

# RESEARCH & PRACTICE IN ASSESSMENT

VOLUME SIX | SUMMER 2011  
[www.RPAJournal.com](http://www.RPAJournal.com)  
ISSN #2161-4210

A PUBLICATION OF THE VIRGINIA ASSESSMENT GROUP





## Editorial Board

Executive Editor  
 Joshua Brown  
 Director of Student Affairs Assessment  
*Liberty University*

Editor  
 Katie Busby  
 Director of Institutional Assessment  
*Tulane University*

## Consulting Editors

Robin Anderson  
 Associate Professor and Associate Director  
 Center for Assessment and Research Studies  
*James Madison University*

Suzanne L. Pieper  
 Associate Director of Assessment for  
 Distance Learning  
*Northern Arizona University*

Marilee J. Bresciani  
 Associate Professor of Postsecondary  
 Education Leadership  
*San Diego State University*

P. Jesse Rine  
 Center for the Advanced Study of  
 Teaching and Learning in  
 Higher Education  
*University of Virginia*

Dorothy Doolittle  
 Professor of Psychology  
*Christopher Newport University*

Pamela Steinke  
 Director of Research, Planning  
 and Assessment  
*Meredith College*

Ghazala Hashmi  
 Professor of English  
 Coordinator, Quality Enhancement Plan  
*J. Sargeant Reynolds Community College*

John T. Willse  
 Assistant Professor of Educational  
 Research Methodology  
*University of North Carolina at Greensboro*

Hillary R. Michaels  
 Independent Consultant in  
 Educational Evaluation

Vicki L. Wise  
 Coordinator for Assessment and Research  
*Portland State University*

Deborah L. Moore  
 Director, Office of Assessment,  
 Evaluation, & Accreditation  
*Christopher Newport University*

## Ex-Officio Members

Ling Whitworth  
 President  
 Virginia Assessment Group  
*Longwood University*

Kathryne Drezek McConnell  
 President-Elect  
 Virginia Assessment Group  
*Virginia Tech*

## RESEARCH & PRACTICE IN ASSESSMENT

The goal of *Research & Practice in Assessment* is to serve the assessment community as an online journal focusing on student learning outcomes in higher education. 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 blind review process. Approximately fifty percent of submissions are accepted for issues that are published twice annually.

## TABLE OF CONTENTS

### Virginia Assessment Group 2011 Annual Conference

#### Leading and Learning through Assessable Moments: Engaging to Improve

Wednesday, November 16th – Friday, November 18th, 2011  
Crowne Plaza Williamsburg at Fort Magruder

Submission deadline for conference proposals is August 31, 2011

For more information visit [www.virginiaassessment.org](http://www.virginiaassessment.org)

#### Call for Papers

*Research & Practice in Assessment* is currently soliciting articles and reviews (book, media, or software) for its December 2011 issue. Work submitted for review in RPA may be related to various higher education assessment themes; whereas articles submitted for peer review should address topics related to student learning outcomes. Contributions will be evaluated by the RPA editorial board and should be submitted no later than September 1, 2011. Items should comply with the RPA Submission Guidelines and be sent electronically to :

[webmaster@virginiaassessment.org](mailto:webmaster@virginiaassessment.org)



The *RPA* journal is now  
available for download for iPad!

VIRGINIA ASSESSMENT GROUP  
[www.virginiaassessment.org](http://www.virginiaassessment.org)

For questions about new or current subscriptions, or membership opportunities email [webmaster@virginiaassessment.org](mailto:webmaster@virginiaassessment.org). We welcome members and non-members to join our community. If you would like to be informed of Virginia Assessment Group events, please contact us and we will add you to our distribution list.

Publication Design by Patrice Brown  
[www.treestardesign.com](http://www.treestardesign.com)  
Layout Placement Design by Kyle Doyon  
Copyright © 2011

#### 4 FROM THE EDITOR

#### ARTICLES

#### 5 **Identifying Barriers in Implementing Outcomes-Based Assessments Program Review: A Grounded Theory Analysis**

- Marilee J. Bresciani

#### 17 **Measuring Students' Cognitive Engagement on Assessment Tests: A Confirmatory Factor Analysis of the Short Form of the Cognitive Engagement Scale**

- Whitney Smiley and Robin Anderson

#### 29 **Can Institutions Have Quality Programming Without Utilizing a Systematic Outcomes-Based Assessment Process?**

- Lauren Weiner and Marilee J. Bresciani

#### SPECIAL FEATURE

#### 36 **Reflective (Ac) Counting: Institutional Research, Evaluation, & Assessment in a Time of Cholera**

- Walter F. Heinecke

#### BOOK REVIEWS

#### 40 **Review of *Academically Adrift***

- Kathryne Drezek McConnell

#### 43 **Review of *Higher Education Assessments***

- Tisha M. Paredes

#### RUMINATE

#### 46 **Standardized**

- Casey Templeton

#### 47 GUIDELINES FOR SUBMISSION

## FROM THE EDITOR

**T**his fifth volume of *Research & Practice in Assessment* marks the first issue for the newly established editorial team led by myself and Katie Busby of Tulane University. Although it is delivered in both an updated form and format, these efforts would not have been possible without the foundational efforts of the previous RPA editors - Robin Anderson, Keston Fulcher, and Allen DuPont. The release of this issue continues the scholarly assessment tradition established by the previous team of editors, yet it unveils new aspects of RPA including our logo, aesthetic layout, and content areas including the special feature, reviews (book, media or software), and a concluding section we have entitled "Ruminate".

It is our desire that the diverse perspectives included in this issue will encourage a rich dialogue on education assessment. As such, the issue begins with a qualitative piece by Bresciani who utilizes a grounded theory approach to investigate whether the use of outcomes-based program reviews have an impact on institutional reward structures. This is followed by Smiley and Anderson who use factor analysis to investigate whether assessment instruments might be improved by incorporating facets of cognitive engagement. Weiner and Bresciani conclude the peer-review section with a comparative piece that analyzes program quality differences and whether outcomes-based assessments are necessary to demonstrate such.

The special feature for this issue is authored by a critical education scholar, Walter Heinecke, who asks assessment professionals and researchers to consider the extent to which numbers and measurement categories are a social construction as well as their possible political implications. Then, McConnell provides readers with a review of *Academically Adrift* (Arum & Roksa, 2011), while Paredes offers a review of *Higher Education Assessment* (Kramer & Swing, 2010). This issue of RPA concludes with a photograph by Casey Templeton entitled, "Standardized" in a section that aims to integrate education assessment and the arts.

The overhaul of RPA was accomplished through the diligent efforts of many individuals. These persons deserve mention for their consistent contributions over the past six months and include: Patrice Brown, Alysha Brubaker, Katie Busby, Kyle Doyon, Terrell Perry, Tisha Paredes, Jesse Rine as well as the blind reviewers and published authors within this issue. Please consider offering us your feedback or critique as with any effort there is always room for improvement. More importantly, we hope you would consider extending the greater assessment dialogue by submitting your scholarly piece to *Research & Practice in Assessment*.

Regards,



Joshua Brown  
Liberty University

## **Abstract**

The purpose of this grounded theory study was to identify the typical barriers encountered by faculty and administrators when implementing outcomes-based assessment program review. An analysis of interviews with faculty and administrators at nine institutions revealed a theory that faculty and administrators' promotion, tenure (if applicable), and/or renewal of contracts are often not dependent on whether they use results from outcomes-based assessment program review to improve their students' learning and development.

### **Author**

Marilee J. Bresciani, Ph.D.  
San Diego State University

### **Email**

mbrescia@mail.sdsu.edu

# **IDENTIFYING BARRIERS IN IMPLEMENTING OUTCOMES-BASED ASSESSMENT PROGRAM REVIEW: A GROUNDED THEORY ANALYSIS**

While conversations proposing standardized testing within higher education abound (Allen & Bresciani, 2003; Department of Education (DOE), 2006; Ewell, 1997a, 1997b; Ewell & Jones, 1996; Maki, 2004; Palomba & Banta, 1999), proponents of outcomes-based assessment program review are still applauding the value and extent that the process can be used to inform decisions to improve student learning and development (Bresciani, 2006; Bresciani, Zelna, & Anderson, 2004; Huba & Freed, 2000; Maki, 2004; Mentkowski, 2000; Palomba & Banta, 1999; Suskie, 2004). As such, practitioners of outcomes-based assessment continue to seek various ways to meaningfully engage in outcomes-based assessment program review in order to find ways to improve student learning and development.

Even so, there are many others who do not believe the process adds value to their day-to-day teaching or administrative duties (Banta, 2002; Wergin, 1999). Regardless of whether individuals agree upon the level of value that engaging in outcomes-based assessment may bring to improving student learning and development, many instructors, administrators, and scholars still experience barriers in the implementation of the process.

The purpose of this grounded theory study was to identify the typical barriers encountered by faculty and administrators at a variety of institutional types when implementing outcomes-based assessment program review. While the intent of the study was to identify barriers and explore strategies that institutions use to address those barriers, a theory emerged as to why the practice of outcomes-based assessment is not pervasive even in institutions whose leadership emphasizes the importance of such a process to improve student development and learning.

## **Literature Review**

### **The Common Institutional Barriers**

Research has been conducted to illustrate the common barriers to implementing outcomes-based assessment. The reasons that outcomes-based assessment is not pervasively practiced or practiced at all are often classified into three categories: (a) time, (b) resources, and (c) understanding of assessment (Banta, 2002; Bresciani, 2006; Bresciani et al., 2004; Palomba & Banta, 1999; Suskie, 2004; Upcraft & Schuh, 1996).

As these categories are dissected further, additional reasons are revealed.

**Time.** Research posits that the manner in which one allocates time is influenced by how one prioritizes one's values (Argyris & Schon, 1978; Argyris & Schon, 1996; Dalton, Healy, & Moore, 1985; Sandeen, 1985). As such, human beings, regardless of their profession, will allocate their time that is devoted to work based on what they value or what they are told to value by those responsible for evaluating job performance (Argyris & Schon, 1978; Argyris & Schon, 1996; Bass, 1985; Bass & Avolio, 1994). Furthermore, while it can be assumed that all people have been given the same amount of time, all do not have the same number of priorities or level of responsibilities pressing upon their time. Thus, it is uncertain as to how decisions are made in accordance to varying number of priorities or responsibilities that press upon the amount of time that each person has to spend.

Certainly, in addition to personal and professional values, personality styles and time management strategies play a role in how people choose to prioritize their work projects in the time that they have allocated to their profession (Hackman, 1990; Kirkpatrick, 1993). The manner in which a person is evaluated and the criteria applied to personnel review also may influence how persons allocate their time at work (Hackman, 1990; Petrini & Hultman, 1995).

**Resources.** Resources have been presented as a reason that people do not engage in outcomes-based assessment including (a) the cost of providing professional development to faculty and administrators in order for them to learn how to engage in quality outcomes-based assessment, (b) the cost of the time re-allocated from actually teaching to the evaluation of teaching or from delivering the program to its evaluation, and (c) the cost of providing retreats so that faculty and administrators can actually reflect on what the outcomes-based assessment data are telling them about their program or curriculum. In addition, there is the cost of the administration and analysis of the evaluation tools used in outcomes-based assessment, as well as the cost of the improvements recommended for the program as the data suggests (Bresciani, 2006; Bresciani et al., 2004; Palomba & Banta, 1999; Suskie, 2004; Upcraft & Schuh, 1996).

The actual costs of implementing outcomes-based assessment often go uncalculated. In an attempt to determine actual costs of engaging in outcomes-based assessment, or rather, the attempt to evaluate the evidence of student learning and development, administrators are unsure as to whether to place the costs in the instructional category, an institutional research category, or an unfunded mandate category (Addison, Bresciani, & Bowman, 2005; Bresciani, 2006). Furthermore, the start-up costs of educating personnel to learn how to implement effective, efficient, and enduring outcomes-based assessment are often never allocated (Palomba & Banta, 1999; Upcraft & Schuh, 1996). Because the actual cost of engaging in outcomes-based assessment has not been systematically calculated, it is difficult to determine whether the perceived or actual costs of professional development are off-set by improved student learning.

**Understanding of assessment.** Higher education apparently has been fraught with "flavor of the day" processes and reporting initiatives, and institutional memory is often long about these fads (Banta, 2002; Palomba & Banta, 1999; Petrini & Hultman, 1995; Upcraft & Schuh, 1996). As such, faculty and administrators are often wary of anything else that comes along in an apparently pre-packaged version or with the threat of an unfunded mandate. While outcomes-based assessment has been around in one form or another for quite some time (Banta, 2002; Bresciani, 2006; Palomba & Banta, 1999; Upcraft & Schuh, 1996), the assumption that it is really here to stay is understandably questioned because the manner in which outcomes-based assessment has been labeled has changed over the years.

Many administrators and faculty simply do not believe that outcomes-based assessment is designed to be a systematic process to improve student learning and

**"Because the actual cost of engaging in outcomes-based assessment has not been systematically calculated, it is difficult to determine whether the perceived or actual costs of professional development are off-set by improved student learning."**

development, rather than a process to sustain itself (Bresciani et al., 2004; Kuh, Kinzie, Schuh, & Whitt, 2005; Upcraft & Schuh, 1996). In addition, the increasing emphasis on accountability, using standardized testing and other performance indicators that often cannot be linked to what is actually occurring in the classroom or within the co-curricular causes further understandable confusion.

In reviewing the three primary published reasons that faculty and administrators do not engage in outcomes-based assessment, a question arises: If an institution claims in its mission statement that it values student learning and development, why is it that that statement does not translate into action such as an institutional expectation for the evaluation of how well that learning and development is done in a manner in which it can be improved?

Outcomes-based assessment has several definitions (Palomba & Banta, 1999). Regardless of which definition one is examining, the idea of continuous improvement is often a common element (Allen, 2004; Banta, 2002; Bresciani et al., 2004; Maki, 2004; Palomba & Banta, 1999; Suskie, 2004). Using continuous improvement in the definition, there is an assumption of purposeful planning for the delivery and evaluation of intended outcomes. In addition, the evaluation process is designed so the information gathered could be used to inform specific decisions about how the intended outcomes can be met at a greater level of quality for the group that was included in the evaluation.

There are several resources designed to assist faculty and administrators with implementation of outcomes-based assessment (Allen, 2004; Bresciani, 2006; Bresciani et al., 2004; Maki, 2004; Palomba & Banta, 1999; Suskie, 2004; Upcraft & Schuh, 1996), yet many faculty and administrators are still having difficulty meaningfully engaging in the process. Why is that?

The purpose of the grounded theory study was to identify the typical barriers encountered by faculty and administrators at a variety of institutional types when implementing outcomes-based assessment program review. While the intent of the study was to examine barriers and explore strategies that institutions use to address those barriers, a theory emerged as to why the practice of outcomes-based assessment may not be pervasive even in institutions whose leadership emphasizes the importance of such a process to improve student development and learning.

## Methodology

To better understand how faculty and administrators are challenged with implementing outcomes-based assessment program review, a qualitative method of inquiry was utilized because the researcher's intent was to uncover rich and descriptive "meaning" (Bogdan & Biklen, 1992). There are several methods by which meaning can emerge and many of them share the common goal of understanding the subject's perspective. Researchers using grounded theory attempt to generate a theory that is closely related to the context of that which is being studied (Strauss & Corbin, 1990).

In grounded theory, the researcher interviews subjects and examines documents, then returns to evaluate the transcripts and documents using open coding analysis in order to identify categories or properties about what is being studied. Next, the researcher returns to the field to interview more subjects, continues with open coding, but also begins to use axial coding to compare the interviews in order to understand the central phenomenon, such as the reason that faculty and administrators are not pervasively engaging in outcomes-based assessment. Axial coding involves the act of constantly comparing words and meanings in order to formulate some common themes across the data. In the case of this study, the purpose of axial coding is to identify categories or conditions that may be contributing to the subjects' inability to engage in outcomes-based assessment and to identify specific strategies, conditions, and contexts

that influence practice or in this case, the student affairs/services professionals' willingness to practice outcomes-based assessment. In the third stage of data analysis, selective coding, the researcher uses the results of open and axial coding to integrate categories in order to identify a theory that further explains the complexities of the research findings (Creswell, 1998).

## **Research Questions**

In an effort to explore how pervasive faculty and administrators were engaged in outcomes-based assessment program review, the following research questions guided the analysis of the case studies and interviews.

1. How pervasively are your faculty and administrators engaged in outcomes-based assessment program review?
2. What do they report as challenges in their ability to meaningfully and manageably engage in the process?
3. How do they address those challenges?

## **Selection of Sample**

Nine institutions were purposefully selected to participate in this study. All nine institutions were considered to be good practice or emerging good practice institutions in implementing outcomes-based assessment program review based on good practice criteria published by Bresciani (2006). There were three community colleges, three comprehensive institutions, and three research extensive universities. At least three faculty and three administrators were interviewed at each institution. At some institutions, due to the opportunity provided to the researcher, more faculty and administrators were interviewed. In addition, documents, such as personnel evaluation criteria and documents (when available), meeting minutes, faculty memos, and institutional websites were also reviewed.

**If an institution claims in its mission statement that it values student learning and development, why is it that that statement does not translate into action such as institutional expectation for the evaluation of how well that learning and development is done in a manner in which it can be improved?"**

## **Limitations**

Limitations of this study include the inability to verify among decision makers the degree to which contributions or improvements to student learning is factored into personnel evaluations. In many instances, criteria for the weight that is placed on such an evaluation could not be ascertained in faculty evaluations nor in evaluations of administrators. In addition, it was unclear, apart from the use of student evaluations of teaching effectiveness, what criteria for improvement or contributions to student learning and development were being used for faculty evaluations. Similarly, there was little evidence of criteria for contributions to student development being used for administrative evaluations.

Another potential limitation is that those who participated in the interviews were motivated to illustrate a positive or negative aspect of the extent that faculty and administrators are using outcomes-based assessment results to improve student learning and development at their institution.

## **Findings and Discussion**

In an effort to explore how pervasively faculty and administrators were engaged in outcomes-based assessment program review, the aforementioned research questions guided the grounded theory analysis of the case studies and interviews.

A grounded theory analysis of the interviews and documents revealed a theory as to why faculty and administrators' engagement in outcomes-based assessment program review may not be pervasive. Faculty and administrators' promotion, tenure (if applicable), and/or

renewal of contracts is often not dependent on whether they use results from outcomes-based assessment program review to improve their students' learning and development. Rather, the promotion, tenure (if applicable), and/or renewal of their contracts appears to be based on a level of productivity that evidence suggests may not directly be related to improved student learning and development. Dependent on institutional type, reviews appear to be focused on either (a) number of peer-reviewed research journals published, (b) number and funding level of grants received, (c) in-class instructional evaluations by students, and (d) level of outreach activity, and/or other performance activities pertaining to increased inputs or outputs that improve performance indicators used for funding but are not necessarily related to student learning and development. To further illustrate the emergence of this theory, the findings are presented and discussed under each research question.

### **Pervasive Involvement in Outcomes-based Assessment Program Review**

In regards to how pervasively faculty and administrators are engaged in outcomes-based assessment program review, the institutions that had leadership commitment to evidence-based decision making had faculty and administrators in every unit of their organization engaged in some level of systematic outcomes-based assessment program review. When asked how this came to be, one administrator's response was illustrative of the many others when he said:

We didn't arrive at this level of involvement overnight. It took years of consistent commitment to building a process that faculty would recognize as first and foremost meaningful to them. They had to see that improvements [in student learning and development] would be made and that they were improvements that mattered to them and what they wanted students to learn.

While faculty, academic support specialists, and student support specialists were necessary to the establishment of a process that generated useful data to informing decisions that led to improvement, not all faculty and administrators were engaged in that process. For the most part, those that engaged in the process did so because they found the systematic practice of outcomes-based assessment valuable to improving student learning and development.

Many of these faculty and administrators reported that they had been implementing the process, albeit informally, long before the leadership called for all to be involved in it. As one participant stated, "I had been doing this [outcomes-based assessment] for years. I just didn't know it was called that [outcomes-based assessment]." Another illustrated the thought by sharing:

We [faculty] have discussions about how to improve student performance all of the time. We have them around the coffee pot; we have them in department meetings; we have them when we are standing in the halls; and we have them right after a student we are concerned about leaves our office. We care about improving student learning because we care, not because someone told us we had to care.

Such comments were common among the faculty represented in the institutions where outcomes-based assessment practice was prevalent. When asked why this was the case, one administrator summarized a common response among many participants when she said:

The faculty who are engaged in outcomes-based assessment in a systematic way would have done it anyway. These faculty experiment with inquiry processes to improve student learning because they want their students to improve. All we [administration] did was help them systematize it, provide some support so they could structure the process into program review and offer some release

time and professional development to get other faculty to discover its benefits. Our faculty don't do it because they are externally [external to the academy] or internally [internal to the academy] rewarded for doing it [outcomes-based assessment]. They do it [outcomes-based assessment] because it provides them with specific information to inform decisions that will improve student learning.

Huba and Freed (2000), Maki (2004), Mentkowski and associates (2000), Suskie (2004), and Palomba and Banta (1999) have repeatedly illustrated that the primary motivation for faculty to engage in systematic outcomes-based assessment comes from the realization they discover, after trying it out, that it does indeed contribute to improved student learning and development. So, when one examines how pervasive faculty involvement is in outcomes-based assessment, does one assume that if faculty experience it and recognize its value that they will then systematically implement it? If so, how do you get the faculty and administrators who are not involved to become involved? As reported by one administrator:

I am not sure which comes first. It is the proverbial chicken and the egg. While well-respected faculty have designed the process and to some extent, I think they have peer-pressured other faculty into getting involved... [pause] Still, not every faculty member is involved. And I don't think they have to be, even though we want them to be.

While administrators and faculty commented on whether they felt the practice of outcomes-based assessment program review on their campuses was pervasive, there was no clear definition as to what pervasive meant and how many faculty or administrators made the practice "pervasive."

In the Merriam –Webster dictionary, pervasive means, "pervading or tending to pervade" ("Pervasive," 2007). If seeking the definition of pervading or pervade, you would discover that it means, "to become diffused throughout every part of" ("Pervading," 2007). If pervasive means getting every faculty member and administrator involved in outcomes-based assessment, then when it comes to how to get faculty involved, several ideas were shared by study participants. Many of the ideas shared range from hiring faculty and administrators who are able to do outcomes-based assessment as advertised for in position descriptions, to providing release time to engage in outcomes-based assessment, to providing other rewards and incentives for involvement, and clarifying expectations for involvement (Bresciani, 2006; Bresciani et al., 2004; Huba & Freed, 2000; Maki, 2004; Mentkowski, 2000; Palomba & Banta, 1999; Suskie, 2004).

**“...the fact of the matter is that the only way I can keep my job...the only way I can keep teaching is if I publish several articles in journals that only accept less than 10 percent of the submissions. I tried to explain this to a student once... when I could not see them because I had to get my research done.”**

Administrators desiring for their staff to become more involved in outcomes-based assessment simply stated that, "getting them [administrative staff] involved? That is easy. They are not faculty; you can just tell them [administrators], it has to be done and they do it. They [administrators] can't hide behind academic freedom." Even with such confidence, it was clear that not every administrator was engaged in outcomes-based assessment, even when the administrative leadership felt they had made the expectations for such involvement very clear.

If pervasive does not constitute getting every person involved in outcomes-based assessment but ensuring that at least someone in every aspect of the organization is engaged in outcomes-based assessment, how does one still manage to get people who are not currently involved to participate in the process?

### **Challenges to Engaging in the Process**

Participants in this study reiterated the common challenges to engaging in outcomes-based assessment: (a) time, (b) resources, and (c) understanding of assessment. For brevity's sake, the researcher will not expand on these challenges since they are reaffirmed in several publications (Bresciani, 2006; Bresciani et al., 2004; Huba & Freed,

2000; Maki, 2004; Mentkowski, 2000; Palomba & Banta, 1999; Suskie, 2004). What may be more compelling at this point is to discuss the barrier that participants, regardless of institutional type, felt was most difficult to address. This barrier is as follows.

Faculty and administrators' promotion, tenure (if applicable), and/or renewal of contracts is often not dependent on whether they use results from outcomes-based assessment program review to improve their students' learning and development. Rather, the promotion, tenure (if applicable), and/or renewal of their contracts appears to be based on a level of productivity that evidence suggests may not directly be related to improved student learning and development. Therefore it appears that the evaluation of personnel processes, particularly for faculty, dissuade faculty in engaging in extensive evaluation of student learning and development. As this faculty member illustrates:

Listen, it is not that I don't care about what my students learn and how well they learn it; I do. I really care. But the fact of the matter is that the only way I can keep my job...the only way I can keep teaching is if I publish several articles in journals that only accept less than 10 percent of the submissions. I tried to explain this to a student once... when I couldn't see them because I had to get my research done... [pause] I stopped myself [from telling the student this]. How do you explain that to a student who needs your help?

Several tenured, full professors, who are actively engaged in outcomes-based assessment on their campuses and who would be described by their senior administrators as the faculty who led the design and implementation of the process that systematically improved student learning affirmed the rewards process for publications. "You can't have junior ( untenured) faculty getting involved in documenting how well their students learn. You just can't. It takes too much time away from their research." These faculty expressed time and time again how responsible they were being by "protecting" assistant professors, so they could get their tenure. "They need to get their tenure, and then we will ask them to assess [student learning]", explained a full professor.

Faculty at less research intensive institutions felt the same pressure to publish first, and then to consider the evaluation of student learning second.

I hate it when people assume we don't care about student learning; that is not the case. It's just that I have seen some teachers go untenured [pause]...the ones who only talk about student learning and measuring how well they do it...if you don't pay attention to your research, you don't get to stay.

In two-year colleges where there is little to no research pressure, this concern is expressed slightly differently.

No, we don't have the research pressures that others have, but we get less time to plan our preps than the high school teachers get. Also, there is no time to sit and chat about we learned from our classroom assessment. I am gathering the data [about how our students learn], but we are so busy teaching, we have no time to talk about what our students are learning.

When these two-year college faculty were asked about the possibility of release time to reflect on student learning results in order to inform conversations where decisions could be made for improvements, some mentioned that their collective bargaining agreements were being interpreted by their union leadership in a manner that would dissuade them from doing this:

I tried to get a group of faculty together where we would talk about what we were finding out in our classrooms. I was visited by a union leader who discouraged me from doing this, telling me that this type of work was outside the scope of our collective bargaining agreement. I was confused by that remark but I haven't had time to look into it further.

This faculty member, along with others in the study, understood that hosting

such conversations was not in the scope of their duties and they would need additional compensation and possibly an additional contract to have such conversations.

Other two-year faculty members engaged in outcomes-based assessment reported that they are doing it as additional work for which they are neither recognized nor reimbursed:

If you look at how our workload is calculated, there is zero time allocated to the practice of outcomes-based assessment. We [my colleagues and I at other 2-year institutions] get evaluated based on how many FTEs [Full-Time Equivalent], SCHs [Scheduled Class Hours], seat hours, or continuing education hours we generate. That is what we hear about; that is what we understand we are to care about...just productivity of increased numbers. How can you find time to focus on student learning when that is all you hear?

Even though faculty of all institutional types may have painted a bleak picture, many of them also re-affirmed that regardless of whether they get evaluated for their contributions to student learning, they will do so, and they will use the data to discuss ideas for improvement with their colleagues. When asked how they manage to evaluate student learning and use the results, many responded with similar answers as this:

I work my ass off. I can't fall behind on publications. It doesn't matter what level [of professor] you are around here. The higher you go [in promotion and tenure], the more work you get; the more responsibility you have. Somebody should do a study of divorce rates and loss of custody of children among faculty and why it occurs. Now, that would be something to study.

Some faculty were able to use the results of their outcomes-based assessment work for publications and thus, were able to "kill two birds with one stone." But many more faculty said that the evaluation of their students' learning was not anywhere close to what their faculty peers would value as research in their discipline.

For many administrators, continuance of employment may be perceived as political or based on countable production and activity, rather than evidence of contributions to student learning and development:

Look at how I evaluate people, there is nothing in there asking them to show evidence of their organization's contribution to student learning or development...nothing. We do it anyway though, because we know it is what we are all about.

One other senior level administrator however, made his expectation very clear, "If you don't tell me how you know you are contributing to student learning, I will tell you that you just made you, and your department candidates for 'outsourcing.'" While this administrator made his expectations very clear, there was no clear evidence that his staff was evaluated and rewarded for the extent they could demonstrate improvements in student learning and development.

Scholars have written about the notion that lasting change cannot occur in higher education unless the rewards structure for making that change also follows suit (Banta, 2002; Cox & Richlin, 2004; Doherty, Riordan, & Roth, 2002; Eckel, Green, & Hill, 2001; Frazier & Frazier, 1997; Hutchings, 2001; Kreber, 2001; Maki, 2004). Given this line of research and the findings in this study, one may wonder if the institutional leadership who are committed to improving student learning and development can sustain the efforts if their personnel evaluation systems are not updated to reflect a change in organizational values.

**"Scholars have written about the notion that lasting change cannot occur in higher education unless the rewards structure for making that change also follows suit."**

## Addressing the Challenges

Apart from being able to arrive at a strategy to address the barrier that personnel

are not rewarded for the extent that they use evidence to improve student learning and development, the institutional expectation for improving student learning and development was still made clear at the institutions that participated in this study. As such, they were eager to share their solutions for addressing the three common barriers identified: (a) time, (b) understanding of assessment, and (c) resources.

**Time.** As previously mentioned, many participants discussed the fact that there was no new time to allocate to the reflection and improvement of student learning and development. Participants reiterated that time either had to be re-allocated from their personal time or from some other projects. As one administrator stated:

Everyone cares about students here and we are a small institution so many of us carry a lot of responsibilities. In order to get this [outcomes-based assessment] done, we just get it done. We work harder and smarter. That is just the way it has to be.

A faculty member at another institution represented a different approach:

We know that it [outcomes-based assessment] won't get done well if we don't re-allocate time to it. So, that is what we do. We have invested a lot of time in learning how to do this well. We have re-designed our faculty meetings so that we discuss the results and apply what we learned. We are not able to respond to all of the needs we see; we simply can't. But we do move forward with improvements every year. We are very proud of that.

Some common strategies discussed by most of the participants in this study included the following. (a) Re-allocation of time from the doing of the activity to the evaluation of the doing. This may mean investing in release-time from activities for both administrators and for faculty. (b) When an organization cannot provide release time, encourage faculty and administrators to engage in fewer activities so that they can reflect on the data and decisions that need to be made in order to improve learning and development. (c) Discuss results derived from outcomes-based assessment in a manner that is open and inclusive, the practice of which saves time when people are wondering how and why decisions are made. (d) Collaborate with peers to assess student learning, discuss the results, and make decisions. If you share your workload, which includes involving students in every aspect of the process, you save time because you are borrowing ideas from colleagues. They may be able to suggest solutions more quickly because they are not as invested in the history of what led you to do that which you now do and have recently discovered is not as an effective process as you had hoped.

**Understanding of assessment.** Similar to previous research findings, the need to understand what outcomes-based assessment is and why one should engage in it was prevalent (Bresciani, 2006; Bresciani et al., 2004; Huba & Freed, 2000; Maki, 2004; Mentkowski, 2000; Palomba & Banta, 1999; Suskie, 2004). In particular, participants spoke of the importance of balance in understanding how results would be used and the level of expectations for engagement in assessment among faculty, administrators and top level leadership:

It is a delicate balance. You need the expectations communicated from the leadership that this is a process that will help us systematically improve students' learning, but it needs to be communicated in a manner that allows faculty the freedom to develop the process that is most meaningful to them; a process where they can discover whether the results will actually be helpful in improving learning.

Another participant explained:

To keep this (outcomes-based assessment) from being seen as an unfunded mandate, we made sure to connect the required documentation to what we already were doing with planning and program review. Doing this was so

helpful to everyone involved; they could focus on the meaning of doing it, rather than griping about having to do it.

Participants illustrated the importance of making sure everyone, including top-level leadership could understand (a) what outcomes-based assessment is and why it is being required, (b) the connection of assessment to planning, (c) how to use the results both as a practitioner (e.g., administrator or instructor) and as a leader, and (d) how to connect the results to external benchmarks or indicators of success. However, in order to understand how to do all this, one participant remarked:

You really need to provide training or education or whatever you want to call it. You can't just expect that faculty and administrators are going to simply understand this just because you have shown them evidence of how the process can improve student learning. You really have to educate ...and that takes time...

A commitment to providing professional development takes time and it takes resources in the manner of investing in the professional development (Bresciani, 2006; Bresciani et al., 2004; Huba & Freed, 2000; Maki, 2004; Mentkowski, 2000; Palomba & Banta, 1999; Suskie, 2004).

**Resources.** While time can certainly be considered a resource, this section focuses on the types of resources that could be provided to faculty and administrators such as professional development to learn all aspects of the outcomes-based assessment process, including follow-up one-on-one assistance, assistance with documentation, and facilitated reflection for interpretation of results and decision-making. Most of the participants agreed that early investment in professional development was needed in order for the process to be sustainable and effective:

We invested a great deal of money in up front professional development. When we decided to do this, we actually followed a corporate model for ‘re-tooling’ our workforce. Now, since we have done this for a while, we really have more of a ‘train-the-trainer’ model so the departments absorb the professional development costs as they ‘orient’ faculty and staff to the way they do things.

Many scholars (Battino et al., 2006; Eddy, 2005; Kemp et al., 2006; Stanley, Watson, & Algert, 2005) discuss the importance of investing in faculty training and development in order to “re-tool the workforce” for changes in the industry such as ways to evaluate and improve student learning and development. These scholars, while not fully adopting a corporate model explain the importance of realizing that needs in higher education change, society changes, and students change; thus, to not invest in the re-equipping of our faculty and staff is irresponsible and short-sighted.

While the participants in this study found varying ways to provide and fund professional development, they all did so and continue to do so in the ways that make the most sense to each institution. However, the documentation of the process and the results and decisions still posed challenges. One participant illustrated this when she said:

I do care about how well my students learn and develop, but outcomes-based assessment takes a great deal of time to document. We can do this easily, but documenting it all? [pause] I am just not sure that the improvements that result from engaging in it are worth the time [invested in it].

When pressed for an explanation, this faculty member returned to the earlier expressed concern that what she really needed to document were the items that went into her promotion reviews; how she used outcomes-based assessment to improve student learning was not required in that paperwork.

### Recommendations

Clearly, more research needs to be conducted in order to determine the extent that faculty and administrators are not reviewed and promoted or tenured (if applicable) based on their contributions to student learning and development. Data are inconclusive.

**“You can’t just expect that faculty and administrators are going to simply understand this just because you have shown them evidence of how the process can improve student learning. You really have to educate...and that takes time...”**

Even so, it should be noted that when reviewing all personnel evaluation guidelines, none of these institutional processes asked for faculty and administrators to provide evidence, based on a systematic evaluation process, as to how student learning and development was improved. Furthermore, while faculty and administrators recognized that this information was not provided, there appeared to be very little discussion around how to change existing processes. Further research on why that may be the case would aid additional conversations on the matter.

Given that faculty and administrators may not be evaluated and subsequently rewarded for improving student learning and development, it appears that regardless of what tool (e.g., CLA, ETS) or what process (e.g., outcomes-based assessment, CQI) is used to evaluate student learning and development, pervasive improvement in student learning and development may not come about if faculty and administrators' promotion and review processes are not influenced by the use of outcomes-based assessment for improvement. On the contrary, if faculty and administrators responsible for designing learning and development activities are told what tools and what processes they have to use without being able to develop these tools and measurements themselves, it may be likely that given these findings, there will be even less motivation and subsequent reward to improving student learning and development. Certainly there may be less clarity in regards to how to specifically improve student learning (Maki, 2004; Suskie, 2004).

## References

- Addison, P. D., Bresciani, M. J., & Bowman, K. (2005, April 25). Peer comparison analysis of offices of assessment in higher education. *Net Results: NASPA's E-Zine for Student Affairs Professionals*. <http://www.naspa.org/membership/mem/nr/article.cfm?id=1494>
- Allen, M. J. (2004). *Assessing academic programs in higher education*. Bolton, MA: Anker Publishing.
- Allen, J., & Bresciani, M. J. (2003, January/February). Public institutions, public challenges: On the transparency of assessment results. *Change Magazine*, 35(1), 20-23.
- Argyris, C., & Schon, D. (1978). *Organizational learning*. Reading, MA: Addison Wesley.
- Argyris, C., & Schon, D. A. (1996). *Organizational learning II: Theory, method and practice*. Reading, MA: Addison-Wesley.
- Banta, T. W. (2002). *Building a scholarship of assessment*. San Francisco, CA: Jossey-Bass.
- Bass, B. M. (1990). *Bass and Slodzehill's handbook of leaders*. New York, NY: Free Press.
- Bass, B. M., & Avolio, B. J. (1994). *Improving organizational effectiveness through transformational learning*. Thousand Oaks, CA: Sage Publications.
- Battino, W., Clem, J., Caine, R. N., Reigeluth, C. M., Chapman, C., Flinders, D. J., & Malopinsky, L. V. (2006). New systems produced by systemic change. *TechTrends: Linking Research and Practice to Improve Learning*, 50(2), 51-56.
- Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education*. New York, NY: Allyn and Bacon.
- Bresciani, M. J. (2006). *Outcomes-based academic and co-curricular program review: A compilation of institutional good practices*. Sterling, VA: Stylus Publications.
- Bresciani, M. J., Zelna, C. L., & Anderson, J. A. (2004). *Assessing student learning and development. A Handbook for Practitioners*. United States: NASPA.
- Cox, M. D., & Richlin, L. (Eds.). (2004). *Building faculty learning communities.: New directions for teaching and learning*. San Francisco, CA: Jossey-Bass.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.
- Dalton, J., Healy, M., & Moore, J. (1985). Planning a comprehensive values education program. In J. Dalton (Ed.), *Promoting values development in college students* (monograph series, v.4, pp. 1-16). Washington, DC: NASPA.
- Department of Education (DOE). (2006). The Commission on the Future of Higher Education Draft Report of 8/9/2006. Retrieved August 10, 2006, from <http://www.ed.gov/about/bdscomm/list/hiedfuture/reports/0809-draft.pdf>.

- Doherty, A., Riordan, T., & Roth, J. (2002). *Student learning: A central focus for institutions of higher education*. Milwaukee, WI: Alverno College Institute.
- Eckel, P., Green, M., & Hill, B. (2001). On change V—Riding the waves of change: Insights from transforming institutions. Washington DC: American Council of Education.
- Eddy, P. L. (2005). Faculty development in community colleges: Surveying the present, preparing for the future. *Journal of Faculty Development*, 20(3), 143–152.
- Ewell, P. T. (1997a). From the states: Putting it all on the line—South Carolina's performance funding initiative. *Assessment Update*, 9(1), 9–11.
- Ewell, P. T. (1997b). Identifying indicators of curricular quality. In G. J. Gaff, L. J. Ratcliff, and Associates (Eds.), *Handbook of the undergraduate curriculum: A comprehensive guide to purposes, structures, practices, and change* (pp. 608–628). San Francisco, CA: Jossey-Bass.
- Ewell, P. T., & Jones, D. P. (1996). *Indicators of good practice in undergraduate education. A handbook for development and implementation*. Boulder, CO: NCHEMS.
- Frazier, A., & Frazier, M. A. (1997). *A roadmap for quality transformation in education*. Boca Raton, FL: CRC Press.
- Hackman, J. R. (1990). *Groups that work (and those that don't)*. San Francisco, CA: Jossey-Bass.
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses / Shifting the focus from teaching to learning*. Needham Heights, MA: Allyn & Bacon.
- Hutchings, P. (2001). *Opening lines: Approaches to the scholarship of teaching and learning*. Menlo Park, CA: Carnegie Publications.
- Kemp, J. E., Lee, S., Pascoe, D., Beabout, B., Watson, W., & Moore, S. (2006). Foundations for systemic change. *TechTrends: Linking Research and Practice to Improve Learning*, 50(2), 20–26.
- Kirkpatrick, D. L. (1993). Riding the winds of change. *Training & Development*, 47, 28–32.
- Kreber, C. (Ed.). (2001). Scholarship revisited: Perspectives on the scholarship of teaching. *New Directions for Teaching and Learning*, 86. San Francisco, CA: Jossey-Bass.
- Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2005). *Student success in college: Creating conditions that matter*. San Francisco, CA: Jossey-Bass.
- Maki, P. (2004). *Assessing for student learning: Building a sustainable commitment across the institution*. Sterling, VA: Stylus Publishing, LLC.
- Mentkowski, M. (2000). *Learning that lasts: Integrating learning, development, and performance in college and beyond*. The Jossey-Bass higher and adult education series. San Francisco, CA: Jossey-Bass.
- Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco, CA: Jossey-Bass.
- Pervading. (2007). In Merriam-Webster Online Dictionary. Retrieved October 11, 2007, from <http://www.merriam-webster.org/dictionary/pervading>
- Pervasive. (2007) In Merriam-Webster Online Dictionary. Retrieved October 11, 2007, from <http://www.merriam-webster.org/dictionary/pervasive>
- Petrini, C., & Hultman, K. E. (1995). Scaling the wall of resistance. *Training & Development*, 49, 15–18.
- Stanley, C. A., Watson, K. L., & Algert, N. E. (2005). A faculty development model for mediating diversity conflicts in the university setting. *Journal of Faculty Development*, 20(3), 129–142.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Suskie, L. (2004). *Assessing student learning: A common sense guide*. Bolton, MA: Anker Publishing.
- Upcraft, M. L., & Schuh, J. H. (1996). *Assessment in student affairs: A guide for practitioners*. San Francisco, CA: Jossey-Bass.
- Wergin, J. F. (1999, December) Evaluating department achievements: Consequences for the work of faculty. AAHE Bulletin.

## **Abstract**

Motivational theory is often used to develop strategies for boosting student effort on assessments, particularly in low stakes situations. Increasing students' cognitive engagement on such assessments may also impact student effort. However, before such interventions can be evaluated, a sound measure of cognitive engagement must be identified. This study examines the factor structure of a scale (CE-S) modified to measure students' cognitive engagement specifically on assessment tests. A 2-factor model of cognitive engagement supports the interpretation of two subscale scores. The relationship between these subscale scores and scores on measures of motivation and goal orientation further supports two separate subscales of cognitive engagement. Future research and implications for use of the CE-S in assessment practice is discussed.

### **Authors**

Whitney Smiley, M.A.  
James Madison University

Robin Anderson, Psy.D.  
James Madison University

**Email**  
smileywf@dukes.jmu.edu

# **MEASURING STUDENTS' COGNITIVE ENGAGEMENT ON ASSESSMENT TESTS: A CONFIRMATORY FACTOR ANALYSIS OF THE SHORT FORM OF THE COGNITIVE ENGAGEMENT SCALE**

As with K-12 institutions, higher education institutions are feeling the pressure from the state governing bodies to provide evidence that learning is occurring, in return for the hard-earned tax dollars the states dispense to colleges and universities. In response, many higher education institutions are designing methods to assess student learning and development as evidence of the effectiveness of their academic programs. These assessments are typically viewed as low-stakes for the students because there are no consequences regardless of how they perform. However, if institutions want to demonstrate what students are learning to stakeholders, students must be motivated to put forth effort on the test (Wise & DeMars, 2005). It often falls to assessment specialists to ensure that assessment data are collected in a meaningful way, especially in low-stakes situations.

While students may receive no direct consequences from their performance on such assessments, these tests often represent a high stakes situation for the institution. Failure to provide evidence that programs are effective could result in serious consequences at the hands of accrediting organizations and state governing bodies. It is of little surprise that the low-stakes nature for students on such assessments would make institutions skeptical about using findings inferred from low-stakes assessment data. Research findings indicating that low motivation hinders the validity of inferences made from student scores (Wise & Demars, 2005), further support such institutional concerns. Concerns regarding the impact of low motivation on assessment results have prompted assessment practitioners to employ motivational theory in an attempt to find ways to encourage students to put forth effort. However, relying on motivational theory alone may exclude other factors that play a role in student effort on low-stakes assessment.

One factor that is less understood is the role that cognitive engagement plays in student effort. Newmann, Wehlage, and Lamborn's (1992) definition of cognitive engagement, "the student's psychological investment in and effort directed toward learning, understanding, or mastering the knowledge, skills, or crafts that academic work is intended to promote" (p.12), is specific to academic work situations and is therefore relevant for assessment contexts. For example, students may put forth more effort on

assessments that they find more cognitively engaging. Thus, assessment specialists may be able to improve student effort by utilizing more cognitively engaging assessments. We expect that if students are more engaged the costs associated with taking the test (i.e. effort, time, etc.) will be reduced and students should get more out of the test, boosting the value they place on the assessment. As Wigfield and Eccles (2000) pointed out, value is a tradeoff between what students get out of the test and the costs associated with taking the test. This increased engagement and the resulting boost in value placed on the assessment may result in increased effort. However, these are empirical questions and current assessment practices have largely ignored cognitive engagement as an area of research.

## Cognitive Engagement

School and government policies have been put in place to require students to attend schools; however, engagement in academic settings is tough to mandate. Newmann et al. (1992) point out that disengaged students can disrupt the classroom, skip classes, fail to complete assignments, etc. However, the more typical disengaged student can come to class every day, complete all of their work, behave well, and yet have neither excitement nor commitment to the material. They may in turn lack mastery of the material. Of course, while attendance can be regulated, engagement cannot. In situations where attendance is regulated but engagement is lacking, students may become bored and uninvolved throughout the school day; in many cases, they might as well be absent (Newmann et al., 1992). Because of this, it is important to study cognitive engagement so that policy and practices can be developed to reduce the likelihood of such cognitive absences. This is especially important in low-stakes assessment testing situations where students are mandated to attend but cannot be mandated to engage. If students are not engaged while taking the test, institutions will have assessment results, but what inferences can we draw from these results?

The construct of cognitive engagement can be talked about in a myriad of ways. Appleton, Christenson, and Furlong (2008) reviewed several definitions of cognitive engagement and were able to classify the definitions into eight types: engagement, engagement in schoolwork, academic engagement, school engagement, student engagement, student engagement in academic work, student engagement in/with school, and participation identification. Measuring cognitive engagement during assessments would fall under the student engagement with academic work subtype.

Cognitive engagement in academic work has been defined by Marks (2000) as, “A psychological process involving the attention, interest, investment, and effort students expend in the work of learning” (pp. 154–155). Newmann et al. (1992) defined cognitive engagement in academic work as, “The student’s psychological investment in and effort directed toward learning, understanding, or mastering the knowledge, skills, or crafts that academic work is intended to promote” (p. 12). Both of these definitions involve psychological investment and effort. The Newmann et al. definition is the more specific one stating that the construct involves engagement for the purpose of mastering knowledge, skills, or crafts; whereas, Marks’ definition does not address the issue of purpose for engagement. The definition used by the current study more closely aligns with Newmann et al.’s definition. We are most interested in students’ psychological investment directed toward a specific academic event (assessment testing). Students may complete academic work and perform well without being engaged in mastery of material. In fact, a significant body of research indicates that “students invest much of their energy in performing rituals, procedures, and routines without developing substantive understanding” (Newmann et al., 1992, p. 12). Our understanding of cognitive engagement can be furthered by distinguishing among behaviors as on a continuum between deep and shallow engagement (Greene & Miller, 1996). Students who exhibit behaviors that allow them to master academic work are seen to have deep cognitive engagement, while students who exhibit behaviors such as rote memorization and rituals they perceive will help them do well without developing mastery of the material are demonstrating shallow engagement. In the context of assessment testing,

**“...assessment specialists may be able to improve student effort by utilizing more cognitively engaging assessments.”**

deeply engaged students will come in and make sure they read each answer carefully and try to formulate thoughtful answers while students who simply come in and provide vague, unrelated, or not well thought out answers, exhibit behaviors associated with shallow engagement.

To further understand how cognitive engagement may impact student performance, one must understand how cognitive engagement differs from other related constructs. For example, it is important to distinguish between cognitive engagement and motivation. Effort is incorporated into both of the above definitions of cognitive engagement. Motivation scales often include items designed to assess effort as a subscale of motivation (e.g. the Student Opinion Survey; Sundre, 1999). However, engagement implies more than motivation, although motivation is necessary for cognitive engagement. Motivation is more of a general trait; that is, one can be a motivated person without being engaged in a specific task (Appleton, Christenson, Kim, & Reschly, 2006; Newmann et al., 1992). However, cognitive engagement is context dependent. This can be shown in the research of Marks (2000) who found that students in his sample reported higher cognitive engagement behavior in their mathematics courses than in their social studies courses. Marks concluded that this difference shows that cognitive engagement can change across contexts, or in this case, educational experiences.

Another construct that might be confused with cognitive engagement is goal orientation. Goal orientation refers to the reason a person engages in an academic task. Initially, research was focused on two types of goal orientation: performance and mastery (Dweck, 1986; Nicholls, 1984). Performance goals involve competence relative to others whereas mastery goals are seen as competence related to task mastery. However, over time goal orientation has grown to include five different orientation types including the original two, as well as work-avoidance, performance-avoidance, and mastery-avoidance. These avoidance items are used to distinguish between people who want to perform well on a task, versus people who want to avoid performing badly at a task (Baranik, Barron, & Finney, 2010). In Newmann et al.'s (1992) definition of cognitive engagement, they make it clear that the goal of an engaged student is mastery of knowledge, which is a factor in goal orientation.

Consistent with Newmann et al.'s (1992) definition of cognitive engagement, Meece, Blumenfeld, and Hoyle (1988) found a significant relationship between goal orientation and engagement patterns. They found a strong positive correlation between the task mastery subscale of their goal orientation measure on the Science Activity Questionnaire (i.e. a child's goal to learn something new and understand his or her work, or learn as much as possible) and active cognitive engagement. Also, scores on the ego/social scale as well as the work-avoidant scale on the same measure correlated positively with superficial cognitive engagement. This research shows that while these constructs are highly correlated, they are also likely two separate constructs. The difference between these constructs is also contextual. Goal orientation refers to a general orientation toward learning (Meece et al., 1988) whereas cognitive engagement in academic tasks refers to a specific task and can change across tasks.

### **Problems with Measuring Cognitive Engagement**

As expressed above, cognitive engagement is an important construct to measure within the context of assessment practice because higher cognitive engagement could result in more effort exerted from students on low-stakes assessment tests. As Newmann et al. (1992) point out, simply attending an environment (assessment day, classroom, or computer lab) and completing necessary work (assessment tests) are not good indicators of cognitive engagement. Rather, engagement is a construct that is used to describe internal behaviors such as effort to learn and quality of understanding. In order to make valid inferences regarding students' level of cognitive engagement across different tasks, researchers must have a measure of cognitive engagement that produces reliable scores

and demonstrates evidence for the validity of the inferences made from those scores. Currently, many of the instruments used to measure cognitive engagement are focused on a specific discipline and cannot be used across a variety of tasks. For example, the Science Activity Questionnaire (SAQ) is designed to assess engagement in the context of science activities. The Attitudes towards Mathematics Survey developed by Miller, Greene, Montalvo, Ravindran, and Nichols (1996) assesses academic engagement in mathematics courses. Items on these two scales may not be suitable for tasks outside of the science and/or mathematics classroom.

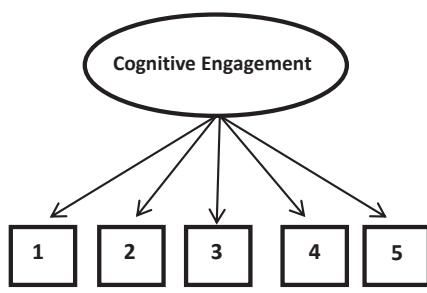
### Current Study

In the current study, faculty members wanted to examine cognitive engagement within the context of a large-scale arts, humanities, and literature assessment situation. The original items on the Attitudes Towards Mathematics Survey (Greene & Miller, 1996) were modified to address, specifically, student engagement on a low-stakes general education fine arts and humanities assessment instrument. Some of the original cognitive engagement items had to be excluded because they were irrelevant to the assessment context. Any time test users shorten a scale (American Educational Research Association [AERA], American Psychological Association [APA], & The National Council on Measurement in Education [NCME], 1999; Smith, McCarthy, & Anderson, 2000) or change the context of the questions (Baranik et al., 2010), the test users should re-examine the reliability of scores and the validity of inferences made from those scores. One such re-examination would be to test whether the factor structure of the original scale applies to the adapted measure.

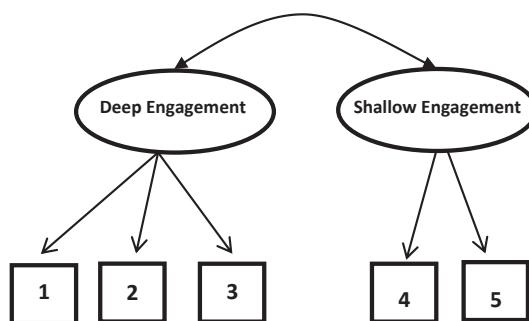
The dimensionality of the scale can affect scoring, which in turn impacts inferences from findings. In order to determine whether student scores should be interpreted as an overall cognitive engagement factor, or as two separate factors, (deep and shallow) as Greene and Miller (1996) suggested, the dimensionality of the adapted scale was examined using a confirmatory factor analysis (CFA). Researchers hypothesized that because the context of the new cognitive engagement scale was more specific (pertaining to one 45 minute testing session instead of an entire course) the items would be more closely related and represent a unidimensional model. A one and two factor CFA was run to test this hypothesis. For a priori hypothesis models, see Figure 1. Researchers examined global and local fit indices

Figure 1

1-factor model:



2-factor model:



to determine which model best represents the data. In addition, researchers established the internal consistency (Cronbach's alpha) of the instrument based on the factor structure as recommended by Cortina (1993). Finally, researchers examined the relatedness of this

scale to constructs that have shown to be correlated to cognitive engagement, specifically goal orientation and motivation. The development of a sound measure of cognitive engagement for students in large-scale assessment situations could assist faculty and assessment specialist in examining empirical questions such as, “Which assessment tests are most engaging for participants?”

### **Participants and Procedure**

Assessment specialists gathered responses to the short form of the cognitive engagement instrument (CE-S) from students participating in university-wide assessment day activities at a mid-sized, mid-Atlantic university. All incoming freshmen and students with 45–70 earned credits are required to participate in the university’s assessment activities. First-year, incoming students complete assessments in the fall on the last day of freshmen orientation. Students with 45–70 earned credits complete assessments in the spring. Students are assigned to testing rooms according to the last two digits of their university identification number. Using this method, the assessment specialists were able to randomly assign students to complete a specific battery of assessment instruments based on their room assignment. Participants included 243 students who completed the assessment activities during the fall of 2010 as incoming freshman or in the spring of 2011 after having earned 45–70 credits. The assessment specialists assigned the students in this study to an assessment battery that included the university’s fine arts and humanities assessment tests.

### **Instruments**

In addition to completing the university’s open-ended, constructed-response fine arts and humanities general education assessments, each participant completed a series of student development instruments. Among these instruments were scales designed to measure participants’ overall goal-orientation, as well as their motivation and cognitive engagement associated with the fine arts and humanities assessment.

**Cognitive Engagement – Short form.** The CE-S was adapted from a cognitive engagement scale written by Greene and Miller (1996). Five items were adapted from the scale and reworded to specifically refer to the specific large-scale assessment context. Participants are asked to respond to each question using a 1 to 5 scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree). Three of the questions are used to measure meaningful cognitive engagement while two questions were used to measure shallow cognitive engagement. Greene and Miller found a Cronbach’s alpha of .90 for their longer version of the meaningful engagement subscale and .81 for their longer version of the shallow engagement subscale. The current study examines the internal consistency of the shorter CE-S scale (For the CE-S items, see Appendix).

**Student Opinion Scale.** The Student Opinion Scale (SOS; Sundre, 1999) is a 10-item questionnaire used to measure examinee motivation. This scale is frequently used to help faculty understand motivation during low-stakes testing situations. Participants are asked to respond to each question using a 1 to 5 scale (1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree). The questionnaire contains two subscales measuring importance and effort. Thelk, Sundre, and Horst (2009) used Cronbach’s alpha as a measure of reliability and found the subscales to have an alpha value ranging between .80 to .89 for Importance and .83 to .87 for Effort. In the current study, internal reliability was found to be slightly lower for both the effort ( $\alpha=.74$ ) and importance ( $\alpha=.77$ ) subscales.

**Achievement Goal Questionnaire.** The Achievement Goal Questionnaire contains 12 goal orientation items (Finney, Pieper, & Barron, 2004), plus four work avoidance items (Pieper, 2004), and new mastery-avoidance items from Elliot and Murayama (2008). The AGQ consists of five subscales: mastery-approach, performance-avoidance, work avoidance, performance-approach, and mastery-avoidance that coincide

**“If students are not engaged while taking the test, institutions will have assessment results, but what inferences can we draw from these results?”**

with achievement goal theory mentioned previously. Cronbach's alpha for the subscales range from .65 to .89.

## Results

### Data Cleaning and Screening

Before running the models, the data were checked for outliers and normality. Data were screened for univariate and multivariate outliers. A graphical plot of the cognitive engagement scores was used to screen for univariate outliers. Researchers used a SPSS macro written by DeCarlo (1997) to screen for multivariate outliers. Analyses suggest that there are no outliers. Univariate normality was assessed by examining skewness and kurtosis. All of the skewness and kurtosis values fell below the recommended cutoffs of |2| for skewness and |7| for kurtosis (Bandalo & Finney, 2010; see Table 1). A histogram with an overlying normal curve was used to examine normality for each item. The responses appeared to depart from the normal curve, a possible function of the

Table 1.

Item Correlations and Descriptive Statistics (N = 243)

Item	1	2	3	4	5
1	1.00				
2	.46	1.00			
3	.31	.39	1.00		
4	-.28	-.13	-.32	1.00	
5	-.39	-.31	-.26	.62	1.00
M	2.81	3.67	3.45	2.81	2.42
SD	1.04	0.88	0.98	1.04	0.84
Skew	0.14	-0.88	-0.66	0.45	1.03
Kurt	-0.91	0.62	-0.22	-0.66	1.26

categorical nature of the data. Evidence of multivariate non-normality was also found using Mardia's normalized multivariate kurtosis; therefore, the researchers decided to use robust diagonally weighted least squares estimation methods

### Factor Analysis

The asymptotic covariance matrix used for the analyses was produced in PRELIS 2.71, and the confirmatory factor analyses were conducted using LISREL 8.72 (Jöreskog & Sörbom, 2005). A unidimensional model was fit to the data to obtain evidence that the CE-S items are measuring cognitive engagement as a single construct. A two-factor model was fit to the data to see if the items are measuring cognitive engagement as two separate factors as previously found by Miller et al. (1996). Hu and Bentler (1998, 1999) recommend reporting at least one absolute fit index and one incremental fit index in addition to  $\chi^2$ . Therefore, four global fit indices were examined to evaluate model fit: the  $\chi^2$ , the standardized root mean square residual (SRMR), the robust root mean square residual (RMSEA), and the robust comparative fit index (CFI). The  $\chi^2$  test is an absolute fit index that is sensitive to sample size. Like the  $\chi^2$ , the SRMR and RMSEA are absolute fit indices, meaning that they assess how well the hypothesized model reproduces the sample asymptotic covariance matrix. It is recommended the SRMR and RMSEA values be .08 or less (Browne & Cudeck, 1993; Hu & Bentler, 1999). The CFI is an incremental fit index and, unlike the other indices, larger values indicate adequate model fit. Hu and Bentler (1998, 1999) recommend a cutoff of .95 or above.

**“...simply attending  
an environment  
and completing  
necessary work are  
not good indicators  
of cognitive  
engagement.”**

Table 2 shows the fit indices for the one and two factor models. None of the fit indices for the one factor model are within the suggested cutoffs. However, all of the indices for the 2-factor model are within the recommendations set forth by previous research. Localized misfit in the 2-factor model was investigated by looking at the

Table 2.

Fit indices for the one and two factor models (N = 243)

Model	$\chi^2$	df	SRMR	RMSEA	CFI
Two Factor	10.29	4	.05	.08	.97
One Factor	43.63	5	.10	.18	.82

\* Note: RDWLS estimation used

Table 3.

Standardized Polychoric Residuals for the One and Two Factor model (N = 243)

Two Factor Model:

	Item 1	Item 2	Item 3	Item 4	Item 5
Item 1	--				
Item 2	.95	--			
Item 3	-1.70	.80	--		
Item 4	-.17	1.78	-1.70	--	
Item 5	-.73	.33	.41	--	--

One factor Model:

	Item 1	Item 2	Item 3	Item 4	Item 5
Item 1	--				
Item 2	3.54	--			
Item 3	-.02	1.85	--		
Item 4	1.56	3.07	-.08	--	
Item 5	1.18	1.54	1.71	7.71	--

Table 4.

Standardized Factor Pattern Coefficients, Correlations, and Cronbach's Alpha for the Two-Factor Model (N = 243)

Items	Deep	Shallow	Error Variance	$R^2$
1	.69		.53	.47
2	.64		.59	.41
3	.56		.69	.31
4		.68	.54	.46
5		.91	.17	.83
Deep	1.00			
Shallow	-.57	1.00		
Cronbach's $\alpha$	.57	.71		

standardized polychoric residuals. The 1-factor model has several areas of local misfit that exceed the recommended cutoff of |3|; while the 2-factor model had no areas of misfit (see Table 3). Since the 2-factor model had appropriate values for both the fit indices as well as the standardized polychoric residuals, we championed this model. Reliability for the two subscales was also examined. While the deep subscale reliability ( $\alpha=.56$ ) is not acceptable for program-level inferences, it is higher than expected considering the number of items in the subscale. The two-item shallow subscale has an impressive reliability of .71, indicating it may be appropriate for program-level inferences (Nunnally, 1978). No  $\Delta X^2$  was reported as the fit indices for the one factor model clearly did not represent the data.

Having championed the 2-factor model, we looked at the parameter estimates (See Table 4) to understand how much of the variance in the item is accounted for by the latent factor (or how much variance was due to measurement error). The standardized coefficients ranged from .56 to .91 and were all significant at  $p < .05$ . Squaring these standardized estimates produced the  $R^2$  for each item.  $R^2$  values ranged from .31 to .83. These values indicate that items such as item 3 had low variance accounted for (31%) by the latent factor (deep cognitive engagement) and large amounts of unexplained variance. Item 5 on the other hand had a large amount of variability explained by the latent factor (83%). The standardized error variances ranged between .17 and .69 for all items. Finally, the factor intercorrelations were estimated (Table 4). The deep and shallow factors had a moderate negative correlation (-.57) suggesting that as deep engagement increases, shallow engagement decreases.

### Relationships with External Variables

Table 5 shows the correlations between the two subscales of the CE-S with the SOS total score and each subscale of the SOS and AGQ. The deep subscale is positively related to the SOS total score as well as each SOS subscale, suggesting that as deep engagement goes up so does both effort and importance. However, these correlations are only moderate in nature, suggesting that these two constructs are related but different. The deep subscale also has low to moderate correlations with the AGQ subscales. As expected based on previous literature, AGQ mastery performance subscale scores are related to the deep subscale of the CE-S. There was no significant correlation with the deep subscale of the CE-S to

**“In order to make valid inferences regarding students’ level of cognitive engagement across different tasks, researchers must have a measure of cognitive engagement that produces reliable scores and demonstrates evidence for the validity of the inferences made from those scores.”**

Table 5.

Correlations with External Variables (N = 243)

	Deep	Shallow
SOS total	.54**	.25**
Effort	.45**	.13*
Importance	.43**	.29**
Mastery Approach	.29**	.14*
Performance Approach	.13*	.11
Mastery Avoidance	.22**	.17**
Performance Avoidance	.05	.13*
Work Avoidance	-.25**	-.13*

\*\* Correlation is significant at the .01 level.

\* Correlation is significant at the .05 level.

the performance avoidance subscale of the AGQ and only a slight negative correlation with the work avoidance subscale of the AGQ. While several of the correlations are statistically significant, correlations between the shallow subscale scores and the AGQ subscales were all small (less than  $r = .17$ ).

## Discussion

After examining both a 1-factor and a 2-factor solution, we have championed a 2-factor model of cognitive engagement as measured by the CE-S. This is consistent with Miller et al. (1996) and Meece et al. (1988). Therefore, in this case, shortening a parent questionnaire and changing the context to be more specific did not affect the factor structure of the instrument. Reliability of the subscale scores was higher than expected considering the small number of items composing the two subscales.

We also looked at external correlations which seem to support that cognitive engagement is related to other constructs in expected ways. The positive and moderate correlation between deep cognitive engagement and motivation shows that the two constructs are related, yet distinct from one another (Appleton et al., 2006; Nemann et al., 1992). Deep engagement is also positively related to mastery approach and not related to performance avoidance, which is consistent with Meece et al. (1988). Shallow engagement showed much smaller correlations with these variables, further supporting the two-factor model by showing that the deep and shallow are related to other variables in different, yet predicted ways.

## Future Research

In the future, more work should be done to continue to develop the CE-S as a psychometrically sound instrument for cognitive engagement. As mentioned earlier, this work is important to both assessment and educational practices. The development of additional items designed to tap into the deep and shallow engagement factors may improve subscale score reliability. However, we do still want to make sure that we retain only a small amount of items to make sure that use of the cognitive engagement instrument is feasible and easy to add into existing assessment processes.

In addition to adding items, this study should be replicated with a new sample of participants to examine the stability of the 2-factor model. Another future direction could be developing a cognitive engagement scale to examine cognitive engagement on selected-response assessments, as the CE-S was developed for use with construct-response assessments only. The ultimate goal of this instrument development process should be to develop a general cognitive engagement instrument that can be used flexibly across all assessments.

Once a sound instrument of cognitive engagement is fully developed, future research can examine empirical questions related to assessment practice. One example of an interesting question that could be relevant to an assessment specialist is, “which assessment produces higher cognitive engagement in different contexts, (open-ended vs. multiple choice, paper and pencil vs. computer based testing, etc.)”? Once a good measure is established, assessment specialists may also want to model the relationship among cognitive engagement, effort, and performance. Understanding the connectedness of these constructs may assist in the development of interventions designed to increase students’ cognitive engagement on low-stakes assessments. Also of interest may be whether students are giving quality responses on constructed response tests, making sure rapid responding is diminished on multiple-choice assessments, and investigating whether participants are skipping fewer questions when compared to less cognitively engaging assessments.

## Conclusion

Cognitive engagement currently is under-researched in applied assessment contexts. The study of this construct may provide unique information regarding

**“Students who exhibit behaviors that allow them to master academic work are seen to have deep cognitive engagement, while students who exhibit behaviors such as rote memorization and rituals they perceive will help them to do well without developing mastery of the material are demonstrating shallow engagement.”**

students' effort and performance on assessment tests beyond that currently understood through motivation theory alone. Considering the factor structure and reliabilities of the CE-S scale, this scale appears to have potential as a psychometrically sound measure of deep and shallow cognitive engagement. The addition of a few quality items would likely increase the utility of the measure. The establishment of such a method would allow assessment practitioners to test empirically multiple hypotheses regarding the role of cognitive engagement in assessment practice.

## References

- American Educational Research Association (AERA), American Psychological Association (APA), & The National Council on Measurement in Education (NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools*, 45, 369–386.
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology*, 44, 427–445.
- Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 125–155). Florence, KY: Routledge Education.
- Banta, T. W. (2008). Editor's notes: Trying to clothe the emperor. *Assessment Update*, 20(2), p. 3–4, 15–16.
- Baranik, L. E., Barron, K. E., & Finney, S. J. (2010). Examining specific versus general measures of achievement goals. *Human Performance*, 23, 155–172.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98–104.
- DeCarlo, L. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2, 292–307.
- Dweck, C. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040–1048.
- Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, 100, 613–628.
- Finney, S. J., Pieper, S., L., & Barron, K. E. (2004). Examining the psychometric properties of the Achievement Goal Questionnaire in a general academic context. *Educational and Psychological Measurement*, 64, 365–382.
- Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, 21, 181–192.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424–453.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Jöreskog, K. G., & Sörbom, D. (2005). LISREL 8.72 [Computer software]. Lincolnwood, IL: Scientific Software.

- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, 37, 153-184.
- Meece, J. L., Blumenfeld, P. C., & Hoyle, R. H. (1988). Students' goal orientations and cognitive engagement in classroom activities. *Journal of Educational Psychology*, 80, 514-523.
- Miller, R. B., Greene, B. A., Montalvo, G. P., Ravindran, B., & Nichols, J. D. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others, and perceived ability. *Contemporary Educational Psychology*, 21(4), 388-422.
- Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11-39). New York, NY: Teachers College Press.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91, 328-346.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.
- Pieper, S. L. (2004). Refining and extending the  $2 \times 2$  achievement goal framework: Another look at work-avoidance. Doctoral dissertation, James Madison University, 2004. *Dissertation Abstracts International*, 64, 4436.
- Pike, G. R. (2006). Assessment measures: Value-added models and the collegiate learning assessment. *Assessment Update*, 18(4), 5-7.
- Shermis, M.D. (2008). The collegiate learning assessment: A critical perspective. *Assessment Update*, 20(2), 10-12.
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12, 102-111.
- Sundre, D. L. (1999). Does examinee motivation moderate the relationship between test consequences and test performance? (Report No. TM029964). Harrisonburg, Virginia: James Madison University. (ERIC Documentation Reproduction Service No. ED432588).
- Thelk, A. D., Sundre, D. L., Horst, S. J., & Finney, S. J. (2009). Motivation matters: Using the Student Opinion Scale to make valid inferences about student performance. *The Journal of General Education*, 58, 129-151.
- Terenzini, P. T. (2010). Assessment with open eyes: Pitfalls in studying student outcomes. In J. Fredericks Volkwein (Ed.), *New Directions in Institutional Research Special Issue: Assessing Student Outcomes*, (S1), 29-46.
- Wigfield, A., & Eccles, J. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68-81.
- Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10(1), 1-17.

## Appendix

Earlier in today's assessment session you completed two assessment tests designed to assess your performance on learning goals associated with JMU's General Education Cluster 2 (Fine Arts and Humanities). These assessments were the humanities test and the aesthetics test. The humanities test asked you to respond to two separate texts while the aesthetics test asked you to respond to a painting, musical work and play. Please consider these two particular assessments when responding to the following items.

- 1) When approaching the questions on the Cluster 2 assessments, I planned out or organized my response prior to writing my answer.
- 2) When preparing to answer the questions on the Cluster 2 assessments, I stopped to reflect on my experience with the works (text, video, music, painting) presented.
- 3) When experiencing the works (text, video, music, painting) presented in the Cluster 2 assessments, I considered issues related to culture when considering their meaning or significance.
- 4) When answering the questions on the Cluster 2 assessments, I considered how those reviewing the answers would want me to respond.
- 5) When answering the questions on the Cluster 2 assessments, I looked for clues of how to respond with the test itself.

## **Abstract**

The researchers explored whether implementation of a systematic outcomes-based assessment process is necessary for demonstrating quality in service learning programs at a two-year and a four-year institution. The findings revealed that Western Community College and the University of the Coast maintained quality service-learning programs, which met their established learning goals and reflected the community's needs, by incorporating numerous components of a systematic outcomes-based assessment process, even though this process had not been established.

### **Authors**

Lauren Weiner, Ed.D.  
Cuyamaca College

Marilee J. Bresciani, Ph.D.  
San Diego State University

**Email**  
lauren.weiner@gcccd.edu

# **CAN INSTITUTIONS HAVE QUALITY PROGRAMMING WITHOUT UTILIZING A SYSTEMATIC OUTCOMES-BASED ASSESSMENT PROCESS?**

For many students, service learning not only expands their educational horizons, but also makes them more aware and compassionate of those who live in communities very different from their own. Yet, despite the expressed benefits of service learning and the increasing numbers of institutions offering service-learning programs, it is not known whether they are truly successful unless evidence of their success is provided. The use of outcomes-based assessment is one of the processes that generates evidence of program effectiveness. While scholars have written about the ways in which effective outcomes-based assessment must be implemented (Bresciani, 2006; Bresciani, Moore-Gardner, & Hickmott, 2009; Maki, 2004; Palomba & Banta, 1999; Suskie, 2004; Upcraft & Schuh, 1996), few studies have been performed to identify whether all the components of effective outcomes-based assessment must be present in order for quality programs to be identified. Thus, the purpose of this cross-case comparative study was to find out whether implementation of a systematic outcomes-based assessment system is necessary for demonstrating quality in service learning at a community college and a private four-year institution.

### **Review of Literature**

Bresciani (2006) noted that outcomes-based assessment provides vital information to enhance learning through "systematic evaluation of student learning and development" (p. 1). In addition, according to Bresciani et al. (2009), "Outcomes-based assessment is intended to inform decision-making processes" (p. 14) and "goes one step beyond typical evaluation" (p. 16). Bresciani noted that a strategy is devised to accomplish the intended goals after determining what the curriculum is to achieve "in regard to its services, research, student learning, and faculty/staff development programs" (p. 14). A systematic method of evaluation, outcomes-based assessment is a continuous process, which recurs on a timeline in order to ascertain "whether the program improvements contribute to the intended outcomes" (Bresciani, 2006, p. 14).

As noted in the literature, there are several integral components that make up a quality outcomes-based assessment process (Bresciani, 2006; Bresciani, Zelna, & Anderson, 2004; Palomba & Banta, 1999; Suskie, 2004). These components should include: (a) defining mission statements, (b) defining goals, (c) distinguishing between

institutional, division, and programmatic learning outcomes, (d) mapping learning outcomes, (e) planning the delivery of your outcomes, (f) choosing a method and criteria to assess, (g) reporting and interpreting results, (h) using results to make decisions and recommendations, and (i) establishing a timeline for implementation and improvement (Bresciani et al., 2009).

The researchers were interested in learning from the students, faculty, staff, and community partners of Western Community College and the University of the Coast whether these two institutions incorporated a systematic outcomes-based assessment process when evaluating the effectiveness of their service-learning courses. However, if the institutions utilized outcomes-based assessment, how did they handle this process?

From this study, the researchers were able to determine that Western Community College and the University of the Coast evaluated the effectiveness of their service-learning programs by (a) the success of their collaborative partnerships, (b) the successful connection of their students' curricular and co-curricular experiences, (c) the availability of reflection opportunities, (d) the depth of the students' feedback and assessment opportunities, and (e) the sustainability of the program. Although there were many similarities between Western Community College and the University of the Coast, only the University of the Coast incorporated a one-unit for credit leadership course, which was designed to train selected service-learning student leaders to work as site coordinators. Site coordinators serve as intermediaries between the service-learning students, faculty, community partners and the Community Service-Learning Center to make sure that the academic aspects of the course are being met along with the needs of the partners. The following study illustrates the process used to discover how and which of these portions of the systematic outcomes-based process are present.

**“Outcomes-based assessment is intended to inform decision-making processes and goes one step beyond typical evaluation.”**

## Methodology

### Research Design

This cross-case comparative study explored through a constructive lens whether implementation of a systematic outcomes-based assessment process is necessary to determine whether quality of service-learning programs are present in a four-year private institution and a public community college, referred to as the University of the Coast and Western Community College. The researchers selected a multiple case study as the specific methodology type since it works very well for gathering information to explore the factors that may lead to sustainable, high-quality service learning programs and since this model is very appropriate “whenever it is human inquiry that is being considered” (Guba & Lincoln, 1989, p. 82). The researchers also chose the constructive method since its flexibility enabled them to pursue avenues of investigation that may arise during the investigation process, making this paradigm open to finding out new information and substantiating it (Guba & Lincoln, 1989). The researchers defined quality based upon an extensive literature review of the characteristics of quality service-learning programs and by the research process that uncovered components of a systematic outcomes-based assessment.

The researchers utilized multiple sources of data collection, including one-on-one interviews and focus groups with students, faculty, staff members, and community partners and on-site observations as well as document analysis of (a) students' journals from class, (b) faculty or staff reports, (c) student assessment, (d) faculty and staff assessments, and (e) partnership assessments. The researchers also reviewed the assessment plans of students, faculty and or staff members, and community partners.

### Site Selection and Sampling Method

The researchers selected Western Community College's Service-Learning Program and the University of the Coast's Community Service-Learning Center for the study since these particular institutions have long established service-learning programs with strong

ties to their respective communities. In addition, these two institutions are very different from each other. For example, Western Community College is large, two-year public educational institution relying on state funding and the University of the Coast is a private, four-year church-based educational institution with an endowment of hundreds of millions of dollars.

There were 26 participants in the study, five of whom were males. In addition to one-on-one interviews with 18 persons including (a) five students, (b) six faculty members, (c) three staff members, and (d) four community partners, there were two focus groups with a total of eight students. The researchers utilized purposeful sampling and asked the service-learning administrators at both institutions to identify faculty or staff members for one-on-one interviews, who had different areas of service-learning expertise and varying degrees of experience in service learning. Based on the recommendations of the service-learning coordinators, the researchers interviewed community partners from (a) both small and large non-profit organizations, (b) long established and newly formed non-profit organizations, (c) those working with different age groups, different socio-economic groups, and different ethnic groups.

In framing the interview protocols for one-on-one interviews and focus groups, the researchers selected a set of specific questions derived from a review of literature, which added focus and meaning to the research question. These particular questions provided insight into what factors the participants believed were essential components of quality service-learning programs. Another integral component of this cross-case comparative study were observations at partner sites and a document analysis of official documents, such as strategic plans, assessment plans, student evaluations, reflections, and mission statements. These public documents not only enriched the study and helped the researchers determine if these two institutions of higher education had the characteristics of quality service-learning programs, but also supplemented the other methods of the study, providing more depth to the researchers' findings.

### **Data Analysis**

Using a "process of inductive reasoning, thinking, and theorizing" (Taylor & Bogdan, 1998, p. 140), the researchers analyzed recorded transcripts of the one-on-one interviews and focus groups. From the transcripts, the researchers captured from students, faculty, staff members, and partners their feelings about their service-learning programs including (a) their reactions to their service-learning experience, (b) their definition of sustainability, (c) their interactions with all of the entities that comprise the program, and (d) any issues of concern.

**"In our opinion, Western Community College and University of the Coast are successful since both institutions use a variety of ways to determine learning outcomes for their service-learning courses."**

### **Findings and Discussion**

While determining whether each institution has a systematic outcomes-based assessment process in place, and if not, identifying what components of a systematic outcomes-based assessment process exist, the researchers learned there is close communication among all of the entities of both service-learning programs. Although the University of the Coast's Community Service-Learning Center and Western Community College's Service-Learning Program do not have a systematic outcomes-based assessment process, both programs have numerous components of the process in place, such as (a) incorporating feedback, including surveys and evaluations from students, faculty, and community partners as part of an internal assessment process to make necessary changes that have improved their respective service-learning programs; (b) having the service-learning administrators interface with faculty on an individual basis to review the service-learning component for their courses and devise learning outcomes; (c) providing numerous orientation and training opportunities, including workshops to assist faculty in developing learning outcomes; and (d) encouraging close communication, enabling the entities in the partnership to interface and share their ideas and concerns. Both institutions communicated "a shared purpose and vision...defining

goals" (Bresciani et al., 2009, pp. 36-37), which should be present in an outcomes-based assessment.

A formalized feedback and assessment process is an important factor leading to the success of both service-learning programs. Both the service-learning coordinator at Western Community College and the director and assistant director of the University of the Coast's Community Service-Learning Center place great emphasis on both feedback and assessment by maintaining close communication with all entities providing (a) guidance about the purpose, goals, and logistics of service-learning programs, (b) information about the community partners and how their mission may complement the academic aspects of service-learning courses, and (c) information about the service-learning sites to assist faculty in designing and implementing programs to achieve the expected learning outcomes. This information not only educates students, faculty, and partners, but also enables them to solicit feedback, strengthening the assessment process.

Although neither the University of the Coast nor Western Community College utilize a systematic outcomes-based process, faculty members meet with service-learning program administrators to discuss learning outcomes. According to Bresciani (2006), in an outcomes-based assessment, it is important to "engage in these conversations faculty and staff who are well respected in their disciplines and in their research" to lessen "...the common misconceptions that outcomes-based assessment is just another fad or a process built by administrators to 'check up' on the work of the faculty" (p. 19). Bresciani also noted that the information compiled through this process promotes discussions about responsibility and may lead to "opportunities for improvement" (p. 15).

From the interviews and document analysis, the researchers determined that Western Community College established several outcomes in its Unit Assessment Report (2008-2009, pp. 1-4) that included (a) increasing the service-learning skills of its student advocates by providing leadership training to plan a community service project, (b) advancing critical and reflective thinking skills to encourage connections between curricular and co-curricular experiences, and (c) increasing student knowledge about the organizations which are their community service-learning partners. However, the Unit Assessment Report tracks only a few objectives and does not reflect the work of the entire program. According to the University of the Coast Strategic Plan, "Our key success factors are our outcome measures of success. They measure how successfully the University of the Coast Community Service-Learning Center is achieving our Mission, Vision 2012, and our Core Values on a year-by-year basis" (2008-2012, p. 8).

In our opinion, Western Community College and the University of the Coast are successful since both institutions use a variety of ways to determine learning outcomes for their service-learning courses. Neither institution utilizes a systematic outcomes based assessment that provides integration between course outcomes and program outcomes, the purpose of which is to provide "quality assurance and external accountability in higher education" (Bresciani, 2006, p. 13). From the comments of the participants in the study, the researchers learned that the instructors at Western Community College and the University of the Coast make their students aware of the expected learning outcomes from their courses. For instance, in an ongoing systematic manner, both institutions have (a) devised the learning outcomes for their courses, (b) gathered feedback, (c) assessed it, (d) made necessary changes, and (e) evaluated the results of the changes to determine if students are achieving the expected learning outcomes purposefully established, a process that has characteristics of an outcomes-based assessment (Bresciani, 2006).

The document analysis not only provided insight into the assessment procedures at the University of the Coast and Western Community College, but also reinforced the findings from the interviews and focus groups. For example, Western Community College's Unit Assessment Report 2008 went into great detail regarding Western Community Colleges outcomes such as increasing the service-learning skills of its student advocates by

providing leadership training to plan a community service project and increasing student knowledge about the organizations which are community service-learning partners. The University of the Coast's service-learning rubric was very specific and provides guidelines and structure. It can be an excellent tool to assess academic achievement of the curricular and co-curricular experiences and may provide as much or more structure for systematic assessment than a formalized plan would provide. In addition, in the University of the Coast Community Service-Learning Center Strategic plan (2008-2012), the Community Service-Learning Center placed a great deal of emphasis on the assessment and evaluation of its service-learning programs and incorporated these categories into the document.

Although Western Community College is not utilizing an ongoing systematic outcomes-based assessment at this time, the coordinator is working with the faculty and administrators to establish overall learning outcomes of the service-learning program and not just the learning outcomes for particular courses. She is putting into place "a systematic process of evaluation . . . repeated at a later date to determine whether the program improvements contribute to the intended learning outcomes" to incorporate components of an ongoing systematic outcomes-based assessment (Bresciani, 2006, p. 14). According to McEwen (1996), "Carefully designed service-learning experiences can lead to profound learning and development outcomes for students, the primary reason that institutions of higher education engage in service-learning" (p. 53). Matt, assistant director of the service-learning programs at the University of the Coast, also discussed the importance of establishing learning objectives in designing programs and "getting a buy-in" from all of the stakeholders. "Getting their [stakeholders'] input and seeing how they think the program should be run . . . helps guide you," Matt continued. Even though the University of the Coast does not have a systematic on-going outcomes-based assessment process, they are devising learning outcomes not only for individual courses, but also for their overall service-learning program.

The close interaction and ongoing quality checks between both the service-learning administrators at University of the Coast and at Western Community College and their respective stakeholders is instrumental in achieving the intended outcomes of their service-learning programs. Constant communication between all of the entities is essential to preserve quality, especially when there are changes in the staffing at a community partner's service-learning site. Based on the study, an important component of Western Community College and the University of the Coast's success is the fact that both institutions have strong and stable service-learning departments where their staff members maintain close ties with all of the stakeholders.

In addition, both institutions' service-learning departments work directly with faculty, helping them to assess learning outcomes for the course on an individual basis, but the learning outcomes are not integrated with program outcomes. However, the importance of utilizing a formalized systematic on-going outcomes-based assessment process cannot be overlooked. For example, if there were changes in the hierarchy of the service-learning programs at either institution and a new coordinator or director replaced Michelle or Wendy, gaps in leadership may develop, and their successful service-learning programs may suffer.

The University of the Coast and Western Community College's service-learning programs succeeded without a systematic ongoing outcomes-based assessment review due to the fact that they implemented the characteristics of a quality service-learning program, including informally and continuously soliciting feedback and making evaluations. However, other two-year and four-year institutions may not experience the same results without a systematic outcomes-based assessment process if they do not incorporate the same quality characteristics as the University of the Coast and Western Community College.

Even though an outcomes-based assessment process does not exist at the

University of the Coast's Community Service-Learning Center and Western Community College's Service-Learning Program, based on this study, both institutions appear to have quality service-learning programs as noted by (a) their sustainability; (b) the satisfaction of their students, faculty, and community partners; (c) the agreement on the intended learning goals and how the program meets the needs of the community; (d) and the learning that takes place, which is reflected in the feedback and reflection activities. From the study, the researchers determined that both institutions (a) have clearly defined their mission statements through an excellent communication network, (b) have carefully identified the goals of their individual service-learning programs even though this identification is not always a collaborative effort between the faculty and service-learning staff, but at a later time is mutually discussed and understood, (c) selected appropriate assessment methods through a variety of different processes appropriate for the particular program, (d) report and interpret results to all of the partners, and (e) use the results to make decisions and recommendations for program improvement.

While faculty members and service-learning administrators at Western Community College and the University of the Coast devise learning outcomes and are working towards distinguishing between and integrating the different types of learning outcomes, the process is still under development at both institutions. The institutions' missing pieces are (a) not incorporating a segmented process for the developing the achievement of learning outcomes, (b) the lack of integration between service-learning administrators and faculty in identifying whether the learning outcomes have been achieved, (c) the lack of a systematic process for conducting assessments, (d) not systematically identifying, the results of the learning outcomes, (e) not consistently utilizing the results of the assessment for program improvement or recommendations, and (f) not consistently repeating the evaluation process to measure whether the improvements or recommendations have resulted in the achievement of the learning outcomes.

### **Conclusion and Implications for Practice**

The findings of this study reveal that while a systematic outcomes-based assessment process did not exist at the University of the Coast and Western Community College, there was evidence of quality in both programs that were studied since both institutions demonstrated several effective assessment practices. For example, they placed great emphasis on feedback and assessment between their various constituent groups. Western Community College has also increased the frequency of its program reviews and has incorporated learning outcomes into the process, and the University of the Coast has created a Service-Learning Student Learning Rubric. In addition to surveys and the rubric, some professors teaching service-learning courses have students write literacy case studies. For example, according to third-year student, Alice, literacy case studies not only help to evaluate the progress service-learning students made tutoring youngsters, but also enable them to make suggestions for improvement. The University of the Coast uses both formal assessments with a paper trail and informal assessments, and according to Wendy, director of the University's Community Service-Learning Center, "With the partners, we moved from formal to more checking in at the beginning and end of each semester with personal conversations." Feedback has always been an essential part of the University's assessment process and, according to the University of the Coast Community Service-Learning Center Annual Report (2005-2006), great emphasis is placed on student feedback to determine what students are learning from their community experiences. Feedback from this objective came from an evaluation by 256 students replying to 15 statements answered on a Likert scale; "97% of the students responded favorably that the service experience helped them better understand their organization" (Western Community College Unit Assessment Report, 2008-2009, p. 4).

The researchers recommend that both institutions continue to expand on these assessment practices, such as the feedback and reflection activities that they have in place. However, the level of quality could only be identified because the researchers conducted a

**"...unless an institution has a commitment from its leadership to continue a program regardless of the quality of data they are able to produce in a systematic way, it may be wise to engage in a systematic outcomes-based assessment process or at the very least, engage researchers to conduct a thorough case study so that quality data and evidence of decision-making can emerge and be documented."**

thorough cross-case comparative study in order to find the data that revealed the quality of the programs. Thus, unless an institution has a commitment from its leadership to continue a program regardless of the quality of data they are able to produce in a systematic way, it may be wise to engage in a systematic outcomes-based assessment process or at the very least, engage researchers to conduct a thorough case study so that quality data and evidence of decision-making can emerge and be documented.

While it may not be necessary for both institutions to incorporate an ongoing systematic outcomes-based assessment process at this time because of their institutional leadership commitment to continue service-learning, the researchers believe it would be helpful for them to do so given the scope of their service-learning programs and the continued expansion of their outreach student and community outreach efforts. After a systematic outcomes-based assessment process is developed and in place, it may take less time to generate accurate data to determine whether there is a quality program, thus, saving time later in replicating a similar study to the one conducted at Western Community College and the University of the Coast.

### References

- Bresciani, M. J. (2006). *Outcomes-based academic and co-curricular program review*. Sterling, VA: Stylus Publishing.
- Bresciani, M. J., Moore-Gardner, M. M., & Hickmott, J. (2009). *Demonstrating student success in student affairs*. Sterling, VA: Stylus Publishing.
- Bresciani, M. J., Zelna, C. L., & Anderson, J. A. (2004). *Assessing student learning and development: A handbook for practitioners*. Washington, DC: National Association of Student Personnel Administrators.
- Guba, E., & Lincoln, Y. (1989). *Fourth generation evaluation*. Newbury Park, CA: Sage Publications.
- Maki, P. L. (2004). *Assessing for learning: Building a sustainable commitment across the institution*. Sterling, VA: Stylus Publishing.
- McEwen, M. K. (1996). *Enhancing student learning and development through service-learning*. In B. Jacoby & Associates (Eds.), *Service-learning in higher education* (pp. 53-91). San Francisco, CA: Jossey-Bass.
- Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco, CA: Jossey-Bass.
- Suskie, L. (2004). *Assessing student learning: A common sense guide*. Bolton, MA: Anker Publishing.
- Taylor, S. J., & Bogdan, R. (1998). *Introduction to qualitative research methods*. New York, NY: John Wiley.
- Upcraft, M. L., & Schuh, J. H. (1996). *Assessment in student affairs*. San Francisco, CA: Jossey-Bass.

# Reflective (Ac) Counting: Institutional Research, Evaluation, & Assessment in a Time of Cholera

Walter F. Heinecke, Ph.D.

Associate Professor of Research, Statistics, & Evaluation

*University of Virginia*

heinecke@virginia.edu

Institutional researchers and assessment professionals in higher education are living in interesting and challenging times, one might say in times of crisis. In the post-Reagan era government and its agencies and public professionals have come under fire for being ineffective and reluctant to reform and protected as special interests (House, 1993). As federal and state coffers have shrunk in a time of rising costs of higher education, university budgets are tightening while at the same time calls for accountability are increasing (U.S. Department of Education, 2006). The discourse of higher education quality has changed to focus on “return on investment” as the criteria for college and university success. Free market economics take the spotlight off of private enterprises but shift attention of policy makers to public organizations and as House (1993) asserts, “As a consequence, higher education...is blamed for social and economic problems that originate elsewhere, such as in the economic structure itself” (p. 59). As a consequence, Terenzini (2009) asserts that assessment of student outcomes originating from external mandates is here to stay but that there are significant “conceptual, administrative, political and methodological issues” (p. 45) making such initiatives problematic. As a result, institutional researchers are caught in a conflicting press for efficiency, accountability and excellence. The research on the role of IR professionals reflects more emphasis on “accountability and performance” and “outcomes assessment” among other job tasks (Volkwein, 2008, p. 9). Here I attempt to offer perspective related to the “issues intelligence” as well as the “contextual intelligence” that make up the “organizational intelligence” of IR (Volkwein, 2008, p. 5; see also Terenzini, 1999).

As in other social arenas the use of the term crisis always requires close inspection. Crises are socially constructed and they offer an opportunity to examine what an institution is doing and how it is conducting its operations. When a crisis is declared, it stands to reason that we should ask who is defining the crisis. Is the crisis defined in such a way as to determine new solutions and new patterns of operating? Are the calls for greater accountability, evaluation, and assessment motivated by forces to improve or discipline higher education? In

effect, times of crisis are good times to take a step back from the fray of day-to-day demands of standard operating procedures and examine what we are doing and why. This applies to evaluation and assessment professionals in institutional research and assessment offices within colleges and universities. We have to ask ourselves: Are our evaluation activities merely tools for “informing and legitimizing the unpopular steps that government must take, which often means budget cutting”? Are our evaluation activities merely lending scientific authority to questionable political decisions? “When professionals work in bureaucracies, their autonomy is often challenged. So professional versus bureaucratic interests is a central conflict. The conflict is increasingly manifested in the

**“Are our evaluation activities merely  
lending scientific authority to  
questionable political decisions?”**

higher education system, a stronghold of professional knowledge and legitimization. Governments have curtailed funding and increased their control over universities” (House, 1993, p. 53). Do productivity and efficiency trump the needs of our clients? To what degree should institutional researchers follow government policy and to what degree must they follow the standards of their own profession? Ultimately these are matters of professional ethics and social justice” (House, 1993, p. 55).

The field of evaluation itself exists in a political and social context and is at a crossroads of sorts. Writing in the 1990s, House’s (1993) words seem prophetic:

Exactly in what form evaluation develops depends on how modern market societies develop. If these societies become more authoritarian, a distinct possibility in reaction to managing turbulent societies and sluggish economies, evaluation could be used for repressive purposes. On the other hand, if modern market societies become less ideological and more willing to consider new social possibilities, then evaluation could become more useful. ...Being involved with

government programs means that evaluation is always connected to ideological and political issues (pp. 28-29).

The attacks on colleges and universities, “inspired by the ideology of the free market” and “... directed at the content as well as the costs, of higher education” (House, 1993, p. 59) are not solely from external bodies such as business interests, commissions, accrediting organizations, think tanks, or the media, they come from within the academy as well. When colleges and universities are criticized by researchers, working with foundation support, putatively because undergraduates don’t learn much as a result of attending our universities (Arum & Roksa, 2011), the stakes go up and the press for assessment goes up as well. “Measures imply a need for action” (Stone, 2002, p. 167). We get forced into an efficiency and accountability frame with other frames such as equity and quality pushed from the discussion. The tendency is to become myopic in our vision and practice of assessment.

In 1993, House wrote:

...(T)raditional autonomy that higher education institutions have enjoyed is coming to an end. The question is not whether we should have accountability, but rather what kind of accountability and evaluation we can have that will protect the vital internal processes of research and teaching that are essential to the improvement of society, and that will help protect higher education institutions against the economic and ideological assaults that are certain to come (pp.71-72).

Many colleges and universities and their accrediting bodies are currently undergoing transitions brought on by changes in the economy and the political landscape with more and more universities acting like for-profit enterprises. The accreditation process has shifted from one focused on self-study primarily based on inputs to the educational process, to an evidence-based, external accountability-oriented, outcomes-based model which focuses significantly on student learning outcome measures and assessment. The pressures of the new economy bring new challenges to institutional research and assessment offices as they conduct traditional duties of

**“...there may be a tendency to measure types of student outcomes that are easily measurable but not very meaningful.”**

program review, assessing faculty and student outcomes for accreditation and other accountability-related tasks. There is always the danger that evaluation, which has always been part of the legitimating functions of government, may

devolve into mere institutional impression management rather than a useful tool for formative improvement. Will evaluation become a mere tool of “scientifically informing, legitimizing, and controlling” (House, 1993, p. 33)?

There is a tension between public and professional accountability in higher education. Institutional researchers and evaluators are often in “conflicting and ambivalent circumstances” balancing demands of state or public accountability, professional accountability and consumer accountability (House, 1993, pp. 34-35). Institutional researchers are often caught in the middle of varying definitions of quality held by faculty, accreditors, students, parents, government officials, and trustees” (Volkwein, 2010, p. 10; see also Terenzini, 2010). Changes in the current era of neoliberal ideologies may encourage tendencies to ignore tried and true methods of university program review that include the perspectives of faculty members in the process. Or there may be a tendency to measure types of student outcomes that are easily measurable but not very meaningful. Are evaluators’ loyalties to central or local concerns or to political or professional authority?

Evaluation and assessment are traditionally approached as a highly technical-rational field dominated by assumptions of neutrality, objectivity, rationality and technical details of psychometrics. Rarely do we stop to examine the assumptions of our job and why we are doing it. In addition, in college and university settings, assessment is part of a top-down institutional framework where questioning basic assumptions is not rewarded and is often times counter-productive to expectations of job performance and persistence within the bureaucracy. However, scholars of assessment understand that the field is embedded in a social, political, and historical context that shapes how we practice. We are part of a profession with professional goals and responsibilities. This is discernable from the agendas of professional groups such as the Virginia Assessment Group’s agenda for its 2011 annual meeting which lists among other goals:

- Engaging faculty in assessment policy and decision-making
- Developing leadership in assessment across campus
- Data quality versus quantity

## Numbers and Data Quality

Institutional researchers attend primarily to quantitative indicators, survey data, assessment scores, performance data, and financial data. Most assessments and measures rely on numbers. Numbers are seemingly

technical, neutral, and rational but upon inspection, politics and interpretation are behind mere numbers. Numbers appear to have some cultural cache as the language of science. As Stone (2002) writes in her book *Policy Paradox*:

"Numbers...are measures of human activities, made by human beings, and intended to influence human behavior. They are subject to conscious and unconscious manipulation by the people being measured, the people making the measurements, and the people who interpret and use measures made by others" (p. 177). Stone explains why counting is political not technical:

- Counting requires decisions about categorizing, about what (or whom) to include and exclude.
  - Measuring any phenomenon implicitly creates norms about how much is too little, too much, or just right.
  - Numbers can be ambiguous, and so leave room for political struggles to control their interpretation.
  - Numbers are used to tell stories, such as stories of decline ("we are approaching a crisis").
  - Numbers can create the illusion that a very complex and ambiguous phenomenon is simple, countable, and precisely defined.
  - Numbers can create political communities out of people who share some trait that has been counted.
  - Counting can aid negotiation and compromise, by making intangible qualities seem divisible.
  - Numbers, by seeming to be so precise, help bolster the authority of those who count.
- (2002, p. 176)

She goes on to warn us about reactivity in measurement because measures always carry "implicit norms." We attend to what we measure. When we have a hammer, everything becomes a nail. When we have the light of some off-the-shelf- assessment, we look for our lost keys there. Statistics are often used by those wishing to begin a reform effort (e.g., accountability in higher education), to show how bad things have become in order to support a narrative of decline. Numbers never stand on their own; they are always interpreted as part of a story line. Measurement exerts a strong force over behavior, and not always in a positive manner. Raising the level stakes of standardized tests means teachers will teach to the test or worse cheat on the test. The statistics and data we collect are collected in the interests of the organization and are

therefore often inseparable from the goals and agendas of those organizations. Stone (2002) laments: "Because policy measurement is always linked to benefits and penalties, the measured try to influence the measurers, occasionally with outright bribes, but more often with pleading, cajoling, and selective disclosure" (p. 182).

According to Stone (2002), so much of what we

**"Numbers...are measures of human activities, made by human beings, and intended to influence human behavior. They are subject to conscious and unconscious manipulation by the people being measured, the people making the measurements, and the people who interpret and use measures made by others."**

count depends on the categories we have constructed to count around. Category construction is a qualitative, interpretive, and ultimately a political activity. The discretion used in determining if one behavior fits in this or that category influences the numbers we produce, yet we rarely stop to question the category construction. Program costs can include just the cost of the program or it might also include the cost to design the program. The selection of measures almost always indicates some preferred solution to a problem. For instance, a foundation may be supporting research about the lack of student learning in higher education and may choose measurements that suggest a particular solution that positions the foundation to obtain more funding for its program of change and improvement.

In the current climate of defining higher education policy problems, measurement plays a strategic role. Critics assert the policy problem is that colleges and universities are inefficient and do not produce learning. Numbers constructed by researchers that purport to show that undergraduates do not learn much in their first two years of college (Arum & Roksa, 2011) are used to support that a policy problem is growing. Measurement is always tied to a purpose: accountability, program improvement, cost efficiency, etc. Stone (2002) asserts: "Numerals hide all the difficult choices that go into a count" (p. 177).

Volkwein (2008) proposes a typology of four possible IR purposes and roles. These vary by the purposes and audiences (formative/internal for improvement vs. summative and external for accountability) and whether the organizational role and culture emphasizes the administrative/institutional vs. academic/professional roles. If the role is one of administrative/institutional the job becomes one of describing and gathering facts about the institution (under the formative purpose) or to perform the role of spin-doctor in institutional impression

management. If IR departments and individuals take on the academic and professional role, the job becomes one of analyzing alternatives as the IR takes on the role policy analysis (under the formative purpose) and IR as scholar and researcher providing impartial evidence of effectiveness (under the summative purpose). Given the external conditions facing higher education in a time of neo-liberal ideologies, the commercialization of higher education, and increased political clamoring for accountability, I want to argue that institutional researchers engage the professional identity in terms of organizational roles (policy analysts and scholar/researcher) and that they think carefully about the use of numbers as they face the challenges of the new era in higher education. In the context of the massive changes in society that are having such a significant impact on higher education at this historical moment, we must ask: at what point does “thoughtful noncompliance” (Stein, 2004) enter into the thinking of institutional researchers? How will we use the authority of numbers?

## References

- Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses*. Chicago, IL: University of Chicago Press.
- House, E. R. (1993). *Professional evaluation: Social impact and political consequences*. Newbury Park, CA: Sage.
- Stein, S. J. (2004). *The culture of education policy*. New York, NY: Teachers College Press.
- Stone, D. A. (2002). *Policy paradox: The art of political decision making*, (Rev. ed.). New York, NY: W.W. Norton.
- Terenzini, P. T. (1999). On the nature of institutional research and the knowledge and skills it requires. *New Directions for Institutional Research*, 1999(104), 21.
- Terenzini, P. T. (2010). Assessment with open eyes: Pitfalls in studying student outcomes. *New Directions for Institutional Research*, 2010, 29-46.
- U.S. Department of Education (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, DC: U.S. Department of Education.
- Volkwein, J. F. (2008). The foundations and evolution of institutional research. *New Directions for Higher Education*, 2008, 5-20.
- Volkwein, J. F. (2010). The assessment context: Accreditation, accountability and performance. *New Directions for Institutional Research*, 2010, 3-12.

## BOOK REVIEW:

*Academically Adrift: Limited Learning on College Campuses.* Richard Arum and Josipa Roksa. Chicago, IL: University of Chicago Press, 2011. 272 pp. ISBN-13: 978-0-2260-2856-9. Paperback, \$25.00

Reviewed by Kathryne Drezek McConnell, Ph.D.  
Virginia Tech

To say that *Academically Adrift* is the “must read” for every assessment professional for 2011 is an understatement. Richard Arum and Josipa Roska’s book has received more attention than most publications about student learning in higher education, which means the work of assessment on college and university campuses will once again be under scrutiny from a varied group of constituents and otherwise interested parties. As such, careful reading and consideration of *Academically Adrift* may prove essential to all assessment professionals attending committee meetings and cocktail parties alike. While the overall message of the book – that institutions must take greater responsibility for student learning – resonates, it is in the details of the research design and subsequent conclusions that one must proceed cautiously before extrapolating beyond the logical limits of the study.

**“To say that *Academically Adrift* is the ‘must read’ for every assessment professional for 2011 is an understatement.”**

The relatively short book is organized into five chapters plus a detailed “Methodological Appendix,” leading the reader first through a brief history of *College Cultures and Student Learning* (the title of the first chapter) to highlight “four core ‘important lessons’”, specifically that:

1. Colleges and universities, and the students who attend them, are ‘academically adrift’;
2. Gains in student performance, as measured by the Collegiate Learning Assessment (CLA), are disturbingly low;
3. Individual learning is characterized by persistent and/or growing inequality; and
4. Student performance on the CLA varies greatly both within and across institutions.

(Arum & Roksa, 2011, p. 30)

The three following chapters serve to delineate each lesson. The second chapter guides the reader through such topical minefields as racial and ethnic inequality

in CLA performance, educational experiences, and outcomes. Chapter three examines students’ academic engagement and their experience of college life by looking at self-reported data on their interaction with faculty and peers, courses taken, amount of homework completed, as well as how they were financing college. Both chapters lay the groundwork for the fourth chapter, in which they explore the relationship between the self-reported data, other available student information (e.g., SAT/ACT scores, GPA, demographics), and CLA scores. These relationships coalesce as the proposed conceptual framework (p. 118), which visually represents the relationships between pre-college factors, pre-test scores on the CLA, college factors, and CLA post-test scores. Chapter four essentially concludes that not only is American undergraduate education broken, but everyone – administrators, policy makers, faculty, parents, and the students themselves – is complicit.

The final chapter, with its charge-taking title *Mandate for Reform*, argues that rather than dismantling the entire higher education system within the United States, “all higher education institutions could focus increased attention on the academic component of undergraduate learning without fundamental challenge to the existing system” (p. 129). Appropriate changes – and by appropriate, the authors remind the reader that they mean changes designed to improve students’ performance on the CLA’s measures of critical thinking, complex reasoning, and written communication (p. 141) – include such things as having students take more courses that require 40 pages of reading per week and 20 pages of writing per semester, and de-emphasizing current trends toward collaborative learning in classes. Throughout the book, Arum and Roksa artfully intersperse results from a variety of other studies of higher education, from empirical investigations of student engagement to one faculty member’s anthropological exploration of students’ experiences of their first year at her own institution, to make their case. *Academically Adrift* makes for a very compelling narrative.

Because the book is so compelling, it is important to proceed cautiously in how one uses it. As a touchstone for important, perhaps even difficult conversations about what a college or university values educationally and how it translates those values into measurable and meaningful student learning outcomes for *all* students, *Academically Adrift* holds great promise. As a specific roadmap toward reforming undergraduate education, it is less powerful. Despite the tone of its title, the final chapter of the book is more circumspect, as the authors reference some of the limitations of the study by explicitly acknowledging the need to move beyond using observational data of the kind they examined to longitudinal studies that

utilize experimental and quasi-experimental designs. Furthermore, they stipulate that while the CLA is a promising measure from a “sociological perspective”, they admit that “we are not at a stage of scientific knowledge where college students’ learning outcomes can be measured with sufficient precision to justify embracing a coercive accountability system without significant reservations” (p. 141). But do such admonitions and caveats go far enough to acknowledge the limitations of the CLA specifically, or the challenges and limitations of assessing student learning more generally?

One of the most compelling lessons supposedly learned from this study is compromised by sample size issues. The authors note that while there is variation between institutions, there is even more variation within institutions, and that there are high-performers (students who scored in the top 10% of the CLA growth distribution) at every campus in the sample. The next logical step would be to look within and carefully examine the data at individual institutions to see what worked and what did not when it came to developing critical thinking, complex reasoning, and writing skills. Due to issues with the sample, however, such an analysis is not possible. So while institutions matter, and while at every institution in the sample something worked to help develop students’ learning as measured by the CLA, it is difficult to know what else might help explain variance within institutions otherwise absent the conceptual framework from *Academically Adrift*.

While the authors go to great lengths to demonstrate the representativeness of the convenience sample used in their study, they give less attention to the myriad specific psychometric issues raised about standardized testing for accountability in general (e.g., Banta, 2008) and the CLA specifically (Pike, 2006; Shermis, 2008). Without rehashing in detail territory well-covered by measurement experts, instead it may be helpful to remember a few words of advice offered by Patrick Terenzini in his piece “Assessment with Open Eyes: Pitfalls in Studying Student Outcomes”. First published in 1989, and reprinted in 2010, this thoughtful essay holds several nuggets of truth applicable to any reading of *Academically Adrift*. First and foremost, the reader must remember that “research design is a series of compromises” (Terenzini, 2010, p. 38).

The authors’ major compromise was in the use of the dataset itself; they themselves admit to feeling frustrated at the lack of existing longitudinal datasets for

higher education like those found for K-12 education (Arum & Roksa, 2011, p. 19). Key to this data was the use of the CLA as the sole performance measure of student learning. The CLA utilizes a *value-added* framework; that is, it is designed to try to isolate the change in student learning that is attributable to or *caused by* the institution. Again, Terenzini reminds us that *value-added* is both metaphor and research design. As a metaphor, it borders on the intuitive, but as a statistical operation, *value-added* is more problematic, and “potentially more dangerous” (Terenzini, 2010, p. 42). The danger rests in its unreliability, whether due to a lack of random assignment (despite idiosyncrasies in admissions processes, students are not randomly assigned to attend different institutions), ceiling effect, or regression toward the mean from the pre- to the post-test (Terenzini, 2010), never mind the question of student motivation on performance measures devoid of context or connection to their courses or other educational experiences. And while the authors suggest that institutional “internal self-assessment efforts ideally would be built on a diverse set of measures tracking teaching

and learning within an institution” (Arum & Roksa, 2011, p. 139), of which the CLA may be one measure, they give short shrift to what such diverse measures beyond the CLA might include. For example, they make passing reference to research commissioned by the Association of American Colleges & Universities (AAC&U) about the potential power of capstones and portfolios for student learning outcomes assessment without any mention

of that same organization’s national effort to develop a common set of rubrics to measure outcomes like critical thinking and written communication, among others, under the auspices of the VALUE (Valid Assessment of Learning in Undergraduate Education) project. From an assessment perspective, that omission is a glaring one when writing about measuring student learning and, more importantly, improving the educational experience at the undergraduate level—an area in which AAC&U is arguably and increasingly viewed as the primary organizational voice from an advocacy and policy perspective.

The disconnect, from an assessment perspective, is not that *Academically Adrift* explored potentially ugly truths about undergraduate education, but that its very tone and tenor lends itself to grander generalizations than current data supports. For the assessment professional, the power behind *Academically Adrift* is that it raises critical questions and posits directions for future research that should be explored by institutions. In turn, the assessment

**“As a touchstone for important, perhaps even difficult conversations about what a college or university values educationally and how it translates those values into measurable and meaningful student learning outcomes for all students, *Academically Adrift* holds great promise. As a specific roadmap toward reforming undergraduate education, it is less powerful.”**

challenge of *Academically Adrift* will be identifying its strengths while addressing its limitations to others without killing the messengers. Arum and Roksa's work, if viewed as a thought-provoking first step in what should be a long line of research, may be another opportunity for promoting student learning assessment on campus.

### References

- Banta, T.W. (2008). *Editor's notes: Trying to clothe the emperor*. *Assessment Update*, 20(2), p. 3-4, 15-16.
- Shermis, M.D. (2008). *The collegiate learning assessment: A critical perspective*. *Assessment Update*, 20(2), p. 10-12.
- Pike, G. R. (2006). *Assessment measures: Value-added models and the collegiate learning assessment*. *Assessment Update*, 18(4), 5-7.
- Terenzini, P. T. (2010). *Assessment with open eyes: Pitfalls in studying student outcomes*. In J. Fredericks Volkwein (Ed.), *New Directions in Institutional Research Special Issue: Assessing Student Outcomes*, (S1), 29-46.

## BOOK REVIEW:

*Higher Education Assessments: Leadership Matters.* Edited by Gary L. Kramer and Randy L. Swing. Lanham, MD: Rowman & Littlefield Publishers, Inc., 2010. 288 pp. ISBN-13: 978-1-4422-0620-5. Cloth, \$49.95

Reviewed by Tisha M. Paredes,  
Old Dominion University

*Higher Education Assessments: Leadership Matters*, edited by Gary L. Kramer and Randy L. Swing (2010) is a compilation of landmark assessment research from multiple higher education assessment scholars: Angelo, Banta, Ewell, Kuh, Pasarella, Terenzini, and Tinto. Each chapter in *Higher Education Assessments* summarizes assessment research and best practices, thus providing higher education leadership with what they “need to know and do to lead assessment successfully on campus” (Kramer & Swing, 2010, p. XV). Although the book focuses on what senior leadership needs to know about assessment practices, different types of leaders, such as assessment professionals and faculty, will benefit from reading this book.

### Book Summary

*Higher Education Assessments* is divided into three parts: (a) Leading Assessments on the Campus, (b) Bridging Learner Outcomes: Finding Common Ground, and (c) Assessments That Transform the Learning Culture. The first part, Leading Assessments on the Campus, sets the tone for the book by focusing on what senior leaders need to know about assessment, ultimately providing leaders with a broad assessment primer. The first section of Chapter 1 provides leaders with a common definition and background of assessment as well as a broad overview of assessment methodologies. The last part of the chapter discusses the challenges leaders may face and suggests ways to foster a culture of assessment. Bers and Swing (2010) assert that when assessments are carried out just to say an assessment was performed, then assessments become unusable and unjustifiable. They suggest that campus leaders continually guide assessment efforts and focus the campus on using assessment results for improvement. To achieve a sustainable assessment program, Bers and Swing list six imperatives that must occur. From setting expectations to building and sustaining an infrastructure in which assessment professionals “serve as mentors, coaches, and consultants” (Bers & Swing, 2010, p. 22) to using the results and celebrating the success, each imperative guides senior leaders thought the process of developing a sustainable assessment program. Without all of these components, Bers and Swing suggest that leaders

will have a difficult time establishing a quality assessment program.

Chapter 2, “Assessment Frameworks That Can Make a Difference in Achieving Institutional Outcomes,” builds upon Chapter 1 by discussing areas in which senior leadership should concentrate, such as identifying components of an institution-wide assessment plan and building a culture of evidence that reinforces institutional goals and mission. A common theme throughout chapter is that building a systematic, contiguous, and purposeful assessment program requires long-term commitment and buy-in from all institutional stakeholders. The last section of this chapter is dedicated to providing senior leaders with a framework to create a culture of evidence by incorporating factors and principles of assessment plans into a model for the institution.

The next section in *Higher Education Assessments*, Bridging Learner Outcomes: Finding Common Ground, concentrates on student-centered assessment practices that account for a diverse student body, incorporate out-of-class experiences, and lead to program improvements. Chapter 3, “Assessment and Student Diversity” starts out with an overview of the historical changes to the U.S. higher education system that created opportunities for students from diverse backgrounds to enroll in colleges and universities. Torres (2010) asserts that diversity incorporates more than race

### “What can senior leaders do to rally assessments around improving student success, learning, and development?”

and gender and assessment practices should consider students’ multifaceted background. Different types of diversity assessments, such as structural diversity, curricular and co-curricular activities that influence attitudes, and assessment of campus climate are outlined. The last section of this chapter focuses on the disaggregation of data by different groups (i.e. gender, race, first-generation) so campus leaders can focus on students who are succeeding and those who need assistance. Torres (2010) states the importance of focusing intervention efforts:

When institutions take on the attitude that an intervention for all students will eventually help those who actually need the help, the situation that results has the student who needs the intervention looking for the needle in the haystack. Focusing interventions on the population of students who need the assistance is the more direct manner in which an institution can truly improve success rates of students; well-intentioned interventions

that are not focused on the problem will likely produce marginal gains in student success. (pp. 67-68)

The second chapter in this section, “Assessment in Student Services That Foster Student and Program Successes” (Chapter 4), starts out with the assertion that assessment initiatives in student affairs have grown to produce valuable information on programs and services that lead to student success. An overview of student involvement and engagement theories provides the reader with a basic understanding of how student experiences can lead to student success. Lastly, the chapter examines assessment models for student services which provide the reader with best assessment practices in student affairs.

The main argument in the next chapter, “Documenting Student Learning: Valuing the Process” (Chapter 5), is the notion that when or if leaders only

**“...when assessments are carried out just to say an assessment was performed, then assessments become unusable and unjustified.”**

perform assessment functions because of external pressures, then student learning is lost and improvements cannot be made. Smith and Barclay (2010) assert that assessments designed to elicit cognitive and metacognitive processes from students increases the likelihood that students will learn concepts presented to them. They discuss cognitive engagement, a concept in which students are exposed to complex learning environments that promote deeper levels of learning, such as synthesizing and integrating knowledge (Smith & Barclay, 2010). After assessment for learning concepts are explored, Smith and Barclay discuss assessment projects, such as the Association of American Colleges and Universities’ (AAC&U) VALUE (Valid Assessment of Learning in Undergraduate Education) project and University of South Florida’s Cognitive Level and Quality of Writing Assessment (CLAQWA) program as examples of best practices that elicit deeper learning from students.

Lastly, Chapter 6 “Learning Outcomes, Assessment, and Program Improvement,” focuses on what faculty need to know about creating assessment plans that improve what students learn in their courses (Osguthorpe, Bradley, & Johnson, 2010). The chapter reviews common questions faculty should ask when they are creating learning outcomes, developing assessments of the outcomes, and evaluating possible changes. At the end of the chapter, a set of principles designed to help faculty and administrators evaluate the quality of their assessment plan is presented.

The last section in the book *Higher Education Assessment, Assessments That Transform the Learning Culture*, is focused on assessments that transform the

culture. The first chapter in this section, “Student Engagement and a Culture of Assessment” (Chapter 7), provides a student engagement framework and outlines nine characteristics of assessments that are associated with improving student learning (Kinzie, 2010). These nine characteristics direct leaders through the process of developing an assessment culture that promotes student learning and success. Kinzie (2010) suggests that institutions that want to enhance student experiences and student learning need to focus on assessing student engagement and that the “information gained through student engagement data is useful in all institutional assessment agendas intent on improving student learning” (p. 138).

Chapter 8, “Assessment in the Disciplines” changes the focus of the book, shifting to assessments at the college, department, and unit/program level (Muffo, 2010). The author suggests that senior leaders should know about disciplinary-level assessment, paying particular attention to disciplines with accreditation organizations, in order to understand what programs and disciplines are doing to assess student learning and why they do it. After a brief explanation of accreditation organizations, Muffo (2010) spends some time explaining engineering’s accrediting body – Accreditation Board for Engineering and Technology (ABET). He suggests that ABET’s model can be used by many other disciplines because they outline student learning outcomes that are transferable (i.e. functioning on multidisciplinary teams, communicating effectively, and engaging in lifelong learning). The last section of this chapter discusses future trends in federal and state-level accountability, and it is suggested that higher education institutions need to provide leadership and directions to federal and state governments by adopting accreditation-like models of assessment where applicable.

The next chapter of the book, “Assessment That Transforms an Institution” (Chapter 9), provides the reader with a guide on how to create an assessment culture that “becomes embedded in the culture as a systematic and continuous process of improvement and as a way to provide valid and reliable information for accountability” (Gray, 2010, p. 180). Gray (2010) emphasizes that by making assessment meaningful and manageable to faculty, leaders can provide a sustainable institution-wide assessment structure. Throughout the chapter, the author presents a clear and compressive plan that will make assessment transformative. At the end of the chapter, the question “why read this chapter?” is asked. As a response, a letter from an academic dean and provost is presented. Dr. Andrew Phillips details how he came to value assessment and suggests how assessment should be framed as “a vehicle for BETTER education, for BETTER learning, for BETTER program design” (p. 209) by leadership in

order to gain faculty acceptance of assessment initiatives.

Chapter 10, “Putting Student First as Partners in the Learning Experience” discusses how institutions can partner with students and engage them in the learning process. Kramer and Miller (2010) suggest that by engaging students as partners in the learning process, institutions will be able to align student expectations and experiences with institutional outcomes, create assessment measures that are intentional and useful, and make certain that changes to experiences or programs will affect student success. The authors provide several examples of how to engage students in the assessment

process and give readers an overview of a few best practices from various institutions. At the end of the chapter the authors ask the question Randy Pausch proposed to his students: “If we were to vanish tomorrow, what would we want as our legacy?” (Kramer & Miller, 2010, p. 233). Their hope is that, after reading their chapter, our answer will be “put students first as partners in the learning enterprise” (Kramer & Miller, 2010, p. 233).

Finally, *Higher Education Assessments* concludes with an epilogue from Gary L. Kramer and Randy L. Swing. They start out asking the question “what can senior leaders do to rally assessments around improving student success, learning, and development?” (Kramer & Swing, 2010, p. 237). Throughout the book, the chapter contributors have given the reader several ways to answer this question, but all agree on one factor that has a significant impact on building an assessment culture centered around improving student learning – leadership. Without a leader who monitors assessment practices and uses assessment data to drive decisions, the campus will not adopt a sustainable assessment culture. Kramer and Swing (2010) conclude with several strategies for senior leadership that will help them build a culture of assessment, all of which support the idea that assessment practices must be “purposeful, integrated, and built on and reinforced by one another” (p. 240).

## Review

*Higher Education Assessments* is a good basic resource for senior leadership and new and seasoned assessment professionals. The book provides a comprehensive compilation of best practices in assessment research. One of the main strengths of this book is that most of the chapters rely on seminal assessment research studies and scholarly endeavors. Using assessment concepts from leaders in the assessment field, such as Angelo, Banta, and Kuh, provides a complete assessment resource that is ground in researched best practices. Additionally, although chapter authors acknowledge that there is no one size fits

all assessment program, examples from other institutions provide the reader with a guide for implementing the assessment concepts discussed in the book.

A limitation of this book is that the intended audience is campus presidents, vice presidents, provosts, and deans. With the exception of a few chapters, many of the concepts and ideas discussed are more for assessment directors and coordinators to implement, not senior leadership. *Higher Education Assessments* gives senior leaders a road map on how to make assessment transformative on their campuses; however, deans and provosts will not be the people directly helping faculty

with assessment practices within the major. To truly transform a campus, all levels of professionals involved with assessment need to be involved and knowledgeable about assessment practices.

Finally, although a seasoned assessment professional may have extensive knowledge of many of the concepts offered in *Higher Education Assessments*, there are some ideas that will help assessment professionals gain the support of senior leadership at their institutions. Assessment professionals can utilize some of the suggestions in the book to revitalize their campus-wide assessment program. Ultimately *Higher Education Assessment: Leadership Matters* is a welcomed addition to assessment literature and is another good assessment resource for assessment professionals.

**“...diversity incorporates more than race and gender and assessment practices should consider students’ multifaceted background.”**



## Ruminate: Integrating The Arts And Assessment



### Standardized

**Casey Templeton**

Photographer  
Richmond, Virginia  
[www.caseytempleton.com](http://www.caseytempleton.com)

Image copyright belongs to Casey Templeton Photography. No part of this image may be extracted from RPA or reproduced in any form without permission from the artist. For additional use or distribution, please contact the artist at [casey@caseytempleton.com](mailto:casey@caseytempleton.com)



### **Article Submissions:**

Articles for *Research & Practice in Assessment* should be research-based and include concrete examples of practice and results in student learning outcomes assessment or other education assessments. The readers of *Research & Practice in Assessment* are associated with myriad institutional types and have an understanding of basic student learning and assessment practices. Articles for publication will be selected based on their degree of relevance to the journal's mission, compliance with submission criteria, quality of research methods and procedures, and logic of research findings and conclusions. Approximately fifty percent of submissions are accepted for publication.

- Style: Submissions should be formatted in the 6th edition of APA.
- Format: Manuscripts should be submitted electronically. The standard manuscript form has 1 inch margins on all sides located in 12 point New Times Roman font.
- Blind Review: Identification information should be limited to the title page. Do not include self references within the manuscript, electronic file, or references (i.e. "in press").
- Length: articles should be 8-16 typed, double spaced pages (2,000-4,000 words)
- Abstract: An abstract of 150 words or less should be included.
- Copyright: Reviews shall not have been registered for copyright or published elsewhere prior to publication in *Research & Practice in Assessment*.
- Double Submission: Articles currently under review elsewhere shall not be considered for publication in *Research & Practice in Assessment*.
- Deadlines: Submission deadlines for the 2011 issues are April 1 and September 1, which are published in July and December respectively.

### **Book Review Submissions:**

Book reviews are significant scholarly contributions to the education literature that evaluate publications in the field. Persons submitting reviews have the responsibility to summarize authors' works in a just and accurate manner. A quality review includes both description and analysis. The description should include a summary of the main argument or purpose and overview of its content, methodology, and theoretical perspective. The analysis of the book should consider how it contrasts to other works in the field and include a discussion of its strengths, weaknesses and implications. Judgments of the work are permitted, but personal attacks or distortions are not acceptable as the purpose of the review is to foster scholarly dialogue amongst members of the assessment community.

- Style: Submissions should be formatted in the 6th edition of APA.
- Format: Manuscripts should be submitted electronically. The standard manuscript form has 1 inch margins on all sides located in 12 point New Times Roman font.
- Length: reviews should be 5-8 typed, double spaced pages (1,250 - 2,000 words).
- Copyright: Reviews shall not have been registered for copyright or published elsewhere prior to publication in *Research & Practice in Assessment*.
- Deadlines: Submission deadlines for the 2011 issues are April 1 and September 1, which are published in July and December respectively.

### **Notes In Brief and Special Features:**

The Board of Editors will consider notes in brief and special features such as invited articles or reports. These submissions should be no more than 3 typed pages in length and should address topics associated with student learning outcomes assessment. The topics addressed should be of interest to the readership of the journal and adhere to the formatting guidelines of submitted articles.

### **Submissions for Ruminate:**

Ruminate concludes each issue of *Research & Practice in Assessment* and aims to present matters related to educational assessment through an artistic medium such as photography, poetry, art, and historiography, among others. Items are encouraged to display interpretive and symbolic properties. Contributions to Ruminate may be submitted electronically as either a Word document or a jpg file.

**All items may be submitted in a Microsoft Word file attached via email to:**  
**[webmaster@virginiaassessment.org](mailto:webmaster@virginiaassessment.org)**