means were 3.21 and 3.00 respectively, indicating these individual practices were not helpful. The highest means were reported for F2F Feedback and Electronic Feedback, with means of 5.11 and 4.92 respectively, indicating these specific resources were helpful. Regarding correlations between resources, statistically significant correlations were most notably found between resources of similar format. For example, relatively static sources of information (i.e., OIE Website and External Resources) showed a mild, statistically significant correlation of .47. Similarly, static templates or examples (i.e., OIE Rubric and Divisional Example) demonstrated a high, statistically significant correlation of .82. Perhaps not surprisingly, resources incorporating some form of dynamic, personalized interaction (i.e., F2F, Electronic Feedback, RT Off Cycle, and OIE Off Cycle) produced multiple statistically significant correlations (see Table 2).

Researchers constructed individual KCC scores by calculating an average of each participant's responses to the three questions in the Assessment Environment section.

Table 2
Inter-Item Correlations for Specific Resources

	OIE Website	External Resources	F2F	Electronic Feedback	RT Off Cycle	OIE Off Cycle	OIE Rubric
External Resources	0.47*		0.40*	0.18	0.41	0.57**	0.21
F2F	0.45**	0.40*		0.71**	0.86**	0.90**	0.26
Electronic Feedback	0.45**	0.18	0.71**		0.68**	0.71**	0.48**
RT Off Cycle	0.48*	0.41	0.86**	0.68**		0.87**	0.28
OIE Off Cycle	0.57**	0.57**	0.90**	0.71**	0.87**		0.50**
OIE Rubric	0.24	0.21	0.26	0.48**	0.28	0.50**	
Divisional Example	0.08	-0.03	0.40**	0.47**	0.27	0.46*	0.82**

Note. $n = 61$. **Denotes significant at the p<0.01 level; *denotes significant at p<0.05.

Research question two examined the variation in utility of specific resources among the different divisions represented. F2F and Electronic Feedback were perceived by participants to have the most utility in three of the five divisions represented, which included the division of Vice President - Academic Affairs, President, and Vice President - Student Affairs and Enrollment Management. The divisions of Vice President - Business and Finance (VPBF) and Chief Information Officer/Information Technology (CIOIT) rated OIE Off Cycle as the most useful, followed by F2F Feedback.

Knowledge of and Confidence in Assessment

Participants reported the least useful resources to be the OIE Website and External Resources.

Research question three addressed participants' perceptions of their own knowledge of the assessment process and their confidence in applying that knowledge. Participants responded to a series of Likert-scaled questions focusing on Assessment Environment, with responses ranging from *Very untrue (1)* to *Very true (5)*. Items addressing knowledge of and confidence in assessment were: 1) *I have a solid understanding of what constitutes good assessment practice*; 2) *I am confident I can successfully conduct assessment activities in my unit*; and 3) *I am confident I can successfully report assessment activities in my unit* (see Table 3).

Table 3
Mean Scores, Knowledge of and Confidence in Assessment

	Q1	Q2	Q3
Mean	4.05	4.08	4.05
Median	4.00	4.00	4.00
Mode	4.00	4.00	4.00
Std. Dev.	0.85	0.97	0.88
Variance	0.71	0.94	0.78
Skewness	-0.78	-1.18	-0.85
Kurtosis	0.30	1.14	0.26
Range	3.00	4.00	3.00

Note. $n = 61$.

In all three cases, mean scores reported were all slightly higher than 4.00, indicating that, in the aggregate, participants felt it is at least *Somewhat true* that they understand what constitutes good assessment processes, they can conduct assessment, and they can report their assessment activities. As with utility of individual practices, however, there is variation when results were viewed by division. Participants from the divisions of VPBF

and CIOIT have comparatively less confidence in all three areas. Emil and Cress (2014) noted that perceived skill can affect engagement. Therefore, although it may be true in the aggregate, these common barriers to engagement in assessment may not apply in this case. If the results are in fact a true reflection of participants' perceptions of their knowledge and confidence, some divisions may be more likely to engage than others.

Common barriers to engagement in assessment may not apply in this case… some divisions may be more likely to engage than others.

Correlational Analyses

After review of the descriptive statistics for each item, correlational analyses were utilized to investigate the relationship between knowledge of and confidence in the assessment process and the utility of specific resources in place. To facilitate these analyses, the KCC for each participant was derived from participants' responses to the same three Assessment Environment items listed above. All three items used in the composition of the KCC demonstrated high statistically significant correlations suggesting concurrent validity (see Table 4).

Table 4
Correlational Relationships between Variables Contributing to KCC

	Practice	Conduct	Report
Practice		0.79*	0.71*
Conduct	0.79*		0.87*
Report	0.71*	0.87*	

Note. $n = 61$. *Denotes significant at the $p<0.05$ level.

Correlations between the KCC score, individual practices, and number of assessment cycles in which participants have engaged were then reviewed (see Table 5).

Participants' KCC Scores and Utility of Specific Resources

As shown in Table 5 below, of the eight specific resources identified for this study, only two were shown to have statistically significant relationships with participants' KCC scores. Using Pearson's correlation, both Electronic Feedback and resources on the OIE Website demonstrated statistically significant positive relationships with participants' KCC scores at the $p < 0.05$ level.

Table 5
Correlational Relationships between Participant KCC Scores and Utility of Specific Resources

	Cycles	OIE Website	External Resources	F2F	Electronic Feedback	RT Off Cycle	OIE Off Cycle	OIE Rubric	Divisional Example
KCC	0.11	0.38*	0.22	0.25	0.32*	0.04	0.18	0.24	0.17

Note. $n = 61$. *Denotes significant at the $p<0.05$ level.

Before conducting regression analyses, the researchers conducted a second set of descriptive and correlational analyses, excluding all responses of (1) *I did not know about this resource* or (2) *I knew about this resource but did not use it* from section two of the survey instrument. This manipulation of the data permitted analyses of the perceived utility of each specific resource as reported only by participants who actually used each resource. Descriptive statistics are presented in Table 6 below. The sample size varies due to the number of participants who used each resource.

In the aggregate, participants who have used the specific resources the OIE supports reported the least useful resources to be the OIE Rubric and the OIE Website, with means of 4.48 and 4.60 respectively. The highest means were reported for OIE Off Cycle and F2F, with 5.18 and 5.17 respectively. These targeted times for interaction with assessment coordinators and the IE Review Team and OIE staff provided the opportunity to encourage needed reflection and engagement in the assessment process as indicated in the literature

Table 6
Descriptive Statistics for Utility of Specific Resources Manipulated

	OIE Website	External Resources	F2F	Electronic Feedback	RT Off Cycle	OIE Off Cycle	OIE Rubric	Divisional Example
N	35	33	60	60	43	44	44	45
Mean	4.60	4.55	5.17	4.98	5.02	5.18	4.48	4.62
Median	5.00	5.00	5.00	5.00	5.00	5.00	4.00	5.00
Mode	5.00	5.00	6.00	5.00	6.00	6.00	4.00	5.00
Std. Deviation	0.85	0.79	0.83	0.87	0.91	0.92	1.02	1.05
Variance	0.72	0.63	0.68	0.76	0.83	0.85	1.05	1.10
Skewness	-0.03	-0.16	-.051	-0.44	-0.44	-0.75	0.13	-0.28
Kurtosis	-0.50	-0.25	-0.82	-0.58	-0.85	-0.55	-1.67	-1.08
Range	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00

Note. n varies from 33 to 60.

(Gebelica, Van den Bossche, De Maeyer, Segers, & Gijselaers, 2014). Both are needed for participants to see the benefit of assessment beyond external factors and to develop confidence and skill in the process (Emil & Cress, 2014).

Like students, participants in this study appreciated both face-to-face and electronic feedback provided during the institution's annual review process.

Statistically significant results for this question of KCC correlation differed when conducting analyses based on the entire sample for the study versus only those participants who have actively participated by using a particular resource. In the aggregate, only Electronic Feedback and the OIE Website demonstrated statistical significance. When removing participants who had not used specific resources from the correlation, Electronic Feedback continued to produce statistical significance, but the OIE Website did not. Instead, four additional individual resources, including F2F, RT Off Cycle, OIE Off Cycle, and the OIE Rubric, demonstrated statistically significant relationships with KCC scores. The work of Panadero and Romero (2014) is corroborated in the reported utility of the institutional rubric in that it is helpful for participants to have an idea of what their final products should look like, and the OIE rubric provides that guidance. Overall, however, the opportunities for personal or electronic interaction continued to have the most perceived utility. These findings are similar to those of Rodgers et al. (2013), which also supported consultation with assessment professionals and the use of feedback, and Kahlon et al. (2015), which promoted formative feedback, particularly in a face-to-face setting.

Further analysis was next conducted to explore the relationship between participants' KCC scores and those specific resources with statistically significant relationships to the participants' KCC scores. As shown in Table 7 below, of the eight specific resources identified for this study, when considering only those participants who have used the specific resources provided, five resources were shown to have statistically significant relationships with participants' KCC scores, as opposed to two when considering all participants. Using Pearson's correlation, F2F, Electronic Feedback, RT Off Cycle, OIE Off Cycle, and the OIE Rubric demonstrated statistically significant relationships with KCC at the $p < 0.01$ level as depicted in Table 7.

Table 7
Correlation between Participant KCC Score and Utility of Specific Resources Maniputlated

	OIE Website	External Resources	F2F	Electronic Feedback	RT Off Cycle	OIE Off Cycle	OIE Rubric	Divisional Example
KCC	0.33	0.29	0.35**	0.34**	0.54**	0.55**	0.42**	0.14

Note. ** Denotes significance at the $p < 0.01$ level (2-tailed)

Regression Analyses

Finally, while the correlational analyses indicated significant relationships between participants' KCC scores and five of the specific resources, researchers were interested in

the variance (in participants' KCC scores) accounted for by specific resources. Hierarchical regression was applied using results from correlational analyses and researchers' discretion in composing model steps. Specifically, RT Off Cycle and OIE Off Cycle served as step one of the model and F2F, Electronic Feedback, and the OIE Rubric were selected as step two of the model (see Table 8).

Both steps of the model were found to be statistically significant at the $p < .05$ level, predicting variance within participants' KCC score. Step one of the model accounted for approximately 31% of the variance, with step two adding a slight increase of approximately 8%. Review of histograms suggested normal distribution of residuals; however, collinearity statistics (i.e., tolerance and VIF) suggested caution (Field, 2018).

Table 8
Linear Regression Model Summary

	R	Adjusted R^2	SE	*b*	β
Step 1	0.60	0.31	0.71		
Constant				1.236	
RT Off				-0.2	-0.02
OIE Off				0.58	0.62
Step 2	0.67	0.32	0.71		
Constant				0.56	
RT Off				-0.46	-0.44
OIE Off				0.60	0.65
F2F				0.12	0.01
Electronic Feedback				0.39	0.37
OIE Rubric				0.17	0.20

Implementing successful institutional assessment processes is important both in terms of external accountability and internal success. The findings from this study support that participants value the opportunities the OIE provides for indirect and direct interaction with members of the OIE staff and the IE Review Team. Although existing literature regarding the benefits of peer review focus largely on academic assessment (Jonsson, 2013; Kahlon et al., 2015), the premise is very much the same. Like students, participants in this study appreciated both face-to-face and electronic feedback provided during the institution's annual review process.

While the OIE cannot control the utilization of specific resources, it can take steps to be certain that those resources it does provide via its website are helpful to those who seek them.

The majority of the OIE's and the IE Review Team's contact with administrative and student affairs units each year is focused on preparing annual assessment plans and reports. IE Review Team members review both documents and provide feedback to those responsible for report preparation. Written feedback is first shared electronically and is then shared during an annual face-to-face review process during which those who contribute to these documents and those who review them discuss opportunities to improve final reports and develop assessment plans for the coming year. IE Review Team members assist responsible administrators and staff in identifying positive attributes, as well as addressing weaknesses. Gebelica, Van den Bossche, De Maeyer, Segers, and Gijselaers (2014) found support for "accurate and timely feedback" in encouraging "active engagement" and "reflective interactions" (p. 93), which is consistent with the findings of this study. This face-to-face review process provides units with dedicated time to work with IE Review Team members and think critically about the objectives they were trying to accomplish, to determine how effective their strategies were in accomplishing those objectives, and to identify what they may need to do differently going forward. These established feedback processes have demonstrated value to participants and may continue to promote productive engagement in the institution's assessment processes if continued.

Although the aggregate mean scores for consultation with OIE staff or IE Review Team members varied slightly when considering all participants versus only those participants who used these specific resources, consultations outside the annual review process were still perceived to be among the top four most useful resources and further corroborated

the benefits of peer feedback (Nicol, Thomson, & Breslin, 2014). Of the participants, 33% were either unaware they have the option of consulting with an OIE staff member outside the annual review process or they chose not to pursue the option. Furthermore, of these, 28% were unaware of this same option for consulting with a member of the IE Review Team. However, for both resources, when considering only those participants who had used them, the most common response was *This resource was very helpful* (6). Additionally, results from the regression analyses suggest that off-cycle consultation with OIE staff serves as a greater predictor of participant confidence in assessment (i.e., KCC score) than even more dynamic (e.g., face-to-face) forms of interaction during recognized assessment periods. Given these facts, the OIE may benefit from better publicizing such options moving forward (Hahn & Lester, 2012).

Both the OIE Rubric and the Divisional Example present additional publicity possibilities for the OIE. Panadero and Romero (2014) concluded that rubrics, when "well-designed…can have a positive impact on performance" (p. 142). As with the opportunities for consultation outside the annual review cycle, 28% of participants were either unaware of the OIE rubric used to evaluate the quality of completed assessment reports or chose not to consult it, and 26% were either not aware of or chose not to consult the Divisional Example designed as an example of strong assessment reporting for each division. For those using these specific resources, mean scores in the aggregate showed that each were almost squarely between *a little helpful* (4) and *quite helpful* (5). Results by division show that only the VPSAEM participants felt the *Divisional Example* was at least *quite helpful* (5), while the OIE Rubric was only *a little helpful* (3), and for all other divisions, reported means for both the OIE Rubric and the Divisional Example were also only *a little helpful* (3). This suggests the OIE may have opportunities for improvement on both of these specific resources.

Finally, although the OIE Website and External Resources were perceived to be at least *a little helpful* (4), in the aggregate, results considering only those participants who used these specific resources highlight additional publicity efforts may be in order. Forty-three percent of participants were either unaware of materials posted on the OIE website or chose not to use them, and 46% were either unaware that External Resources were available or chose not to use them. While the OIE cannot control the utilization of specific resources, it can take steps to be certain that those resources it does provide via its website are helpful to those who seek them. It may therefore be beneficial for the OIE to examine more closely if resources are recognized but not used or truly are not recognized as available options.

This study… provided baseline data for assessment teams to begin a decision-making process and determine… which resources should be continued or modified.

Limitations, Delimitations, and Assumptions

The immediate results of this study are limited to one university, but the results can extend the body of literature that exists relative to administrative and student affairs assessment in higher education. Existing literature often fails to go beyond anecdotal evidence in support of concrete quantitative data and this study provided quantitative data to support which specific assessment resources were perceived to be more helpful than others. Specifically in regard to the regression analyses of this study, tolerance statistics suggested interpretation of results with caution due to multicollinearity concerns. Furthermore, because data were collected to study the impact of administrative and student affairs assessment processes at one large, public southeastern university, generalizability is limited; however, the results should still be of use to assessment practitioners beyond the study setting.

Implications for Practice

Findings from this study are the first step in conducting ongoing programmatic assessment of the effectiveness of administrative and student affairs assessment practices at one large public, southeastern university. Data collected provide the baseline assessment data regarding the perceived strengths and weaknesses of specific resources supported by OIE assessment teams. Additional data provided new insight into participants' perceptions of their own knowledge of and skill in applying assessment processes.

In expanding the assessment process, it is vital to recruit professionals who have demonstrated some skill in applying effective assessment processes. Data from this study

suggest that, in the aggregate, all participants in this study felt it is at least *somewhat true* that they are able to do so. The OIE and the assessment team may consider revising this section of the survey instrument to better determine those individuals who may be best suited to coach others in conducting and reporting assessment activities. It is possible, for example, that participants feel reasonably certain they can perform these activities themselves, but they are far less certain they could assist others in doing so. As the OIE and the assessment team consider revising individual practices and processes, it could be helpful to collect qualitative information from participants regarding ways to improve the utility of each.

Conclusion

The OIE has established and developed assessment resources over time but their impact has not been routinely and formally investigated. Although this study was limited to a single office working with a specific population of administrative and student affairs assessment coordinators, administrators, and staff, study findings corroborate the positive impact of rubrics and peer review and feedback, providing the OIE with a basis for continuing to support many of its existing resources.

This study was intended to help address questions about the effectiveness of the resources in place in support of institutional administrative and student affairs assessment units to help ensure all resources contribute to the effectiveness of the assessment process, and the researchers believe the findings support these efforts. It is important to "ask the tough questions and to get the news that something is not working (or working as assumed) and should therefore be revised or eliminated" (Meyer & Murrell, 2014, p. 4). This study, which may serve as a model for other institutions that support similar resources, provided baseline data for assessment teams to begin a decision-making process and determine, based on evidence collected, which resources should be continued or modified to attain the most beneficial assessment outcomes.

Appendix

*Assessment Resources and Environment Survey Instrument**
*Adapted, with permission, from Rodgers, M., Grays, M.,Fulcher, K, & Jurich, D. (2013)

Thinking about the assessment resources provided on campus, please choose the phrase that best describes your perception of the usefulness of each resource.						
	This resource was very helpful.	*This resource was quite helpful.*	*This resource was a little helpful.*	*This resource was not at all helpful.*	*I knew about this resource but did not use it.*	*I did not know about this resource*
General information about assessment from OIE's website						
General information about assessment from sources other than the OIE website, such as assessment books or conference workshops						
Face-to-Face feedback from IE Review Team Member (during annual review)						
Electronic feedback from OIE and IE Review Team Member (during annual review)						
Consultation with IE Review Team Member (outside annual review sessions)						
Consultation with OIE staff (outside annual review sessions)						
Administrative, Academic, and Student Support Services Rubric						
Rubric and example specific to my division (e.g., VPBF, VPSAEM, etc.)						

Thinking about the assessment environment in your particular division (for example, Business and Finance or Academic Affairs), how would you respond to each statement?					
	Very true	*Somewhat true*	*Neither true nor untrue*	*Somewhat untrue*	*Very untrue*
I have a solid understanding of what constitutes good assessment practice.					
I am confident I can successfully conduct assessment activities in my unit.					
I am confident I can successfully report assessment activities in my unit.					
Number of assessment cycles in which you have participated					
	1	2	3	4	5 or more
Your reporting division					
	President	VPAA	VPSAEM	VPBF	CIOIT

References

Chalmers, D., & Gardiner, D. (2015). An evaluation framework for identifying the effectiveness and impact of academic teacher development programmes. *Studies in Educational Evaluation, 46* (Evaluating Faculty Development), 81–91.

Creswell, J.W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage Publications.

de Vaus, D. (2014). *Surveys in social research*. 6th ed. New York: Routledge/Taylor & Francis Group.

Eaton, J. S. (2013). Accreditation and the next reauthorization of the higher education act. *Inside Accreditation, 9*(3).

Emil, S., & Cress, C. (2014). Faculty perspectives on programme curricular assessment: individual and institutional characteristics that influence participation engagement. *Assessment & Evaluation in Higher Education, 39*(5), 531.

Field, A. (2018). *Discovering statistics using IMB SPSS statistics* (5th ed.). Thousand Oaks, CA: Sage Publications.

Fink, L. D. (2013). Innovative ways of assessing faculty development. *New Directions for Teaching & Learning, 2013*(133), 47–59.

Fishman, S. M. (2017). Utilizing an assessment planning model: Revision through reflection. *American Association of University Administrators, 32*(1), 186–194.

Fulcher, K. H., Coleman, C. M., & Sundre, D. L. (2016). Twelve tips: Building high-quality assessment through peer review. *Assessment Update, 28*(4), 1.

Fulcher, K. H., & Bashkov, B. M. (2012). Do we practice what we preach? The accountability of an assessment office. *Assessment Update, 24*(6), 5.

Gay, L. R., Airasian, P. W., & Mills, G. E. (2009). *Educational research: Competencies for analysis and applications*. Upper Saddle River, N.J.: Merrill/Pearson.

Gebelica, C., Van den Bossche, P., De Maeyer, S., Segers, M., & Gijselaers, W. (2014). The effect of team feedback and guided reflexivity on team performance change. *Learning and Instruction*, 3486–96.

Hahn, T. B., & Lester, J. (2012). Faculty needs and preferences for professional development. *Journal of Education for Library and Information Science*, (2), 82.

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*(1), 63–76.

Kahlon, J., Delgado-Angulo, E. K., & Bernabé, E. (2015). Graduates' satisfaction with and attitudes towards a master programme in dental public health. *BMC Medical Education, 15*(1), 61.

Krzykowski, L., & Kinser, K. (2014). Transparency in student learning assessment: Can accreditation standards make a difference? *Change, 46*(3), 67–73.

Martin, I. H., Goulet, L., Martin, J. K., & Owens, J. (2015). The use of a formative assessment in progressive leader development. *Journal of Leadership Education, 14*(4).

Meyer, K. A., & Murrell, V. S. (2014). A national survey of faculty development evaluation outcome measures and procedures. *Online Learning, 18*(3).

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education, 39*(1), 102–122.

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice, 21*(2), 133–148.

Rodgers, M., Grays, M., Fulcher, K., & Jurich, D. (2013). Improving academic program assessment: A mixed methods study. *Innovative Higher Education, 38*(5), 383–395.

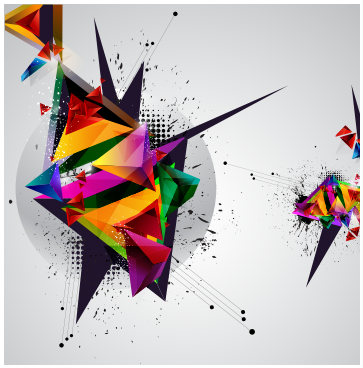
Shutt, M. D., Garrett, J. M., Lynch, J. W., & Dean, L. A. (2012). An assessment model as best practice in student affairs. *Journal of Student Affairs Research and Practice, 49*(1), 65-82.

Slager, E. M., & Oaks, D. J. (2013). A Coaching model for student affairs assessment. *About Campus, 18*(3), 25-29.

Yarber, L., Brownson, C. A., Jacob, R. R., Baker, E. A., Jones, E., Baumann, C., & Brownson, R. C. (2015). Evaluating a train-the-trainer approach for improving capacity for evidence-based decision making in public health. *BMC Health Services Research, 15*(547).

Abstract

Although faculty are an important part of collecting, analyzing, and using student learning data for improvement, significant barriers often prevent faculty from being involved in assessment work outside the classroom. One potential obstacle to faculty involvement in assessment is the misalignment between the work and faculty rewards structures. Through a multi-level sequential development mixed methods study, this paper addresses a gap in the assessment literature by describing how faculty are rewarded and recognized for assessment work outside the classroom. Stipends and course releases, food and drinks at assessment meetings, promotion and tenure consideration, and general campus exposure were the rewards most frequently described by faculty participants. Administrators seeking to improve faculty engagement in assessment work should consider creating specific policies permitting assessment work to be included in the promotion and tenure process to both recognize the importance of assessment work and reward faculty for participating in such efforts.



AUTHORS

Elizabeth E. Smith, Ph.D.
University of Tulsa

Sarah Gordon, Ph.D.
Arkansas Tech University

How are Faculty Rewarded and Recognized for Assessment Work Outside the Classroom?

This study examines the ways in which faculty are recognized and rewarded for assessment work outside the classroom. Assessment work may include a variety of activities such as serving on program, department, college/school, or university-level assessment committees; scoring student artifacts for university-wide assessment; acting as the assessment coordinator for a department or unit; and training other faculty on assessment practices. Faculty play an integral role in collecting, analyzing, and using student learning data; yet, there are barriers to their involvement in these processes (Bresciani, 2011; Cain & Hutchings, 2015). One way to address such barriers is to tie participation to meaningful rewards for faculty. This paper builds upon existing literature by surveying campus administrators and interviewing faculty to explore how faculty are rewarded and recognized for assessment work.

Literature Review

While its roots go back as far as the 1930s, assessment began to develop as a distinct field in the mid-1980s as higher education leaders recognized the opportunity to evaluate student learning data and use it for improvement (Ewell, 2002). Assessment work addresses a myriad of topics including using data to improve student learning, (Kuh, et al., 2015), how to develop learning outcomes and match them with assessment measures (Allen, 2006; Bresciani, 2006; Driscoll & Wood, 2007), and the use of technology in assessing student learning (Light, Chen, & Ittelson, 2012; Yancey, 2009). Three specific questions emerge when considering the role of faculty in assessment in higher education: (a) Why is it important for faculty to be involved in analyzing and using student learning data? (b) How do colleges and universities include faculty in assessment work? and (c) What are the common barriers to faculty involvement in assessment activities? This literature review will answer those questions from the existing literature while noting the gaps this study addresses. Throughout this literature review and paper we use the term "faculty" to refer to full-time teaching faculty—those who may be tenure-track, tenured, or in a different type of full-time teaching

CORRESPONDENCE

Email
elizabeth-smith-43.utulsa.edu

position. As Kezar and Maxey (2014) noted, contingent and adjunct faculty are often not invited to participate in assessment work.

Why is it important for faculty to be involved in assessment work?

The American Association for Higher Education's Nine Principles of Good Practice for Assessing Student Learning declare that "Faculty play an especially important role" in campus assessment efforts, but do not delineate expectations for faculty involvement (American Association for Higher Education, 1996, para. 6). Because faculty traditionally maintain the most frequent contact with students, faculty involvement is key in utilizing program and institutional student learning data for improvement and developing program and institutional learning outcomes (Allen, 2004; Ebersole, 2009). Developing outcomes and conducting assessment work without input from faculty risks a lack of buy-in (Grunwald & Peterson, 2003). When assessment activities are perceived to be forced upon faculty by administration or external agencies, faculty may resist being involved for a variety of reasons (MacDonald, Williams, Lazowski, Horst, & Barron, 2014).

How do colleges and universities include faculty in assessment work?

Faculty play an integral role in collecting, analyzing, and using student learning data.

Institutions can move beyond cursory involvement of faculty in reviewing student learning data. Faculty who engage in collecting, analyzing, and using student learning data with the support of assessment professionals may have more positive attitudes toward assessment (Ebersole, 2009). Faculty at institutions who employ the Boyer model may be more likely to be involved in assessment work because of its connection to the scholarship of teaching and learning (SoTL) (Boyer, 1990; Hutchings, 2010; Secret, Leisey, Lanning, Polich, & Shaub, 2011). Administrators who recognize the SoTL may also accept participating in assessment as scholarship because of their "overlapping agendas, practices, and institutional constituencies" (Hutchings, 2010, p. 11). At institutions where faculty see themselves primarily as teachers, they may connect more to assessment work because of its relationship to the classroom (Hutchings, 2010). Faculty can engage in assessment work by applying an inquiry framework to investigate questions they have about student learning (Cain & Hutchings, 2015; Smith, 2017).

Institutions also involve faculty in assessment work by offering assessment-related professional development, faculty fellowships in assessment offices, and reduced teaching loads for faculty with assessment responsibilities (Ewell, Paulson, & Kinzie, 2011; MacDonald et al., 2014).

What are common barriers to faculty involvement in assessment activities?

Assessment research literature delineates three major obstacles for faculty involvement: time, resources, and understanding of assessment (Bresciani, 2011). Faculty at different types of institutions have varying demands on their time. Those working at research intensive universities have greater expectations of scholarly production while faculty at community colleges or teaching-focused four-year universities have additional teaching and service responsibilities.

Lack of expertise in assessment and scarcity of resources can also prevent faculty involvement in assessment work. While studies suggest robust faculty development to promote faculty engagement in assessment, slashed budgets in higher education leave fewer dollars for faculty training (Grunwald & Peterson, 2003; MacDonald et al. 2014). A lack of expertise and confidence in assessment is related, in part, to poor funding of professional development for faculty and a lack of training for assessment work in doctoral programs (Bresciani, 2011).

How is Assessment Work Recognized in the Promotion and Tenure Process?

Faculty involvement in assessment work is key to ensure aligned institutional assessment processes that reflect opportunities for students to learn. However, one of the greatest fears among faculty when asked to engage in assessment work is that it will "take us

away from the important work of teaching, scholarship, and service" (Crossley & Wang, 2010, p. 276). Because teaching, research, and service are emphasized (and assessment is viewed as something separate), faculty are discouraged from participating in assessment work by the existing rewards structure (MacDonald et al., 2014). Faculty only have a finite amount of time, and spending their time on the activities that are rewarded with promotion and tenure (P/T) is rational (Bresciani, 2011). Thus, Cain and Hutchings (2015) suggest promoting greater faculty involvement in assessment work by tying it to three items of importance for faculty: student learning, their research, and P/T.

Though previous research makes clear that faculty need to be involved in assessment work (Allen, 2004; American Association for Higher Education, 1996; Ebersole, 2009; Reder & Crimmins, 2018), and it is clear that assessment work can be connected to scholarship (Boyer, 1990; Hutchings, 2010; Secret et al., 2011), there has not been any exploration into how assessment work is recognized or rewarded in the P/T process. Understanding the faculty rewards system is vital as it is the system that faculty must work within to remain and advance in academia. Thus, this paper examines the ways in which faculty engaged in assessment work are rewarded and recognized with a specific focus on exploring recognition during the P/T process.

Lack of Fit with Existing Conceptual Frameworks

Having extensively searched the available literature, we approached this topic with the recognition that this work does not fit within an existing conceptual framework. In reading the literature, we found that frameworks used in assessment work only peripherally (if at all) link to faculty rewards and recognition and, more specifically, the P/T process. For example, Boyer's Model of Scholarship (1990, 1996) positions assessment work as part of the SoTL. Although this model provides insight into how assessment endeavors can segue into scholarship, it does not explore the recognition of assessment work in P/T processes.

There has also been extensive research and theoretical development with regard to assessment culture and climate in higher education. For example, Stevenson, Finan, & Martel (2017) drew from work on assessment culture and evaluation capacity-building to create a developmental model for understanding institutional assessment capacity. This model allowed for a means "to speak faculty's perceived truths [regarding assessment] to those with power—power to communicate genuine belief in the value of an ideal assessment culture and support forward movement with policies, recognition, and resources" (p. 44). Those with power may include those who are part of the P/T process, but the model was not created with that specifically in mind.

Lack of expertise in assessment and scarcity of resources can also prevent faculty involvement in assessment work.

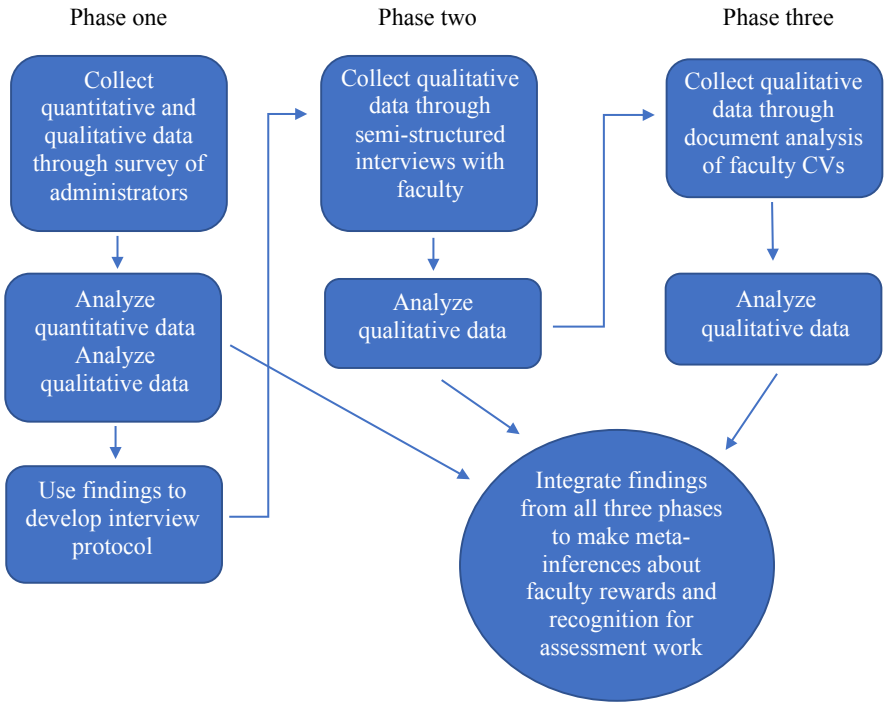
The impetus for this study arose from the work of the authors in various roles within the field of assessment. In our work as directors of assessment, coordinators for accreditation efforts, and as tenure-track faculty members, we have regularly encountered the issues that are clearly and extensively outlined in the literature. In addition to examining the existing literature, our regular attendance at conferences that focus on assessment and our interactions in major assessment online networks indicate that there is not yet a general understanding of rewards for faculty who engage in assessment work, nor is there a common understanding of how (if at all) this work is recognized in the P/T process. The goal of this paper was not to build upon or create a conceptual framework with regard to engaging faculty in assessment work, but to understand the state of how things are now: how are faculty being rewarded and recognized for the work they are doing?

Methods

A multilevel sequential development mixed methods approach (see Figure 1) was utilized for this study (Teddlie & Tashakkori, 2009). Development uses the results from one phase of data collection and analysis to inform the following phases (Greene, Caracelli, & Graham, 1989) and improves validity of the overall findings by using holistic data triangulation to answer the research question (Turner, Cardinal, & Burton, 2017).

To understand how faculty are rewarded and recognized for assessment work, the researchers sought responses from two sample groups: department and school/college

Figure 1
Multilevel Sequential Development Mixed Methods Study Design



The most common types of rewards included faculty awards for assessment work and stipends for assessment work.

administrators who oversee the assessment processes in their departments or school/colleges and faculty who are directly involved in the work of assessment in their department or school/college. The samples were created using convenience and snowball sampling techniques (Creswell, 2014). To obtain the sample group of administrators involved in overseeing assessment at a variety of institutional types (for phase 1 of the study), an email invitation with the survey link was sent out to two listservs for those who are engaged in assessment work in higher education; participating administrators suggested faculty at their institutions who were involved in assessment work to participate in phase two of the study.

The survey used in phase one collected both qualitative and quantitative data. Based on existing research literature to improve content validity, questions on the survey were designed to elicit information from administrators regarding policies and practices pertaining to the reward and recognition of faculty members who participate in assessment work. Survey participants shared about who conducted assessment work at their institutions and how they were trained; the role the participant played in conducting or overseeing assessment work on their campus; how (if at all) faculty were rewarded, recognized, or incentivized for assessment work on their campus; how (if at all) faculty used assessment work in annual appraisals and/or in P/T documents; and whether there were policies or practices in their department/university that encouraged the use of assessment activities in faculty appraisals/P/T documents. The researchers used inductive qualitative analysis to find themes among the responses to the qualitative survey items (Patton, 2002). Descriptive statistics were used to report on the quantitative items.

The second phase included semistructured interviews with 11 faculty members from different institutions. Employing the development approach, findings from phase one were used to formulate interview questions for phase two. Interview questions focused on the faculty members' experience with assessment (outside the classroom), rewards and recognition for assessment work on their campus, if/how they included assessment work on their vitae and/or in P/T documents, and if/how assessment informed their work. Using inductive qualitative analysis, the researchers reviewed the transcripts of each interview, treating the transcripts as texts, and formulated open codes by identifying words or phrases that directly related to the research questions (Patton, 2002). The researchers employed *in vivo* coding to ensure that the responses of participants were utilized in the open codes (Saldana, 2012). After reviewing

all of the open codes, the researchers grouped like codes into themes until agreement was reached among the researchers about the themes.

In phase three, the researchers employed document analysis to review the CVs of interviewed faculty for mentions of assessment work. After the third phase was completed, the researchers used holistic data triangulation to compare the findings from all three phases to make meta-inferences about rewards and recognition for faculty members who participate in department or school/college assessment activities (Turner, Cardinal, & Burton, 2017). IRB approval was obtained prior to data collection.

Results

Administrator Survey

Thirty-seven administrators from 28 institutions completed the survey comprised of open- and closed-ended questions. Twenty participants were senior administrators (holding the title of president, vice-president, dean, assistant/associate dean, or director), five were tenure track faculty, three were nontenure track faculty (with two specifically reporting they belonged to nontenure institutions), five were staff, and four indicated they belonged to some other unspecified classification or did not provide a response. The institutions at which participants worked were a mixture of private and public and represented all six regional accrediting bodies; 77% of respondents belong to programs, departments, or colleges that held specialized accreditation from a national or professional organization (such as ABET, CAEP, AACSB, etc.).

Seventy-five percent of respondents indicated that there were no policies or practices in place to encourage consideration of assessment work in the P/T process.

While 78% of respondents indicated that faculty complete the majority of assessment work at their institutions, 65% percent of respondents indicated that there are no tangible rewards for faculty to participate in assessment work. Among the 35% that responded affirmatively, the most common types of rewards included faculty awards for assessment work and stipends for assessment work. Table 1 outlines all responses to the question "How (if at all) are faculty compensated for assessment work?" Respondents were permitted to choose multiple types of faculty compensation.

Table 1
Responses to the Question "How (if at all) are faculty compensated for assessment work?"

Course release time	9 (25%)
Stipends/pay	11(30.56%)
Travel funds	1 (2.78%)
Grants/contract pay	2 (5.56%)
Other	7 (19.44%)
They are not compensated because assessment work qualifies as service	18 (50%)
No compensation is provided for assessment work	16 (44.44%)
I'm not sure	1 (2.78%)

Typical responses for "Other" were awards and "it depends." For example, one participant noted:

> It depends on the college. Colleges that have graduate program officers and undergraduate program officers (faculty promoted to administrative roles) typically coordinate assessment. However, some colleges have dedicated assessment people (e.g., Education and Pharmacy), while others have faculty assigned to assessment, without formal recognition of the additional tasks required.

Another participant remarked, "It depends on the Dean if assessment work counts toward service, but even when it does–it is minimal." Some respondents added that there were other types of rewards and recognition for assessment work, including public recognition, iPads, invitations to present their assessment work to colleagues, and certificates acknowledging assessment work.

Administrators were also asked two open-ended questions related to the use of assessment work in the P/T process. In response, 12 participants (32%) responded that assessment work was "not at all" used for P/T. Two noted that they were not sure if or how it was used for P/T. The second most common response to this question was that assessment work was most often counted/categorized as service in the P/T process, and many of the responses by administrators indicated that 'service' did not count much toward the P/T process. For example, one participant noted, "I count assessment work (attending assessment retreats, for example) as program service, which is a sub-category under institutional service. It is recognized, but minimally, by Deans." Another remarked, "Not much at all. It counts as service, which doesn't count for much at all." Yet another participant wrote:

> They [faculty] evidence their roles [in assessment work], but I do not believe that assessment carries any weight in [P/T], even though administration first said that it would/should. They have positioned it more and more as service, and service is barely counted compared to teaching and peer-reviewed top-tier publishing.

Four administrators noted that faculty members have asked for letters of recommendation from the assessment director or have included thank you notes from the assessment office in their P/T documents.

Three administrators indicated that some faculty positioned their assessment work as teaching or research. One participant noted, "Several of our faculty have published papers or posters using assessment data which count towards promotion." Another administrator remarked, "Evidence of both teaching and service depending upon involvement. Also trying to move toward research evidence with SoTL projects." Finally, one administrator noted that the connection of assessment to scholarship was a way to "legitimize" assessment work:

> Some faculty have experienced success with scholarly publications based on their assessment work. Increasingly, departments recognize this as legitimate scholarship (SoTL). When the work is based on course-embedded assessments and faculty use findings to improve learning in a class, this can count as evidence for the quality of teaching. The University Guidelines for [P/T] and evaluations recognize the possibility that assessment might support these categories, although some departments have not fully embraced this concept.

Overall, faculty indicated that there are very few tangible rewards for engaging in assessment work.

Seventy-five percent of respondents indicated that there were no policies or practices in place to encourage consideration of assessment work in the P/T process. Of those that did indicate there were such policies or practices, administrators again noted that it was often tied to SoTL. For example, "If [faculty] publish or present on the SoTL, this counts towards promotion." Another participant echoed that in practice, assessment work was a part of teaching or service: "[We have] emerging practices…[for example], mini course assessment reports for evidence of teaching as well as letters of recommendations from our office for being supportive/involved in assessment activities in service capacity." Overall, administrators who oversee assessment work at department or college/school level indicated that assessment work is unlikely to be counted toward P/T or, if it is, it is considered as service, which has varying levels of importance depending on the institution type.

Faculty Interviews

Eleven faculty members were interviewed to learn more about their experiences in participating in assessment work. The faculty members interviewed represented diverse institution types from small liberal arts colleges to large land-grant universities, with nine

working at four-year institutions and two participants representing two-year institutions. Faculty participants also represented a range of academic areas including agriculture, biology, Education, fine arts, and history.

Overall, faculty indicated that there are very few tangible rewards for engaging in assessment work. For example, one faculty member said, "I can't point to [any] tangible [rewards]. I can't see any way, shape, or form that…that I've gained anything in life other than that's just part of the job I'm expected to do." Another said, "I don't see that [anyone] has done anything other than, 'Oh, you're doing what you're supposed to. Great!'" When faculty noted rewards for their work in assessment, the rewards fell into four categories: stipends and course releases, food and drinks, P/T consideration, and general campus exposure.

Assessment work seemed to be helpful for evidence of teaching when faculty could connect assessment data to changes in the classroom.

Stipends and course releases. Though many faculty who were interviewed indicated there were no rewards for assessment work, some noted they received stipends (ranging from $250 for reviewing artifacts or participating in assessment workshops to more than $10,000 for summer salary) or course releases (ranging from one course buyout to "a 60% release") for their work in assessment. The faculty member receiving a $10,000 stipend for their work as an assessment coordinator was an outlier among faculty; stipends for assessment responsibilities were much lower among the other faculty interviewed. For those that received a stipend and/or course release, it was considered a reward and a motivator for continuing to do the work, even if the amount of the stipend was incongruent with the amount of work required. One participant noted that the stipend was an incentive to stay involved, saying "One of the reasons I stayed involved in that assessment committee so long is that up until about two years ago, the chair and the vice-chair got paid summer salary [because it was a lot of work]."

Food and drinks. Faculty also described having food and drinks at assessment meetings as a reward or incentive. One participant indicated that they have assessment "data parties" that were catered and another indicated that faculty who help with assessment activities are put into a drawing for their department to receive a catered lunch. Three participants noted that at events where faculty were expected to participate in assessment work (such as data parties, "assessment day," or peer reviews), there would be food. For example, one participant indicated that when they have an assessment day, it is the "expectation for all the faculty [to participate]…we do try to provide coffee and lunch. You know anything edible or drinkable they can give is, I think, appreciated." Another noted, "When we do the peer review it's a four-hour event, but we have refreshments. We have wine at the end. It's a nice event for everybody and you sit with different departments." Even "small" refreshments were appreciated; one participant said that at assessment meetings "there were…snacks and fizzy drinks and all that, and so it's like anything small like that…is appreciated."

P/T consideration. Faculty generally described assessment work as a component of their service, but the influence of service on P/T differed by participant and their institution's culture. Like the administrators surveyed in phase one, some faculty we interviewed indicated that service was not as important as teaching and research in P/T. For example, one participant said, "It's simply a service line item." However, others considered their assessment work to be a more valuable or visible part of their P/T documents. One noted, "It is a very visible service. I'm going up for full professor this year and that service kind of solidified my campus obligations for that part of my work." Another said, "If [faculty] are on the assessment committee, that's recognized as a major committee so you don't have to be on…one of the other major committees. So, it is recognized that way and seen as a big piece of service." Yet another participant echoed the "big commitment" of assessment as service:

> So part of my…tenure file of my institution–I mean there's a big emphasis on service. You know it's¬–it's a teaching college but also you know there's like–how have you contributed to the school? Having the assessment committee is a big one. It shows commitment to the institution and sort of the big–like taking on the big picture. You know working with different divisions and departments and…so it's under service.

Three participants mentioned that they used their in-class assessment work to provide evidence of teaching effectiveness but did not include assessment work outside the classroom

in the P/T documentation. In particular, assessment work seemed to be helpful for evidence of teaching when faculty could connect assessment data to changes in the classroom. For example, one participant said, "I actually write about how I've changed my course according to the assessment for my teaching evaluations as well." Another indicated, "[it's] under teaching you know…obviously I'm gonna put like different types of assessments that I've developed for classes or that I've collected from my students obviously you know our evaluations are important."

Four participants detailed how they represented assessment work as scholarship for P/T. Sometimes, doing so seemed to be relatively easy for participants, mostly because the assessment work looked like "normal" entries on a CV. For example, one participant said:

> I'd use all my assessment stuff–the publications, my role in leadership roles in assessment organizations, presenting, conferences, things like that–definitely have used those in and I think they would count for others as well–that's part of their scholarship.

Another noted, "I do have a separate section at the end [on my CV] where I put assessment conference presentations and I've got those two or three in there so that's in my research section of my vita because that's where all my presentations are."

Other participants indicated representing assessment work as scholarship was more difficult at their institutions. One participant recalled:

> I can tell you when I was going up for associate professor, we had very direct conversations about whether scholarship in teaching and learning counted as research within the department and then at the faculty and at my dean level as well. And so, I got it in writing that it counted equally to your traditional, you know, our lab science research stuff.

Another participant said:

> I think that there was a portion of faculty and department chairs and some of those, including me, were able to use that work to get P/T as well…It was not an easy argument to make, but it was made successfully by several people, not just myself.

Lastly, one faculty member who had come to be well-connected to assessment work on his campus for a number of years indicated that he created a separate section for assessment work on his CV and in P/T. He recalled:

> I would always squeeze it into my service. [After] about four or five years, I'm not sure, it became a separate category. That really makes it much more noticeable on your appraisal.

Later in his interview, he mentioned that he felt being involved in assessment on his campus was important to his P/T journey:

> It's always mentioned in the appraisal, and I feel like it enhanced it some. I mean, again, it's completely immeasurable. I don't feel like it was a nonfactor. Certainly, if I had spent that time doing more publications I'd been in a better position for raises because publications and grants is really what gets rewarded most of the time, but I feel like it's not insignificant. It's small, but it's not insignificant.

General campus exposure. Two participants noted that the exposure they gained from being a part of assessment work was helpful for them in their careers beyond P/T. One participant had recently transitioned from a faculty role to an administrative one, and she noted:

> I don't know if I would have been considered for this position had I not [had] the successful experience in a campus program review. I don't know, maybe I would have, but I think this did help–the fact that they knew that I could handle the work.

Another participant described the opportunity to connect with campus administrators through assessment work:

> I thought [my work in assessment] gave me some good exposure, like the provost would know who I am and things like that. So…going up for tenure, the provost knew who I was and knew about the work I was doing. So I think it was the icing on the cake type thing.

CV Document Analysis

The 11 faculty members who were interviewed also submitted their CVs and researchers searched for mentions of assessment work in the different sections of their CVs (see Table 2). Nine of 11 participants included a reference to assessment work in the "Service" portion of their CVs and 8 of 11 participants had conducted presentations related to assessment.

Table 2
Faculty CV Content Analysis

CV Category	Experience	Publications	Presentations	Service
Total (n=11)	4	3	8	9

Only one interview participant did not list any assessment work on her vita. When asked about it, she said, "I didn't even think of that…I'm going to put it on [my CV] now.

Discussion

Consistent with existing literature, both administrators and faculty participants in this study acknowledged that few rewards exist to promote faculty engagement in assessment work. Administrators acknowledged the lack of rewards and incentives as a barrier to engaging faculty deeply in this work. Faculty and administrators both related a lack of faculty engagement in assessment work to a disconnection between the work and the current institutional rewards structure. However, most faculty who were interviewed had a positive outlook on assessment work (consistent with Ebersole's 2009 findings) and included assessment work in the service section of their CVs and/or had conducted presentations about assessment.

All of the faculty interviewed discussed their work in relation to P/T, typically as a visible part of their service, which is weighed differently at different types of institutions. One participant even noted, "I could decrease or increase my service and I don't think it would have any impact on my appraisals." Thus, while faculty in the study indicated that service was the most common way assessment work was categorized for P/T documents (and CV analysis supported this), some felt that it wasn't an accurate reflection of the amount of work involved. For most participants, the weight assessment work was given in the P/T process depended on the committee and that was often influenced by institutional culture and/or the amount of experience committee members themselves had with assessment. Participants felt that categorizing assessment work as service was not an adequate recognition of the amount of intellectual labor necessary to engage in and conduct assessment work. Three faculty participants were able to increase the impact of assessment work on their CVs by counting it as scholarship by presenting at conferences or publishing findings. Both administrators and faculty in this study indicated that connecting assessment work to SoTL was the most successful way to represent assessment as scholarship. This connection to SoTL also links the findings of this study to that of Boyer (1990), who advocates for the expansion of research and scholarship to include assessment, as well as to the work of Cain and Hutchings (2015), who emphasize the importance of tying assessment work to three items of importance for faculty: student learning, their research, and P/T.

The issue of recognizing (or not) assessment work in the P/T process may reflect the institution's assessment climate in general.

Most faculty participants reported getting no monetary rewards for coordinating assessment activities. Given the amount of time involved in overseeing assessment work, participants felt monetary compensation was appropriate but not always a feasible part of

their department or university's budget. Three participants reported getting course releases or buyouts to allow them to engage in assessment work; while this was not a monetary payment, faculty viewed it as compensation for time, which was appreciated. With one exception, though, faculty who did receive compensation or course releases for participating in assessment work did not see the compensation as equivalent to the amount of time or level of responsibility required.

Participants noted that engaging in assessment work allowed them to network and get to know campus leaders, which participants saw as important exposure for future career opportunities. Two participants mentioned that they felt their work in assessment was an important consideration for getting other jobs on campus because campus administrators knew them through their involvement in assessment activities. Being publicly congratulated for assessment work or being given awards recognizing their contributions to the university was also valued. Finally, five participants mentioned that providing refreshments at assessment events was motivating and appreciated.

Faculty engaging in assessment work in this study were most motivated by and appreciative of recognition and rewards that seemed to truly value their labor.

This research builds upon previous research on the role of faculty in assessment work. There is not yet a comprehensive conceptual framework for describing the role of faculty in assessment or how institutions reward or recognize faulty for assessment work. Instead of building upon a conceptual framework that doesn't yet exist, we used current and previous literature as a basis for asking how faculty who engage in assessment work are rewarded and recognized. The results of this study indicate that the issue of recognizing (or not) assessment work in the P/T process may reflect the institution's assessment climate in general. For example, categorizing most assessment work as service did not seem to accurately reflect the work involved or the importance of the work itself. Further, most participants in this study indicated that the weight assessment work was given in the P/T process depended on institutional culture and the amount of experience committee members themselves had with assessment.

Limitations

The small sample size of the survey of campus administrators who oversee assessment work is a limitation of this study. Given the small sample size, it is unclear how representative the results are of the perceptions of college and university administrators who oversee assessment work. Additionally, four-year colleges were overrepresented in the survey sample and community colleges were underrepresented. Among interview participants, there were significantly more four-year college faculty included versus two-year faculty. Finally, the use of snowball sampling for the faculty participants may result in limitations in conclusions.

Future Research and Recommendations

There is still much to understand about the role of faculty in assessment work. This paper does not delineate rewards or recognition by institution type, but it is possible, for instance, that community colleges reward and recognize assessment work among faculty differently than universities with high research activity, and future research should examine those potential differences. There may also be differences among rewards for faculty who do assessment work in different fields. For instance, do faculty in education receive more recognition for their assessment activities, as opposed to faculty members in history or biology?

Additional research is needed to examine how and when faculty use assessment work successfully in P/T. Through the interviews in this study, it appears that there is potential to connect assessment activities to P/T (perhaps even outside of the category of service) but administrator surveys indicate institutions lack policies to guide this practice. Faculty interviews reveal that two participants successfully engaged in direct advocacy at their institutions for their assessment work to be considered as scholarship during P/T. Using the participants in this study as a model, faculty involved in assessment work can connect their contributions to teaching, research, and service during P/T to help others see the connection.

Administrators and faculty can collaborate to create rewards and incentives for engaging in assessment work that motivate faculty. Specific policies that permit and encourage

the use of the scholarship of assessment (for instance, presenting at conferences or publishing assessment-related research) in P/T is one way to tie assessment work to the current faculty rewards structure. Administrators interested in supporting faculty assessment work can also share this research with faculty and ask for feedback on how they would like to be rewarded for assessment work.

As briefly mentioned in the discussion, it seems that rewards and recognition for faculty contribution to assessment work may be linked to institutional culture. Future research should explore how institutional culture and climate and recognition of assessment work in P/T may be connected.

Conclusion

The findings from this paper address a gap in the literature regarding how faculty are recognized and rewarded for assessment work. Participants acknowledged that there are few tangible rewards for faculty who engage in assessment work. Faculty engaging in assessment work in this study were most motivated by and appreciative of recognition and rewards that seemed to truly value their labor. Faculty participants recognized that monetary compensation was not always possible; however, providing food, publicly recognizing assessment efforts, and valuing assessment contributions in the P/T process were emphasized.

AUTHORS NOTE

The authors would like to acknowledge the contribution of Andre Foisy who supported this study by participating in the initial data collection.

References

Allen, M. J. (2004). *Assessing academic programs in higher education*. San Francisco, CA: Anker.

Allen, M. J. (2006). *Assessing general education programs*. San Francisco, CA: Anker.

American Association for Higher Education (1996). *Nine principles of good practice for assessing student learning*. Retrieved from http://teaching.uncc.edu/sites/teaching.uncc.edu/files/media/files/file/AssessmentAndGrading/9Principles.pdf

Boyer, E. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Boyer, E. (1996). From scholarship reconsidered to scholarship assessed. *Quest, 48*(2), 129–139.

Bresciani, M. J. (2006). *Outcomes-based academic and co-curricular program review: A compilation of institutional good practices*. Sterling, VA: Stylus.

Bresciani, M. J. (2011). Identifying barriers in implementing outcomes-based assessment program review: A grounded theory analysis. *Research and Practice in Assessment, 6*, 5-16.

Cain, T. R., & Hutchings, P. (2015). Faculty and students: assessment at the intersection of teaching and learning. In G. Kuh, S. Ikenberry, N. Jankowski, T. Cain, P. Ewell, P. Hutchings, & J. Kinzie (Eds.), *Using evidence of student learning to improve higher education* (pp. 95-116). San Francisco, CA: Jossey-Bass.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.) Thousand Oaks, CA: Sage.

Crossley, M., & Wang, L. (2010). Learning by doing: An experience with outcomes assessment. *University of Toledo Law Review*, 41, 269–282.

Driscoll, A., & Wood, S. (2007). *Developing outcomes-based assessment for learner-centered education: A faculty introduction*. Sterling,VA: Stylus.

Ebersole, T. E. (2009). Postsecondary assessment: Faculty attitudes and levels of engagement. *Assessment Update, 21*(2), 1–13.

Ewell, P. (2002). An emerging scholarship: A brief history of assessment. In T. Banta (Ed)., *Building a scholarship of assessment* (pp. 3–25). San Francisco, CA: Jossey-Bass.

Ewell, P., Paulson,K., & Kinzie, J. (2011). *Down and in: Assessment practices at the program level*. Champaign, IL: The National Institute for Learning Outcomes Assessment. Retrieved from http://www.learningoutcome assessment.org/DownAndIn.htm

Greene, J. C., Caracelli, V., & Graham, W. (1989). Toward a conceptual framework for mixed-method evaluation designs. In V. Plano Clark and J. Cresswell (Eds.), *The mixed methods reader* (pp. 121-149). Thousand Oaks, CA: Sage Publications.

Grunwald, H., & Peterson, M. W. (2003). Factors that promote faculty involvement in and satisfaction with institutional and classroom student assessment. *Research in Higher Education, 44*(2), 173–204.

Hutchings, P. (2010). *Opening doors to faculty involvement in assessment*. Urbana, IL: The National Institute for Learning Outcomes Assessment. Retrieved from http://www.learningoutcomeassessment.org/documents/PatHutchings_000.pdf

Kezar, A., & Maxey, D. (2014, July). *Student outcomes assessment among the new non-tenure-track faculty majority*. Urbana, IL: The National Institute for Learning Outcomes Assessment. Retrieved from http://www.learningoutcomesassessment.org/documents/OP21.pdf

Kuh, G. D., Ikenberry, S. O., Jankowski, N.A., Cain, T. R., Ewell, P. T., Hutchings, P, & Kinzie, J. (2015). Making assessment matter. In G. Kuh, S. Ikenberry, N. Jankowski, T. Cain, P. Ewell, P. Hutchings, & J. Kinzie (Eds.), *Using evidence of student learning to improve higher education* (pp. 220–236). San Francisco, CA: Jossey-Bass.

Light, T. P., Chen, H. L., & Ittelson, J. C. (2012). *Documenting learning with ePortfolios: A guide for college instructors*. San Francisco, CA: Jossey-Bass.

MacDonald, S. K., Williams, L. M., Lazowski, R. A., Horst, S. J., & Barron, K. E. (2014). Faculty attitudes toward general education assessment: A qualitative study about their motivation. *Research and Practice in Assessment, 9*, 74–90.

Patton, M. (2002). *Qualitative research and evaluation methods* (3rd ed.) Thousand Oaks, CA: Sage Publications.

Reder, M., & Crimmins, C. (2018). Why assessment and faculty development need each other: Notes on using evidence to improve student learning. *Research and Practice in Assessment, 13*, 15–19.

Saldana, J. (2015). The Coding Manual for Qualitative Researchers (3rd ed.). Los Angeles, CA: Sage Publications.

Secret, M., Leisey, M., Lanning, S., Polich, S., & Shaub, J. (2011). Faculty perceptions of the SoTL: Definition, activity level and merit considerations at one university. *Journal of the SoTL, 11*(3), 1–20.

Smith, E. E. (2017). Using the inquiry process to promote faculty engagement in assessment. *Intersection, 1*(3), 15–22.

Stevenson, J. F., Finan, E., & Martel, M. (2017). Measuring assessment climate: A developmental perspective. *Research and Practice in Assessment, 12*, 41–57.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research*. Los Angeles, CA: Sage Publications.

Turner, S. F., Cardinal, L. B., Burton, R. M. (2017). Research design for mixed methods: A triangulation-based framework and roadmap. *Organizational Research Methods, 20*(2), 243–267.

Yancey, K. B. (2009). Reflection and electronic portfolios: Inventing the self and reinventing the university. In D. C. Cambridge, B. Cambridge, and K. Yancy (Eds.), *Electronic portfolios 2.0: Emergent research on implementation and impact* (pp. 5–16). Sterling, VA: Stylus.