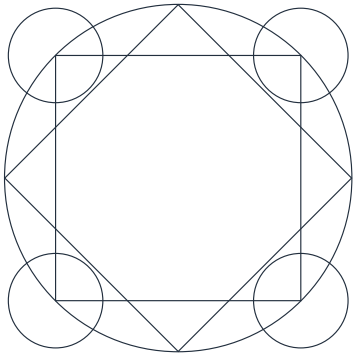

Abstract

A fundamental goal of student learning outcomes assessment in higher education is to use student learning evidence in decision making to improve educational programs. Such use of assessment findings, however, is atypical. This article argues that a narrow conception of use contributes to this conclusion and an accurate appraisal of the contribution of assessment requires a reconceptualization of the aims of assessment and a more inclusive model of possible uses of assessment evidence. To evaluate the heuristic value of a more inclusive model of influence adapted from the field of evaluation, a content analysis was undertaken of program assessment reports at a research university. Results indicate that existing definitions of use suffer from construct underrepresentation; assessment evidence may be more influential than realized, particularly with regard to enhancing understanding of how student learning occurs; and the more inclusive model has potential utility for faculty, administrators, and accrediting bodies.

AUTHORS

Jessica L. Jonson, Ph.D.
University of Nebraska-Lincoln

Tim Guetterman, M.A.
University of Nebraska-Lincoln

Robert J. Thompson Jr., Ph.D.
Duke University

An Integrated Model of Influence: Use of Assessment Data in Higher Education

A fundamental goal of student learning outcomes assessment in higher education[1] is use of student learning evidence to *close the loop* (Banta & Blaich, 2011; Banta, Jones, & Black, 2009; Bresciani & Wolff, 2006; Maki, 2010), that is, completing the assessment cycle that includes planning, gathering, interpreting, and using learning evidence to inform decision making about improving educational programs (Maki, 2010; Palomba & Banta, 1999). However, the realization of this goal has been one of the most important and unaddressed challenges related to assessment (Banta & Blaich, 2011; Kuh & Ikenberry, 2009; Kuh, Jankowski, Ikenberry, & Kinzie, 2014). Even the most well-designed and thorough studies of student learning have concluded that the use of available learning evidence is uncommon (Blaich & Wise, 2011). Concern with fidelity of assessment has resulted in a consideration of factors that hamper and facilitate use of assessment results (Banta & Pike, 2012; Blaich & Wise, 2011; Ewell, 2009; Kuh & Ikenberry, 2009; Peterson & Einarson, 2001). Nevertheless, an important issue that has not been addressed in the literature is whether a narrow conception of what constitutes *use* contributes to the conclusion that assessment results typically do not lead to improved educational practices and student learning. If definitions of use are too narrowly defined, some assessment efforts may be considered failures when those efforts actually may have been very transformative but in unexpected or slowly evolving ways. Accurate appraisal of the extent to which assessment of student learning is contributing to improvement in educational practices and student learning requires both a reconceptualization of the aims of assessment as a process for transforming thinking of internal and external stakeholders about teaching and learning and a more inclusive model of possible uses of assessment evidence.

CORRESPONDENCE

Email
jjonson@unl.edu

Banta (2002) has suggested evaluation literature as a basis for assessment scholarship. The discipline of evaluation provides a framework for the practice of assessment because both involve a systematic method for collecting, analyzing, and using information to

[1] Student learning outcomes assessment in higher education is hereafter referred to as *assessment*.

answer questions about the effectiveness of programs (Banta, 2002; Gray, 2002). In particular, evaluation theory provides a framework for considering a broader model of possible uses of assessment evidence.

This research article follows through on Banta's suggestion by applying theories of use that have evolved over time in the evaluation field to assessment practice (Kirkhart, 2000; Leviton & Hughes, 1981; Patton, 2012; Weiss, 1998). More specifically, Kirkhart's (2000) multidimensional and integrated theory of influence is adapted by this study to reconceptualize the aims of assessment and the construct of use and evaluate the heuristic value of a more comprehensive model of influence. This model potentially can be useful to faculty, administrators, and the bodies that accredit postsecondary institutions when determining the implications of assessment evidence for improving educational practices and policies.

Background

…most institutions have not monitored the use of their assessment information.

The infrequent use of assessment findings to improve academic programs is a key indication in recent literature on the current state of higher education assessment (Banta, 2010; Banta & Blaich, 2011; Blaich & Wise, 2011; Ewell, 2009; Ewell, 2010; Kuh & Ewell, 2010; Kuh & Ikenberry, 2009; Kuh et al., 2014). This lack of use was also recognized over a decade ago in an extensive multi-institutional study of assessment approaches, supports, and uses that found assessment results have very limited impact on academic and faculty related decisions (Peterson & Einarson, 2001). This study also found that most institutions have not monitored the use of their assessment information. Peterson and Einarson (2001) concluded that because of the claims made about the value of assessment and the substantial resources invested, institutions need to prioritize examination and evaluation of uses and effects. This article offers support for Peterson and Einarson's conclusion by suggesting a framework that would allow different dimensions and types of use to be more clearly defined for these monitoring purposes.

Extensive multi-institutional studies of the current state of assessment, like Peterson and Einarson's (2001) work, have been infrequent until the publishing of findings from a survey of provosts and chief academic officers at regionally accredited institutions about what their institutions are doing to gather and use evidence of undergraduate student learning (Kuh & Ikenberry, 2009; Kuh et al., 2014). Kuh et al. (2014) found that since the 2009 survey, the use of assessment evidence appears to be increasing but is still not pervasive enough to guide institutional actions that improve student outcomes. Another large-scale multi-institutional longitudinal study by Blaich and Wise (2011) involved the administration and analysis of multiple measures of student learning. One of the unforeseen results in Blaich and Wise's study was that despite a significant amount of credible data, most institutions had trouble finding tangible uses for the information. They found only 40% of institutions involved in the study had shared results with campus constituencies and only about 25% had actively used the data. Blaich and Wise concluded with sound advice on how to engage institutional communities in the discussion and use of assessment data as a process of inquiry.

However, as this article will demonstrate, approaching assessment as a process of inquiry requires a more broadly defined framework for use that honors a slow, but measurable, four-step process across the assessment cycle. The first step is planning assessment as a process of inquiry that focuses on faculty's questions of interest. The second step is gathering data about student learning and the assessment process. The third step is interpreting and evaluating the data collected by engaging stakeholders in *meaning making*, an epistemological process where social construction of meaning is arrived at through dialogue. Engagement in meaning making process has occurred among faculty and student affairs personnel using both quantitative and qualitative information. Baxter Magolda and King (2007) discussed the use of the meaning making in the interpretation of self-authorship interviews and Driscoll and Wood (2007) discussed its use with faculty learning communities. The fourth and final step is the use of assessment results for improving teaching and learning.

Concerns about the persistent lack of use of assessment evidence has led some to call for studying the effects of outcomes assessment on decision making (Banta 2010; Blaich & Wise, 2011; Ewell, 2009; Ewell, 2010; Kuh & Ikenberry, 2009; Kuh et al., 2014; Spencer Foundation, 2010). Granting agencies, such as the Spencer Foundation (2010), recognized the

use of assessment evidence in decision making as a critical issue in K-12 and higher education by identifying data use and educational improvement as one of their strategic initiatives and offering funding "to support scholarship examining the conditions, contexts, and underlying factors and processes that affect how educational organizations use data and information for improvement" (para. 1). This article is an outgrowth of a grant project funded by the Spencer Foundation's strategic initiative on data use that prompted the need to clarify the meaning of the term *use*. The implicit assumption evident in previous research is that use is only defined by an immediate, observable action for change. However, this narrow definition may underrepresent the construct of interest and limit studies of what conditions, contexts, and underlying factors are most likely to produce meaningful processes and data for decision making and ultimately improvement of student learning. Underrepresentation of use has not been previously discussed in the higher education assessment literature and therefore is a key focus of this study.

Underrepresentation of the Construct of Use

Approaching assessment as a process of inquiry requires a more broadly defined framework for use that honors a slow, but measurable, four-step process across the assessment cycle.

Validity inferences are an important concept in assessment because "broadly defined validity would be nothing less than an evaluative summary of both the evidence for and the actual – as well as potential – consequences of score interpretation and use" (Messick, 1995, p. 742). Messick's (1995) discussion of validity indicates that a comprehensive measure of a construct can encounter two major threats: construct underrepresentation and construct irrelevant variance. In this study, construct underrepresentation of assessment use is of particular concern and involves a definition of a construct that is "too narrow and fails to consider important dimensions or facets of the attribute" (Messick, 1995, p. 742). If the assessment process is only considered successful when the data contribute to immediate, observable actions for change, then the aims of the assessment process and the existing definition of use are too narrow. The field of evaluation has reconceptualized the aims of assessment and developed a broader understanding of what constitutes use (Kirkhart, 2000; Leviton & Hughes, 1981). A consideration of this reconceptualization offers an opportunity to identify important influences and utilizations of student learning data that may be contributing to improvement of academic programs but may not be immediate or observable.

Contribution of Evaluation Theory

The assessment and evaluation communities share a concern about the effects of their assessment and evaluation efforts (Banta, 2002). There is much the assessment community can learn from evaluation because for over 30 years the field has studied use, conducting empirical studies in addition to conceptual work. The guiding assumption about the aims of assessment has been "that evaluations are conducted to provide information for use in decision making" (Preskill & Torres, 2000, p. 26), which has led to a focus on types of use and the factors that contribute to use or nonuse of the information obtained. Several types of use have been delineated, including *instrumental*, in which the findings influence actions or decision making; *conceptual*, in which the evaluation leads to different understandings or enlightenment; *symbolic*, in which the findings are used for advocacy, argument, and persuasion (Leviton & Hughes, 1981); and *process*, in which learning occurs as function of participating in the evaluation process (Patton, 1997). Factors that have been identified as contributing to the use or nonuse of evidence in educational contexts include organizational characteristics, such as the extent to which assessment is being conducted and supported (Peterson & Augustine, 2000); personal factors, such as beliefs about knowledge (Weiss & Bucuvalas, 1980); and information characteristics, such as whether the evidence is fine-grained enough to guide intervention (Ewell, 1989).

Although there have been advances in understanding the effects of evaluation efforts, Kirkhart (2000) argued that "an inclusive understanding of the influence of evaluation has been hampered by the scope and language of past approaches" (p. 5). More specifically, she contended that the term *use* is awkward and has resulted in construct underrepresentation, that is, an inappropriate emphasis on some components while neglecting other critical aspects of the construct (Messick, 1995). Kirkhart advocated for a shift in construct terminology from *use* to *influence*, which called attention to a more comprehensive understanding about the effects of the evaluation endeavor and proposed an integrated theory that conceptualizes evaluation influence in three dimensions: source, intention, and time.

Source refers to the active agent or starting point of the influence (Kirkhart, 2000). There are two well-recognized sources of influence: *Findings-based*, in which the influence stems from the information or data produced by the evaluation (Kirkhart, 2000; Rich, 1977); and *processes-based*, in which the influence stems from the process of conducting the evaluation (Greene, 1988; Patton, 1997). Findings-based evaluation has been further characterized in terms of three types: *instrumental* in terms of direct action taken on the basis of the evidence; *conceptual*, with regard to changes in understanding stimulated by the findings; and *symbolic,* in the sense of the role of the findings in advocacy, argument, and persuasion (Leviton & Hughes, 1981; Weiss, 1998; Weiss & Bucuvalas, 1980). Similarly, process-based influence as experienced by participants has been reported along three dimensions: *cognitive,* in terms of discussion and processing of information; *affective*, in terms of personal feelings of value and worth; and *political*, in terms of recognition and voice to the less powerful (Greene, 1988). This article proposes an adaptation of Kirkhart's categories that findings-based and process-based sources and their respective subcategories and dimensions can be interrelated such that findings-based evaluations may have an affective effect of influence and process-based evaluations may have an instrumental effect of influence.

In addition to source, Kirkhart (2000) proposed two additional dimensions. *Intention* refers to the extent to which the evaluation has intended or unintended influences. Intended use may be directed through both findings-based and process-based sources. Unintended use reflects unanticipated pathways that may also stem from findings-based and process-based sources. *Time* is the third dimension and refers to the period in which the influence occurs, either as an immediate point-in-time event, an end of cycle event, or as a more long-term process.

One of Kirkhart's (2000) contributions has been to foster the recognition and investigation of multiple attributes of the construct of influence. For example, Rossman and Rallis (2000) argue that evaluation serves an educative purpose and they cast evaluation as a process of learning in which evaluation use involves the generation and application of knowledge. This reconceptualization expands the purposes or aims of assessment and evaluation beyond problem solving that leads to decision making to also include learning, that is, new or enhanced understandings that facilitate meaning-making with regard to the implication of assessment evidence for improving educational practices and policies. Moreover, conceptualizing assessment and evaluation as learning enables a connection with constructivist approaches to learning as transformation, not only at the level of the individual but also with regard to organizations. More specifically, Preskill and Torres (2000) focus on both the learning that occurs through the evaluative process and the use of evaluation to facilitate learning, especially transformative learning in organizational contexts. They argued that, "Learning from evaluation and from organization members' subsequent use of what they learn will most likely occur when evaluation is collaborative, is grounded in constructivist and transformational learning theories, and builds communities of evaluation practice" (p. 29). A constructivist and transformational perspective on learning emphasizes the roles of collaboration and dialogue within an organization as essential to the process of making-meaning from participating in the evaluation enterprise. The ideas of Preskill and Torres expand thinking about the aims or purposes of the learning to be accomplished through the evaluation process. These include both personal transformation as well as building communities of practice in which evaluation and assessment are engaged as processes of inquiry. Communities of practice arise as members engage in common activities, rely on one another, and share decision making (Preskill & Torres, 2000; Wegner, 1988). Its members are bound to their institutions and share common problem sets, which they solve through peer review (Herndon, 2006). Communities of practice serve as faculty development mechanisms to foster and sustain dialogue about teaching, learning, and assessment issues and link ideas to effective practices (Jonson & Thompson, 2013). Examples of building these communities of practice have been introduced by Cox (2004) as faculty learning communities and by St. John (2009) in the professional development of graduate students.

If the assessment process is only considered successful when the data contribute to immediate, observable actions for change, then the aims of the assessment process and the existing definition of use are too narrow.

Therefore, the purposes of this article are (a) to enhance the influence of assessment efforts through incorporating advances from the field of evaluation theory; (b) to prompt a conversation about the underrepresentation of the construct of use in assessment as it applies to student learning in higher education; (c) to advocate for a reconceptualization of

assessment aims; and (d) to encourage a shift in focus from asking about the use of assessment information to asking about its influence. To these ends this article undertakes an evaluation of the heuristic value of a model of influence that adopts, and further adapts, Kirkhart's (2000) integrated theory of multiple and integrated dimensions of influence within a framework of conceptualizing assessment as a process of inquiry learning that involves the generation and application of knowledge (Rossman & Rallis, 2000) that can serve multiple and valued purposes (Preskill & Torres, 2000).

Proposed Model

The question to be addressed is: If assessment is conceptualized as an educative process of inquiry and learning that involves the generation and application of knowledge, how might the influence of the assessment enterprise be appraised? Table 1 presents a schema for an elaborated heuristic model with four dimensions of influence that are adapted from Kirkhart's (2000) model.

Table 1

Heuristic Model of Influence: Dimensions, Subtypes, and Definitions

Dimension	Subtype	Definition
Sources of Influence	Findings-based	Based on student learning evidence
	Process-based	Based on evidence about the process of assessment rather than on learning evidence including consideration of methodology or data (e.g., measurement issues, sample size).
Effects of Influence	Instrumental	Involves a direct action or a decision and commitment to take educational practice or policy actions.
	Conceptual/Cognitive	Involves new understandings, ways of thinking, or processing information that may lead to considering action but lacks the actual commitment to act.
	Affect	Involves participant's disposition, emotions, or tendency regarding assessment process or assessment evidence
	Affirmation	Involves a confirmation of the appropriateness or effectiveness of an existing practice, policy, or understanding.
Results of Influence	Improved student learning	Results in evidence of improved student learning.
	Personal transformation	Results in a personal transformation of stakeholders (e.g., feeling empowered and motivated, changes of beliefs).
	Communities of practice	Results in building new or strengthening existing communities of practice.
	Symbolic/Political	Results in generating or sustaining support for policies or practices.
Time of Influence	Immediate	Occurs concurrent with the assessment process.
	End of Cycle	Occurs surrounding the conclusion of an assessment cycle (e.g., end of term)
	Long-term	Occurs in the future or extends beyond the assessment cycle.

Note: The model definitions provided an existing code set for qualitative analysis of programmatic reports.

Conceptualizing assessment and evaluation as learning enables a connection with constructivist approaches to learning as transformation, not only at the level of the individual but also with regard to organizations.

The first dimension is sources of influence. There are two subtypes depending on whether the influence stems from the learning data (findings-based) or from participation in the assessment process (process-based).

The second dimension is effects of the influence, which act as mediating processes or functions between the assessment findings or processes and various educational practices, policies, or outcomes. The effects dimension includes four subtypes. *Instrumental* refers to an influence that prompts a direct action or decision regarding a change to educational practice, such as curriculum, pedagogy, or policy. *Conceptual/Cognitive* refers to an influence on the understanding of some issue, concept, process, or on ways of thinking. *Affect* refers to an influence on the participant's disposition, emotions, or attitudes regarding the assessment process or the evidence collected. *Affirmation* refers to an influence that involves support for the appropriateness or effectiveness of an existing understanding, practice, or policy.

The third dimension is the results of the influence, that is, the outcomes or consequences that result from the application, either intentionally or unintentionally, of the knowledge and understanding generated through the assessment process. Types of results of influence include *improved student learning; Personal transformation*, for example, empowerment, motivation, belief change; *Building/strengthening communities of practice*; and *Symbolic/political*, for example generating financial or institutional support for a policy or practice.

Time of the influence is the fourth dimension and refers to the point in the assessment cycle in which the influence occurs. The cycle is essentially four steps: planning, gathering, interpreting, and using the assessment evidence. *Immediate* refers to an influence that occurs concurrently with one of the steps in the assessment process. *End of cycle* refers to an influence that occurs at the conclusion of the final two steps of the assessment cycle (e.g., interpreting or using). *Long-term* refers to an influence that occurs in the future or extends beyond the conclusion of the assessment cycle.

The proposed model is multidimensional but the dimensions are not meant to be considered mutually exclusive. That is, a particular assessment effort could have multiple *sources, effects*, and *results*. Furthermore, dimensions may also be interrelated. For example, effects of influence can serve as mediational processes between the sources and the results that could occur at the end of a cycle or be a long-term consequence.

Method

To illustrate the application of the proposed model to assess activities in higher education, the authors collected qualitative data from 19 narrative reports documenting assessment methods, results, and conclusions. These reports come from a large Midwestern research university and were a part of an institutional process for assessing general education courses. The programmatic reports in the sample include the fields of humanities, social sciences, formal/natural sciences, and professional/applied sciences. The sample, however, was not selected to be representative of any particular population because the goal of the study was not to generalize the results but to demonstrate heuristic value of the proposed model and clarify its application to examples of assessment.

If assessment is conceptualized as an educative process of inquiry and learning that involves the generation and application of knowledge, how might the influence of the assessment enterprise be appraised?

The qualitative data analysis consisted of content analysis of 28 distinct cases documented in the 19 narrative assessment reports. Content analysis is a systematic technique for examining text using explicit coding rules (Stemler, 2001). The analysis focused on identifying influence using an a priori coding scheme in terms of the dimensions of sources, effects, and results of influence on programmatic conclusions about teaching and learning. The documentation in the assessment reports did not lend itself to classifying the fourth dimension of time because reports documented influences at a particular point in time and programs varied in terms of where they were in their assessment cycle; often somewhere between gathered evidence and implementation of changes. The assessment reports also did not elicit or document participants' dispositions or feelings regarding the assessment process or findings; therefore, classification of the affect subtype of the effect dimension was not possible. To ensure qualitative validity, the three authors independently coded the cases using a predetermined coding scheme for dimensions represented in the proposed model (Table 1). The researchers then compared their codes to determine the extent of agreement and to discuss any disagreements to clarify definitions. Next, qualitative codes were tallied based on each dimension presented in Tables 2 and 3 for Source by Effects and in Table 4 for Source and Effects by Result. The raters' independent classifications of sources of influence, effects of influence, and results of influence agreed for 93% of the cases. Raters disagreed on 2 of the 28 cases because of differences in the definition of process-based sources of influence, which will be discussed later. To further illustrate these findings, example report excerpts representing each type of influence appear in Tables 2 and 3. Words or phrases that might identify a specific course or program were removed and, in some cases, replaced with more neutral terms to protect the confidentiality of these programs.

Results and Discussion

Dimension 1: Sources of Influence

Tables 2 and 3 detail the number of cases that were classified for the source by effect dimensions and provide examples for each dimension. All of the analyzed reports documented some type of influence and nine of those reports documented two distinct influences for a total of 28 cases. Raters were able to quickly and efficiently make differential judgments for each dimension.

Table 2

Coding Results: Findings-based Sources of Influence and Effects of Influence Dimensions (n = 28 cases)

Effects of Influence	Case Examples
Instrumental (n = 6)	*Students write at an acceptable level but continue to struggle with expressing their ideas in a concise and readable way. Several adjustments have been made to help students produce acceptable writing. Specifically, a phased sequence of topic selection, outlining, rough drafts reviewed by instructor and peer groups members has been used.*
	Paper structure is being modified to focus on the application of core concepts. Specifically, common ... problems are presented to students; students select one and (i) identify three core concepts related to the problem, (ii) develop an intervention, (iii) discuss how the intervention will impact core concepts, and (iv) develop a plan to assess the efficacy of the intervention. Hopefully this will force student to relate what they learn to the (real world environment)
Conceptual (n = 6)	*Faculty submitted samples of various types of questions with a range of difficulty, indicating that students found defining terms and identifying images easier than the more interpretive, analytical question that links work with more than one issue/answer.*
	Based on the letters evaluated, students demonstrated good technical knowledge, but some are in need of improved writing skills. Additional writing exercises may need to be incorporated into subsequent courses, so that students have additional opportunities to enhance their writing skills.
Affirmation (n = 9)	*Results of artifacts that demonstrate students' mastery of the student learning outcome seem to affirm the effectiveness of the department's efforts.*
	Assessment results confirmed what we already knew from the (exam) reports that student learning (of) the learning outcomes is high. No program changes are planned.

Note: The narrative reports did not contain enough information to code the time dimension.

Table 2 includes examples of findings-based sources, and Table 3 presents examples of process-based sources. Findings-based sources (n=21) were three times more common than process-based sources (n=7). However, only 25% of the 28 cases were classified as findings-based, instrumental influences, that is, a direct action or decision about a program based on learning evidence. Findings-based, instrumental influences are typically emphasized by accreditation bodies because it represents an observable behavioral change made with the intention of improving student learning. However, intention does not guarantee that an improvement will occur, and often whether the learning improvement does occur is not determined. The results of this study suggest that acknowledging only findings-based, instrumental changes is limiting and shortsighted because other findings-based and process-based sources are just as frequent as, or more so, than findings-based, instrumental influences. The other 75% of cases may not involve immediate, observable changes but are ways assessment findings have affected how one thinks about student learning (conceptual/cognitive, affirmation, personal transformation), how one talks about student learning with others (symbolic, communities of practice), disposition toward future assessment evidence (affect), or how one continues to conduct assessments of student learning (process). In the long term, different types of influences have the capacity to contribute to future improvements in student learning because they are challenging and transforming thinking about teaching and learning through the generation and application of knowledge, as the work from Preskill and Torres (2000) on organizational learning suggests.

A particular assessment effort could have multiple sources, effects, and results.

Seven of the 28 cases (25%) were classified as process-based sources. These cases did not result in an immediate, observable change to the curriculum, co-curriculum, course, or teaching as a result of assessment process. What changed was how one thinks about assessing student learning or representing what students learn so credible information about student learning can be obtained for decision making. For 2 of the 7 process-based cases, the authors disagreed because of differences in the definition of process-based sources. For these two cases, data or findings were used to identify a change, but what prompted that change was not evidence of student learning. Rather, the prompt was a realization that the learning data and how it was collected was not a sufficient or credible representation of the learning outcome the program was trying to measure. Examples include concerns about inadequate measures (e.g., rubrics), design of the assessment process (e.g., pre/post-testing, longitudinal), strategies for sampling student work, or roles/responsibilities of program faculty/instructors. After discussing these differences and clarifying the definition for process-based sources, the raters reached agreement on all seven of the process-based cases. This disagreement does highlight that the influence of process-based sources can be subtle.

Table 3

Coding Results: Process-based Sources of Influence and Effects of Influence Dimensions (n = 28 cases)

Effects of Influence	Case Examples
Instrumental (n = 5)	*It is difficult to assess how well students are interacting with diverse cultures, and engaging with global issues unless the (product) being reviewed incorporates cultural and global issues. Therefore, some type of unifying textbook (that both incorporates global and cultural issues as they relate to (discipline) and had a reading list of (articles) that include global and diverse cultures is recommended.* *A scoring rubric was developed for the final reflection papers in (course). When implementing the rubric, it quickly became clear it was far too complex, so a template was developed with a built-in rubric.*
Conceptual (n = 2)	*A discussion of which faculty member (in the program) should be responsible for submitting a report and sampling for each course, and when, does need to be addressed.* *It is a challenge to compare five courses, taught by many different instructors, with one another. We will experiment with new ways to facilitate evaluation and assessment at the department level (e.g. creating a formal rubric).*
Affirmation (n = 0)	None

Note: The narrative reports did not contain enough information to code the time dimension.

Process-based influence has received some attention in the field of assessment. For example, Fulcher and Orem (2010) include process-based sources of influence within the "Using Results" element of their rubric to assess the quality of assessment and how it guides program improvement "in addition to evaluating the presence of results-driven improvements, the rubric also reviews whether programs address shortcomings to the assessment process itself" (p. 16). Working collectively to make meaning of student learning evidence is not a typical experience for many faculty and academic administrators and the learning curve can be steep particularly with regard to expectations. However, even small steps can help move the conversation forward toward the ultimate goal of changes that directly improve student learning.

Only 25% of the 28 cases were classified as findings-based, instrumental influences, that is, a direct action or decision about a program based on learning evidence.

Overemphasizing the importance of findings-based, instrumental influences can lead to making programmatic changes based on invalid data or no data and has the potential to be as unproductive to the improvement of student learning as making no change. Therefore, the process of inquiry by which faculty are learning how to collect data that is informative and addresses their questions of interest about student learning in the program are important steps toward a meaningful change that may be more likely to improve student learning. One might even argue that greater clarification and communication with students about learning objectives through the articulation of a rubric, for example, may have more impact on student learning than an instrumental change based on poor evidence. Prompting faculty to dialogue about, articulate, and communicate their shared intents for student learning is an example of a collaborative and constructivist approach to assessment.

Dimension 2: Effects of Influence

Approximately 40% of the cases were classified as instrumental while the remaining 60% were classified as conceptual/cognitive and affirmative effects of influence. Instrumental, conceptual/cognitive, and affirmative effects were common with findings-based sources (see Table 2) while only instrumental and conceptual/cognitive effects were apparent with process-based sources (see Table 3). This is somewhat intuitive given that learning from the process tends to lead to changes in how evidence is viewed and collected and not necessarily conclusions about students' learning if evidence collected is not viewed as credible.

Examples of instrumental effects in Table 2 and 3 are distinct not only because a change is identified but also because those changes are based on conclusions from learning evidence. The use of very detailed and specific language suggests a commitment to a planned change or improvement. Details include rationales for changes and specific aspects of the change. For example, detailed language will not just indicate that new material needs to be introduced or assignments need to be changed, but what material will be introduced or how assignments will be changed and why. Instrumental effects provide the kind of tangible response to assessment evidence is closing the loop.

Examples of conceptual/cognitive effects of influence in Table 2 highlight circumstances where evidence revealed an issue or enhanced understanding of how student learning occurs or is represented. Conceptual/cognitive effects differ from previous classifications of instrumental effects in that solutions to issues were either not identified or only suggested, and there was no firm commitment to implement that change. Language from these cases indicate assessment results caused faculty to question current practice or to think more deeply about what they were doing and why. This type of insight has great value in the long term. Although the evidence considered may not have been compelling enough to make an immediate change, it has disrupted conventional wisdom that may lead to a closer examination of the issue in the future when shared collectively through a meaning-making process. In summary, conceptual/cognitive effects reflect deeper and more informed understandings of student learning. Evidence is having an impact but does not necessarily result in observable action at least in the short term. For some disciplines, this type of effect might be more common and more highly valued because it mirrors the process of inquiry used in their disciplines.

However, intention does not guarantee that an improvement will occur, and often whether the learning improvement does occur is not determined.

Cases of affirmation effects typically are confirmatory in nature supporting previous beliefs about the effectiveness of practices and policies and learning concepts. Given the context in which this study was conducted, affirmation effects may also have a persuasive nature. Departments are trying to demonstrate that their courses are meeting the general education guidelines for learning outcomes. As is demonstrated by the example cases in Table 2, these conclusions are often briefly stated and sometimes accompanied by a rationale and only occasionally accompanied by data. Conclusions tend to be stated more strongly and often use persuasive language like "confident" or "pleased." Not referencing evidence in many of these cases does suggest that these conclusions may not be based on evidence but on affect. Therefore, future research might investigate associations between affirmation effects and viewing the purpose of assessment as accountability rather than improvement.

Dimension 3: Results of Influence

As shown in Table 4, the raters were only able to classify results of influence for 3 of the 28 cases (11%). The cases in this study, and even the full reports, did not provide enough information to reliably classify the outcome or consequence that resulted from the effect of the findings. Only improved student learning outcomes and building/strengthening communities of practices results were identified, and these were restricted to findings-based, instrumental influences. Perhaps, the results dimension is related to the time dimension identified by Kirkhart (2000). The results of even findings-based, instrumental influences may not be apparent immediately but only at the end of an assessment cycle or in the long term. It is possible that some assessment processes may never end in a result when specific types of sources or effects come into play. For example, when the assessment process is conducted to meet a mandate rather than as a process of inquiry, affirmative, or affective effects influence might occur without a result.

Table 4

Coding Results: Results of Influence Dimension (n = 28 cases)

Sources of Influence	Effects of Influence	Results of Influence	Case Examples
Findings-based	Instrumental	Improvement of student learning (n = 2)	*Student work revealed that several students were making simple unit errors at the beginning of a new topic. As a result, the instructor used this data as justification for providing more emphasis on the importance of unit for the following course topics. As a consequence, scores for the following quizzes and exams had a noticeable increase. Among the analyzed offerings a variance in the manner in which questions were presented for the purpose of problem solving was noted. In the earlier offerings, quizzes contained a series of simple problem followed by one more complex problem. In later semesters, quizzes were presented with a similar complex problem, but broken down into smaller steps of the same question. Students demonstrated much higher performance in solving this particular type of problem with the new approach.*
Findings-based	Instrumental	Building/ Strengthening communities of practice (n = 1)	*Our department is using the findings as a springboard for continued discussions about student learning. We used the momentum we gathered from our dialogue about (general education) to initiate a teaching brownbag program in the department.*

This does not suggest that the results dimension is irrelevant. As discussed earlier, not identifying effects of influence as an end result is shortsighted. For example, findings-based, instrumental changes are without meaning if follow-up evidence does not show that the intended improvement in student learning was actually achieved. Therefore, the results dimension provides another layer for differentiating types of influences that help move beyond effects dimension to measurable, meaningful changes to teaching and learning. To further explore the results dimension it may be necessary to query programs and their faculty through interviews rather than narrative reports. In-depth interviews tend to be better suited to exploring complex issues (Creswell, 2013), such as this proposed multidimensional, integrated model of influence.

Dimension 4: Time of Influence

Although the analysis for this study did not involve classifying the *time* dimension, it was clear that whether a program completed one or more assessment cycles (identify, gather, interpret, implement) had an impact on the other dimensions. For example, in two cases where improvement of student learning result of influence was identified, the programs completed multiple assessment cycles before reporting. Completing multiple cycles allowed for following up on a change to see if it led to intended improvement in student learning. As illustrated by the building/strengthening communities of practice example in Table 4, not all cases identified for the results dimension completed an assessment cycle. In one case, the program had identified, gathered, and conducted some interpretation of their data, but had not identified an instrumental change. However, the program did plan to implement a cultural change to create occasions for all faculty members in the program to discuss assessment evidence gathered. Therefore, there may be a relationship between the time and results and that interaction may be related to length of time it takes to complete an assessment cycle.

Conceptual/cognitive effects reflect deeper and more informed understandings of student learning. Evidence is having an impact but does not necessarily result in observable action at least in the short term.

Conclusions

The call for performance-based results is growing in higher education. Revisions of the Reauthorization of Higher Education Act resemble NCLB through a call to hold colleges accountable for cost, value, and quality including benchmarks for affordability and student learning outcomes as criteria for receiving financial aid (Kelderman, 2013). This may mean a shift in the focus of accreditation from improvement to external accountability. Ewell (2009) indicates that the paradigms that underlie improvement and accountability differ enough that it creates a tension: "Within the 'Improvement Paradigm,' the intent is to use the resulting information to enhance teaching and learning. Within the 'Accountability Paradigm,' in contrast, the intent is to use information to demonstrate to policy makers and the public that the enterprise they fund is effective and worth supporting" (p. 9); and to demonstrate return on investment, standardized tests and measures are used to compare institutions against standards of performance. This reliance on tests and measures, particularly by policy makers, for accountability purposes is fostered in part by the idea that the only legitimate purposes a test or measure can serve is documenting learning and directly informing decisions about success or failure. As argued in this paper, however, assessment is also responsible for transforming thinking of stakeholders about teaching and learning through the use and discussion of learning evidence. This rising tension is a reason why higher education needs to clarify the intended purposes for the assessment data they collect and clearly understand the multidimensional, integrated ways that assessment evidence can produce program improvement.

The formulation and application of the proposed model of influence in this article demonstrates an important shift in conceptualization of the aims of assessment and terminology from use to influence that was advocated by Kirkhart (2010). Considering influence of the assessment effort in terms of sources, effects, results, and time provides for a fuller appreciation of the aims and effects of the assessment effort and recognizes that organizational learning is facilitated by constructivist and transformational approaches. In particular, the differentiation of the two dimensions of effects of influence and results of influence enables a consideration of the mechanism of effect as well as the consequences, which are critical in closing the loop of assessment efforts. The dimension of time of influence further extends the focus from the immediate and end of cycle periods to the influence that occurs beyond the end point of an assessment cycle. In practice, institutions and the accreditation process may overlook

It is possible that some assessment processes may never end in a result when specific types of sources or effects come into play.

meaningful types of influence by focusing only on findings-based, instrumental influences that occur at the end of the cycle because those influences are more visible. However, it is clear from the applications of a multidimensional, integrated model of influence to examples of assessment activities that several types of influences exist and should be considered when evaluating the contribution of assessment to program improvement. A quarter of the influences identified in this study were from process-based sources of influence and more than half of the effects of influences identified were conceptual/cognitive or affirmative. These influences may have been overlooked as use by both accountability and improvement models. This potential oversight supports the premise of this study that conceptions of assessment use may suffer from construct underrepresentation.

This rising tension is a reason why higher education needs to clarify the intended purposes for the assessment data they collect and clearly understand the multidimensional, integrated ways that assessment evidence can produce program improvement.

These results support the findings of Peterson and Einarson (2001), Blaich and Wise (2011), and others in that while only a small percentage of colleges and universities used assessment information or reported that information impacted improvements in student learning, assessment information had prompted some campus discussion about undergraduate education and teaching and learning (Banta & Blaich, 2011; Ewell, 2009; Kuh & Ikenberry, 2009; Kuh et al., 2014). This study supports the premise that the influence of assessment evidence in prompting of conversations should be valued because it can facilitate meaning making. Furthermore, the heuristic value of the proposed model of influence is also evident in that it facilitates the implementation of Peterson and Einarson's (2001) recommendation that institutions do a better job of monitoring, examining, and evaluating the multiple, integrated ways assessment evidence impacts teaching and learning. Blaich and Wise also suggest that assessment requires people with different experiences and backgrounds to work together and requires communication as a precursor to action. The importance of communication and community as precursors to action encourages the use of a heuristic that recognizes when these steps are occurring and when they are not. Without communication and community the probability of action is greatly diminished. When it does occur and leads to conceptual/cognitive effects or the building/strengthening of community, the possibility of observable, meaningful improvements to teaching and learning are greatly improved. This proposed model of influence values these recommendations from Blaich and Wise and provides one representation of a constructivist approach to how organizations learn from evidence for improvement as proposed by Preskill and Torres (2000).

It is hoped the results of this study will prompt further research and application of the proposed model to assessment practice. Future research might explore how well the model represents the multidimensional, integrated influence of assessment evidence in other contexts. A broader array of quantitative and qualitative research studies could explore further the heuristic value of this model, specifically in terms of the effects, results, and time of influence. In terms of assessment practice, the model could provide more comprehensive documentation and improved representation of assessment's influence for accreditation self-studies, formal program reviews, and self-monitoring of the contributions and impacts of the assessment process. A more comprehensive and valid representation of assessment's influence has the potential to encourage the academy and external stakeholders to view assessment as an enterprise with multiple positive and desirable outcomes that contribute to a value shared by all, more effective educational programs and improved student learning.

References

Banta, T. W. (2002). Characteristics of effective outcomes assessment. In T.W. Banta & Associates (Eds.), *Building a scholarship of assessment*. San Francisco, CA: John Wiley & Sons.

Banta, T. W. (2010). Impact of addressing accountability demands in the United States. *Quality in Higher Education, 16*, 181-183.

Banta, T. W., & Blaich, C. F. (2011). Closing the assessment loop. *Change*: *The Magazine of Higher Learning, 43*(1), 22-27.

Banta, T. W., Jones, E. A., & Black, K. E. (2009). *Designing effective assessment: Principles and profiles of good practice*. San Francisco, CA: John Wiley & Sons.

Banta, T. W., & Pike, B. R. (2012). *Making the case against—One more time* (NILOA Occasional Paper No. 15). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Baxter Magolda, M. B., & King, P. M. (2007). Interview strategies for assessing self-authorship: Constructing conversations to assess meaning making. *Journal of College Student Development, 48*, 491-508.

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

Bresciani, M. J., & Wolff, R. A. (2006). *Outcomes-based academic and co-curricular program review*. Sterling, VA: Stylus.

Cox, M. D. (2004). Introduction to faculty learning communities. *New Directions for Teaching and Learning*, *2004*, 5–23. doi: 10.1002/tl.129

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.

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

Ewell, P. T. (1989). Information for decision: What's the use? *New Directions for Institutional Research, 1989*(64), 7-19. doi: 10.1002/ir.37019896403

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

Ewell, P. (2010). Twenty years of quality assurance in higher education: What's happened and what's different? *Quality in Higher Education, 16*, 173-175.

Fulcher, K. H., & Orem, C. D. (2010). Evolving from quantity to quality: A new yardstick for assessment. *Research and Practice in Assessment, 5*, 1-10.

Greene, J. C. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review, 12*, 91-116.

Gray, P. J. (2002). The roots of assessment: Tensions, solutions, and research directions. In. T.W. Banta & Associates (Eds.), *Building a scholarship of assessment*. San Francisco, CA: John Wiley & Sons.

Herndon, C. (2006). Peer review and organizational learning: Improving the assessment of student learning. *Research and Practice in Assessment, 1*(1), 1-7.

Jonson, J. L., & Thompson, R. J., Jr. (2013). Assessment approaches and perspectives. In R. J. Thompson (Ed.), *Changing the conversation about higher education*. Lanham, MD: Rowman & Littlefield.

Kelderman, E. (2013, February 13). Obama's accreditation proposals surprise higher-education leaders. *The Chronicle of Higher Education*. Retrieved from http://www.chronicle.com.

Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation, 88*(1), 5–24.

Kuh, G. D., & Ewell, P. T. (2010). The state of learning outcomes assessment in the United States. *Higher Education Management and Policy, 22*(1), 1-20.

Kuh, G., & Ikenberry, S. (2009). *More than you think, less than we need: Learning outcomes assessment in American Higher Education*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Kuh, G. D., Jankowski, N., Ikenberry, S. O., & Kinzie, J. (2014). *Knowing what students know and can do: The current state of student learning outcomes assessment in US colleges and universities*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Leviton, L. C., & Hughes, E. F. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review, 5*, 525-548.

Maki, P. (2010). *Assessing for learning: Building a sustainable commitment across the institution* (2nd ed.). Sterling, VA: Stylus Publishing.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.

Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco, CA: Jossey-Bass.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. (3rd ed.). Thousand Oaks, CA: Sage.

Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Thousand Oaks, CA: Sage.

Peterson, M. W., & Augustine, C. H. (2000). External and internal influences on institutional approaches to student assessment: Accountability or improvement? *Research in Higher Education, 41*, 443-479.

Peterson, M. W., & Einarson, M. K. (2001). What are colleges doing about student assessment? Does it make a difference? *Journal of Higher Education, 72*, 629-669.

Preskill, H., & Torres, R. T. (2000). The learning dimension of evaluation use. *New Directions for Evaluation, 88*(1), 25-37.

Rich, R. F. (1977). Use of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding. In C. H. Weiss (Ed.), *Using social research in public policy making* (pp. 199-233). Lexington, MA: Heath.

Rossman, G. B., & Rallis, S. E. (2000). Critical inquiry and use as action. *New Directions for Evaluation, 88*(1), 55–69.

Spencer Foundation. (2010). Strategic initiatives: Data use and educational improvement. Retrieved from http://www.spencer.org/content.cfm/data-use-and-educational-improvement

Stemler, S. (2001). An overview of content analysis. Practical Assessment, *Research & Evaluation, 7*(17). Retrieved from http://PAREonline.net/getvn.asp?v=7&n=17

St. John, E. P. (2009). *College organization and professional development: Integrating moral reasoning and reflective practice*. New York, NY: Routledge.

Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation, 19*, 21-33.

Weiss, C. H., & Bucuvalas, M. J. (1980). *Social science research and decision making*. New York, NY: Columbia University Press.

Wegner, E. (1988). *Communities of practice: Learning, meaning, and identity*. Cambridge, England: Cambridge University Press.