Evidence-Based Storytelling in Assessment

Natasha Jankowski
Contents

Abstract . . . .3
Evidence-Based Storytelling in Assessment . . . .4
Use of Assessment Results . . . .4
Causal Claims. . . . .4
Alternative Conceptions. . . . .6
The Importance of Being Carol Weiss. . . . .7
Assessment and Stories. . . . .8
Evidence-Based Storytelling. . . . .9
Warrants and Arguments. . . . .10
Implications for Practice. . . . .11
Conclusion . . . .11
References . . . .13
About the Author . . . .15
About NILOA . . . .16

Please Cite As:

Abstract

In the field of education, we bump constantly into buzzwords, pass unclear phrases in the hall, and greet acronyms as good friends. Yet rarely is it that we unpack or critically examine what is meant when phrases are uttered. What does it actually mean to “enhance institutional effectiveness”? To “foster synergy”? To “improve student learning”? In the field of assessment, there is a stated desire for “more use of assessment results,” to “close the loop.” There is agreement that the purpose and intention for engaging in assessing student learning is to ultimately “improve student learning” but there is no clear framework for what “improving student learning” entails or what it means to “close the loop.” However, there does seem to be agreement that whatever “use” might mean, there is not enough of it happening let alone with regularity (Kuh et al., 2015). This paper provides an overview of an alternative conception of use through the lens of evidence-based storytelling—an approach that has been used at the National Institute for Learning Outcomes Assessment (NILOA) to refine and encourage evidence-based stories in assessment (Jankowski & Baker, 2019). This occasional paper serves two main purposes: to re-examine what is meant by use of assessment results and to unpack evidence-based storytelling and its connection to assessment.
Evidence-Based Storytelling in Assessment

Natasha Jankowski

In the field of education, we bump constantly into buzzwords, pass unclear phrases in the hall, and greet acronyms as good friends. Yet rarely is it that we unpack or critically examine what is meant when phrases are uttered. What does it actually mean to “enhance institutional effectiveness”? To “foster synergy”? To “improve student learning”? In the field of assessment, there is a stated desire for “more use of assessment results,” to “close the loop.” There is agreement that the purpose and intention for engaging in assessing student learning is to ultimately “improve student learning” but there is no clear framework for what “improving student learning” entails or what it means to “close the loop.” However, there does seem to be agreement that whatever “use” might mean, there is not enough of it happening let alone with regularity (Kuh et al., 2015). This paper provides an overview of an alternative conception of use through the lens of evidence-based storytelling—an approach that has been used at the National Institute for Learning Outcomes Assessment (NILOA) to refine and encourage evidence-based stories in assessment (Jankowski & Baker, 2019). This occasional paper serves two main purposes: to re-examine what is meant by use of assessment results and to unpack evidence-based storytelling and its connection to assessment.

Use of Assessment Results

In some ways, what is implied or intended when assessment professionals say that evidence of student learning was used to improve student learning is that those undertaking assessment of student learning examined results, subsequently made changes, and those changes then led to increased student learning. In essence, what is inferred is a process through which data on student learning are examined and discussed internal to the institution. The process generally sounds very rational, reasonable, and straightforward but leaves the lingering question: Can one ever actually know that the changes made (informed by evidence of student learning) actually enhanced student learning? Was it the assessment results that tipped the scales or something else? What does it really mean to “use” assessment evidence to improve?

Questions on use of assessment results are not new (see Blaich & Wise, 2011) and have led others to distinguish between changes and improvements, asking if changes actually led to improvements and how one might know (see Fulcher et al., 2014 and Smith et al., 2015). Nor is the question of use new to related fields such as use within evaluation findings along with the uptake of research in policy and practice (Johnson et al., 2009; Cousins, 2003; Weiss, 1979). There is, however, a clear thread among these different discussions—a desire for clarity on what is inferred when use is invoked, coupled with a desire for certainty in the form of causal claims.

Causal Claims

One of the end goals of “use” in assessment is to make causal claims about program/experience/institutional impacts on students and their learning. As educators and arbiters of quality in higher education, we want to be able to say that because of our institution, curriculum, support structures, etc. the students who experienced it learned more—that
the learning was a direct result of something we did. For assessment, this includes not only causal connections between a change and subsequent improvement (i.e., the use of data in deciding to modify practice that leads to improvement), but also the establishment of direct causal linkages as a result of assessment data (i.e., this improvement was a result of the assessment process in place). The desire for direct causal linkage between changes and improvements is, in part, one of the reasons for the assess, intervene, reassess model of Fulcher and colleagues (2014). The desire for causal connections between the process of assessment and improved student learning is a long-standing question of faculty—does engaging in assessment actually do anything?

Yet, changes are not synonymous with improvements, as Fulcher and colleagues (2014) present in their work on learning improvement. And simply having lots of data points available does not lead to using them in decisions about altering processes or practice (Blaich & Wise, 2011). While it does make sense that all the data collected from the ongoing processes of assessing student learning should be put to some end and “used” to improve offerings, which in turn would improve student learning since students would experience improved offerings, the argument is focused on improvement in outcomes with certainty that is unlikely in our educational environment.

If instead of examining the model or design of assessment, one was to unpack the causal claim being made throughout in the form of an argument, one might end up with something along these lines:

These Students + Our Unique Institutional Processes and Practices = Enhanced Learning or Learning Promised

The desired argument is that use is direct. We did this. We analyzed data, evaluated a list of possible choice options, chose the best one, implemented it with fidelity (Gerstner & Finney, 2013), and improved learning as a result.

Students are bystanders in this narrative, having education done to them, while educators fiddle with turning the nob just enough to pump out critical thinkers. I write this in jest a bit, with the intent to dig deeper and trouble the concept of use. Further, it is not simply a focus on measurement that led to use being closely associated with desire for clear causal connections. In part, the desire for causal connections is due to the accountability climate and tension inherent in assessment (Ewell, 2009) where assessment is part of answering questions about the value and worth of higher education: Is it worth the cost to attend? Is there value-added by attendance? How did the institution make a difference?

At the end of the day what institutional leaders want to say is that their particular institutional structures plus the support services offered to students, combined with the students served equals enhanced learning. By offering this particular curriculum and these particular experiences to these students, we are accountable because learning improved. We did that. It improved because of us as an accountable educational institution.

But can we know this?

Assessment of student learning occurs in real time on campuses, not in labs with experimental designs in place where all faculty teach the same and the environment can be
controlled. It is at this point in a talk that I would gesture broadly to pandemic education as one example of the inability of education to occur in a controlled environment. But even in a regular academic year, students are mobile, swirling between institutions, stopping and starting again. They work and they have computers in their pockets in the form of phones. It is not far-fetched to ask, how much did the institutional programmatic offerings add to effective written and oral communication, versus life, time, and work? Further, institutions of higher education are loosely coupled systems (Weick, 1976). There is no guarantee that if we put student A into situation B that learning C will happen. On top of that, there are change management and implementation issues that trip up direct causal arguments. There are many small, untracked changes that happen across institutions that can add up over time to large changes. For instance, faculty are constantly making tweaks in their courses to support student learning, which from a program design perspective can mean that the role of a particular course in a program shifts out of alignment impacting program-level learning outcomes data. Or students alter their behavior and evade faculty attempts to instill learning in revised assignments—all of which can add up and impact, over time, the ability to present a causally linked argument that our institutions and offerings were the ones who caused the learning.

**Alternative Conceptions**

Instead of trying to fit complex educational efforts into a rational, linear process of direct use, what if use was thought about in a different way? What might happen if we thought about use from the perspective of meaning making and making an argument? Since we cannot randomly assign students to courses or force faculty to teach exactly the same way every time, what if what was really done when “use assessment results” unfolded was something else entirely? I argue that what we are actually doing is a cognitive process of unpacking theories of change that are shared through evidence-based stories targeted to various audiences. I argue that use occurs in our cognition and the act of making meaning of assessment results is a collaborative and social cognitive process. Storytelling and stories are simultaneously cognitive processes and products of cognition—it is how we understand and make sense of the world (Lewis, 2011). This connection between stories and use as a sense-making process presents an alternative view of use in assessment.

Take for instance the following example of student writing. If upon examining program-level assessment data, faculty find that students are not writing at the expected level on writing focused assignments, there are many possible reasons for the gap in attainment. Some faculty might argue that the issue is one of coverage—that writing was not covered enough in classes and that the solution is to add more coverage of writing throughout the program. Within that possible solution to a perceived root cause are assumptions about how students learn (i.e., in classes), so that if it was not learned the issue was dosage and needing to cover it again until students learn. But a few other faculty may push back and claim that it is widely covered throughout the program and that the real issue is students not making use of the opportunities for support available through the writing center. They argue that instead, each faculty member should encourage students to use the writing center resources in their classes and the problem will be addressed. This narrative puts the onus on the student to be comfortable going to the writing center, know that they need assistance in writing or believe it can help them, and that the timing of the center fits into their schedules and lives. The assumption is that students are simply not using the available supports, and if they did, writing would improve. Lastly, a few faculty
members argue that what students need is intentional and coherent educational pathways with clarity on why students are being asked to write and what it means for their future careers. They argue that students are not learning at the level desired because the students do not see the value in the assignments and if the value was clear, it would align with student values and then faculty would see students’ true writing potential.

I could go on with examples, but the point is that within each of these narratives, faculty are making sense of what the data mean about their program, what the data mean about the possible root cause, and through sharing stories of their sense-making on the data, unveil their beliefs on how education unfolds and the role of students and faculty in the process of learning. Through storying they are making collective meaning.

In order to explore this conception more fully, I turn to the work of Carol Weiss before moving into stories, storytelling, and the connections between stories and evidence that live in the power of the warrant. What is a warrant you ask? Worry not, there is an entire section coming later in the paper on the concept.

**The Importance of Being Carol Weiss**

In 1979, Carol Weiss wrote a paper exploring the many meanings of research utilization focused on what using research actually means and it is incredibly useful to the use conversation in assessment. She strives to clarify the concept to avoid “conceptual confusion” and clean up “glib rhetoric” by presenting seven different ways in which use is conceptualized in the uptake of research findings.

1. **Knowledge-driven model** is predicated on the idea that the sheer fact the data exist will drive it to be used. In essence, the assessment data will be so compelling that once faculty and staff are informed and know, they will act. The work of Blaich and Wise (2011) would suggest otherwise.

2. **Problem-solving model** involves using data in a decision to solve a problem. This approach assumes that there is clear consensus on goals or the desired end state, and that the evidence from assessment will reduce uncertainty by helping to point towards a clear path forward. The focus in the problem-solving model is on communication and getting the right data into the right hands of people at the right moment when they need it for decisions. Finding such perfect conditions seems unlikely.

3. **Interactive model** is one that involves seeking knowledge. It is not orderly or a lock step process of getting data at the right moment for decision-making but is instead messy. This one is more about the ebb and flow of questions between a person and uncertain data over time. In assessment, this would include the collaborative efforts of faculty, staff, and students coming together to reflect upon practice to better understand their different perspectives.

4. **Political model** is when research is used to support a predetermined position. Think of this as the person asking for data that backs up an argument they already made. For assessment professionals that intersect with institutional research, this might be the data request for specific information to appease external reporting bodies.

5. **Tactical model** actually has nothing to do with the results but is all about looks. The model is that the research, or in our case assessment process, is simply happening at
Think about an accreditation visit where someone points to all the assessment unfolding all the time. We are doing so much and there is so much data so clearly, we are doing ok!

6. **Enlightenment model** focuses on the use of evidence to change how someone thinks about or sees things. It involves asking different questions, adding more complexity, and rethinking or approaching something from a different perspective. In conversations with faculty and staff on what changes were made as a result of assessment data, a common discussion point is not about modifying the curriculum but of redefining the problem or thinking about learning in a different way.

7. **Research as the intellectual enterprise of the society** argues that we are part of an enterprise that examines our practice and improves. That engaging in research is a part of what makes society. In assessment, this might be the mantra of continuous improvement where we engage in assessment as a learning organization to advance our own understandings.

Each of the 7 models present a different story, using evidence, on what people are doing (or not doing) and why. They represent not simply the decision or action/inaction but the reasoning behind it as well. In assessment, if use was broadened to include more than direct causal impact stories, one could argue that a lot more use is happening all the time. And while Carol Weiss was instrumental in conversations broadening use of results in research and evaluation, the allure of a narrow focus upon direct use is strong. Johnson and colleagues (2009) conducted a review of research on evaluation use from 1986-2005 and found instances of use in the form of process, instrumental, conceptual, or symbolic. However, they argue that use was limited because instrumental or direct use was found the least. The continual focus on direct use at the expense of others types of use led to work on evaluation influence (Kirkhart, 2000; Alkin & Taut, 2003), and frameworks to better understand context and stakeholder involvement as it relates to possible use (Cousins, 2003) with the goal of increasing use of evaluation findings.

Across the field of evaluation and assessment, both agree that much “use” goes unnoticed or not recorded, because the focus on use has been upon the act of making a decision and how instrumental (or not) the evidence was to that decision. Jonson, Guetterman, and Thompson (2014) agree and argue that assessment “as an educative process of inquiry and learning that involves the generation and application of knowledge” (p. 22) gets lost in the fray of direct, linear, decision-making definitions of use. The process of collectively making sense of data stalls. The story disappears.

**Assessment and Stories**

Stories and assessment are not so unfamiliar, and assessment professional roles can include a focus on a narrative/translator (Jankowski & Slotnick, 2015). Makela and Rooney (2012) write that assessment “is essentially a process of telling a story about our people, programs, and services” (p. 2) that are told to many different people, in many different ways, with many different foci. They argue that the “storyline surrounding an assessment ultimately aims to include enough evidence to make well-reasoned assertions…” (p. 2). In this instance, evidence is defined as that which is used to make a persuasive argument regarding what worked well and what aspects could be improved. In a similar vein, learning improvement stories could be viewed as program intervention stories.
In higher education more broadly, Burton Clark (1972) introduced the concept of organizational saga as the narrative institutional members share of the heroic acts of an institution and its unique development over time. Shulman (2007) reminds us that accounting is a form of narrative and that accountability in higher education is the act of giving an account, arguing that “counting without narrative is meaningless and narrative without counting is suspicious” (p. 11). Further, storytelling can create a new organizational culture or collective identity (Abrahamson, 1998; Butcher, 2006; Feldman, 1990; Whyte & Ralake, 2013) and group storytelling has been used across institutions of higher education to shape national narratives of what worked and how best to address mutually shared challenges in community colleges (Bennett, Bragg, & Kirby, 2015). How do stories do this? Because stories are one of the most powerful means of education (hooks, 2010).

The power lies in the combination of evidence and stories. Storytelling complements abstract analysis, not as a replacement, but as a supplement which enables the imagining of new perspectives and helps the audience identify with the storyteller (Denning, 2001). For assessment, which is inherently local, data are created in particular educational settings that require knowledge of the context (Knight, 2006), meaning stories coupled with evidence can make or break understanding. Without stories, assessment becomes what Marchese (1988) stated was “A technology deployed out of context” (p. 25). Through telling stories assessment professionals provide a way to make sense of the educational experience and communicate that experience to others, not as literal accounts, but about the meaning made from the experience (Lawrence & Paige, 2016). If assessment only tells measurement stories and stories about data collection, the opportunity to engage in making meaning is lost (Jankowski, 2017). Stories also help build authenticity where trust has been lost (Brown et al., 2005) and can be a form of resistance (Lawrence & Paige, 2016), pushing back on dominant models of measuring success. Through connecting stories, evidence, meaning making, and explanation all the pieces are in place for evidence-based storytelling.

**Evidence-Based Storytelling**

Evidence-based storytelling involves the process of using evidence of student learning in support of claims or arguments about improvement and accountability that is told through stories to persuade a specific audience (Jankowski, 2012; Jankowski & Baker, 2019). It couples together evidence gathered from the assessment process, evidence that can be quantitative and/or qualitative. There might be different narratives for different audiences, but the evidence supporting the argument is the same. The story involves giving an account of the assessment process and how an institution knows students are learning by outlining the assumptions and argued causal linkages for why things are done in a particular way—by sharing the meaning made from the data.

These are not stories of data, but stories of processes and practices and the argued impact of those processes and practices on people. Through evidence-based storytelling, those within an institution strive to reach a shared understanding on what good assessment practices entails for them. The distinction of evidence-based storytelling is the connection of the assessment evidence to the claim made in the argument via a focus on the warrant.
Warrants and Arguments

Warrants involve making explicit existing assumptions. Causal assumptions in learning are based on underlying beliefs, but in order to make well-reasoned assertions in support of a larger argument, stories and warrants have to connect. Shadiow (2013) argues,

Within the impressions left by the stories on the storyteller are claims (assertions) about the world, as presented by the people and actions in each. In addition to telling some stories for entertainment, we may tell stories for some prescriptive purpose—to teach a lesson, to illustrate a moral in its narrative arc, or to draw an evaluative contrast. In this, we react, agreeing or disagreeing with the assertions within the stories…Assertions are part of the “understory” in most single incidents. Someone in the story, either through their words or actions, claims something, and in our telling of the story, we judge that claim. Through probing the patterns for assertions embedded within the stories and our tellings of them, we can see how they contribute to bringing current, often unacknowledged, assumptions into view (p. 86).

Warrants help with the explanation part of stories and add support to the assertions being made about student learning. Argued prior, making the case that the institution was the cause of the learning can be rather complicated. Theories offer windows into causal relationships and help explain phenomena (Kitson et al., 2008), but storytelling relates a series of events weaving causality throughout (Miller, 2011). If we think about assessment as Mislevy and Riconscente (2005) suggest, as the process of crafting an argument, then assessment is a means by which “…we arrive at inferences or judgments about learner proficiency based on a set of observations” (p. 1). Assessment practitioners want to make valid claims about students’ knowledge, skills, and abilities. In order to do that, relevant evidence needs to be presented “where criteria for relevance are determined by our warrant” (Mislevy & Riconscente, 2005, p. 2), where assessment is “at heart an exercise in evidentiary reasoning. From a handful of things that students say, do, or make, we want to draw inferences about what they know, can do, or have accomplished more broadly.” (Mislevy & Riconscente, 2005, p. iv).

The Toulmin (2003) argumentation model is insightful here to understand the relationship between evidence and story through the warrant. The model entails the following elements:

1. The Claim, or the conclusion of the argument or statement we want the audience to believe. This entails being accountable or improving student learning in the case of assessment.
2. The Ground, or the foundation or basis for the claim. This is the support, or the evidence gathered from assessment and from educational practice.
3. The Warrant is the reasoning that supports the inferential leap from the grounds or evidence to the claim or conclusion. It is how we get from A to B. This is the story of learning, the institution, the students, and all the different pieces that play a part in the movement between the conclusion and the evidence. This is also where the cognitive process of making meaning in use of evidence lives. It is the story we tell of causality and the argument being made.
There are a few other elements of the model, but for our purposes, these three suffice. What is helpful in the Toulmin (2003) model is that the claim and grounds are linked. One needs to know the grounds for the claim or the evidence the claim is built upon. But the warrant is the why.

The warrant presents the narrative about why and how the claim and evidence connect together. It explains why one thinks that the evidence makes the claim true. In assessment reporting, the warrant is largely forgotten and often jumped over (Jankowski & Baker, 2019). Instead, data are presented and we say that learning happened. What is missing is the exploration not only that it happened, but why it happened the way it did (Rooney & Heuvel, 2004). The warrant helps to moves beyond surface-level problem identification and examine assumptions in order to prevent reoccurrences (Taitz et al., 2010). It helps make meaning of the data and engage in cognitive use through storytelling because it forces education to explain itself.

Taking Toulmin’s (2003) model and focusing on the warrant, the evidence-based story we tell embeds the why. This means the process of use involves selecting a variety of evidence in relation to a claim made about improvement of student learning. The warrant involves outlining and justifying why we think this change, for these students, at this institution, at this time, led to the improvement we are seeing in student learning (Or why we think it will). Without the warrant we cannot tell a story that argues for possible causal pathways.

Implications for Practice

To reinforce the cognitive process of meaning making within use, space and time are needed for dialogue between colleagues. In assessment there is so much doing that there is limited time, if any, built into reflecting upon the data and deciding what it all says, what argument might be made, and what story it tells about students and their learning.

To help with story development and to engage in a “assumption hunting” (Shadiow, 2013), NILOA refined through various workshops and events a toolkit for crafting evidence-based stories (Jankowski & Baker, 2019). One element that was very clear was the importance of audience. Consider the intended audience. In most assessment related reporting there is no clearly defined audience and limited contextual information presented. Further, some audiences do not find certain evidence sources or arguments compelling, privileging some evidence over others. Knowing the audience and using the story to educate on the value of alternative evidence sources while also building shared understanding may prove useful to institutions.

In addition to audience, context is required for meaning making and storytelling. The insight offered through examining local practice in connection to larger contexts may prove useful to determine, what, if anything, may be done to support students in their learning.

Conclusion

Stories are powerful. As we repeat stories to ourselves, we begin to believe them more. Think of the stories of resistant faculty, administration that does not understand, students who do not care. The stories we tell define our institutions. And while a lot of data is
gathered through assessment, meaning needs to be made of the data. To tell audiences what data on student learning means, we must make sense of it ourselves. Higher education has struggled with telling compelling narratives about the impact and value (Kendig, 2015; Gaston, 2014). The response can be one that involves evidence and story, because evidence gives stories substance, but stories give evidence meaning.

Focusing on narrative through evidence-based storytelling is helpful to internal and external communication on what good assessment practice looks like and entails within a specific context and setting. It is also a means to support community and shared understanding in how faculty and staff work to advance student learning together (Simmons, 2006). It can offer a means of healing. Some stories simply need to be told. Some of our stories are tragedies, some of our actions and inactions in higher education involve budget fantasies, and some are tales of heroics and adventures. Practitioners in higher education need space to story together to make sense of their past, sit with the present, and imagine the future.

Some stories simply need to be told. Some of our stories are tragedies, some of our actions and inactions in higher education involve budget fantasies, and some are tales of heroics and adventures.
References


Dr. Natasha Jankowski currently serves as the Director of Assessment Success for AEFIS. Previously she served as the Executive Director of the National Institute for Learning Outcomes Assessment (NILOA) and research associate professor with the University of Illinois Urbana-Champaign. She is co-author of the book *Using Evidence of Student Learning to Improve Higher Education*, as well as co-author of *Degrees that Matter: Moving Higher Education to a Learning Systems Paradigm*, and co-editor of *Student-Focused Learning and Assessment: Involving Students in the Learning Process in Higher Education*. A forthcoming book focuses on equity and assessment. Her areas of focus include assignment design, transparency, evidence-based storytelling, equity, mapping and alignment of learning, and all things assessment. Natasha is the recipient of the Kent State University’s Alumni Award, the Young Alumni Award from the College of Education at UIUC, and the Academic Professional Excellence Award from the University of Illinois. NILOA was the proud recipient of the ACPA Contribution to Higher Education award under her leadership. She holds a Ph.D. in Higher Education from the University of Illinois, an M.A. in Higher Education Administration from Kent State University, and a B.A. in philosophy from Illinois State University.
About NILOA

- The National Institute for Learning Outcomes Assessment (NILOA) was established in December 2008.
- NILOA is co-located at the University of Illinois and Indiana University.
- The NILOA website contains free assessment resources and can be found at http://www.learningoutcomesassessment.org.
- NILOA supports institutions in designing learning experiences and assessment approaches that strengthen the experience of diverse learners within a variety of institutional contexts.
- NILOA works in partnership with a broad range of organizations and provides technical assistance and research support to various projects focused on learning throughout the U.S. and internationally.
- NILOA’s Vision is to broaden the dialogue and conversation on meaningful and sustainable assessment practices that address issues of design and implementation, and position institutions, organizations, and individuals to achieve their goals.
For more information, please contact:

National Institute for Learning Outcomes Assessment (NILOA)
University of Illinois at Urbana-Champaign
51 Gerty Drive
Suite 196, CRC, MC-672
Champaign, IL 61820

learningoutcomesassessment.org
niloa@education.illinois.edu