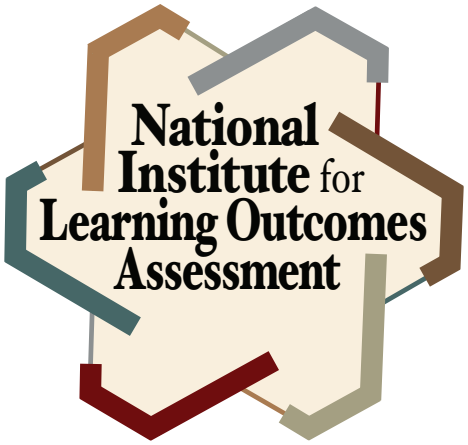
March 2021

Nº 51



The Need for Program Theory and Implementation Fidelity in Assessment Practice and Standards

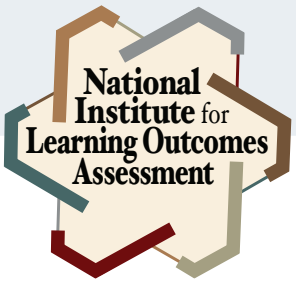
Sara J. Finney, Jennifer B. Wells, & Gavin W. Henning

www.learningoutcomesassessment.org

Contents

NILOA Mission

The National Institute for Learning Outcomes Assessment (NILOA), established in 2008, is a research and resource-development organization dedicated to documenting, advocating, and facilitating the systematic use of learning outcomes assessment to improve student learning.



Please Cite As:

Finney, S. J., Wells, J. B., & Henning, G. W. (2021, March). *The need for program theory and implementation fidelity in assessment practice and standards* (Occasional Paper No. 52). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Abstract

On their own, student learning and development outcomes assessment data have limited utility for improving programming. We believe outcomes data should not be collected until two fundamental questions can be answered: "Why should this programming result in the desired outcome?" (i.e., program theory) and "Was the intended programming actually experienced by students?" (i.e., implementation fidelity). Some assessment professionals may find this proclamation radical. Our call is fueled by the creation of unjustified programming and curriculum, coupled with the collection of outcomes data that are not used for improvement efforts. We contend that it is only after program theory is articulated that faculty and student affairs professionals can collect relevant, useful outcomes data. Moreover, valid inferences from outcomes data are contingent on knowing what programming students experienced. This "expanded" assessment practice has potential to afford better-designed, more impactful, research-informed programming to students. As our students have opportunities to engage in well-implemented, should-be-effective programming, their learning should demonstrably improve. Thus, we call for professional standards and professionals themselves to integrate program theory and implementation fidelity into outcomes assessment practice.

The Need for Program Theory and Implementation Fidelity

Sara J. Finney, Jennifer B. Wells, & Gavin W. Henning

Typical Outcomes Assessment Process

Faculty and student affairs professionals seek to implement programming (e.g., strategies, pedagogies, activities, curriculum) that results in institutions being effectively designed learning environments through which students achieve stated learning goals. Professionals are also expected to assess the intentionally selected program offerings for effectiveness (Finney & Horst, 2019a, 2019b; U.S. Department of Education, 2006). If effectiveness is not achieved, outcomes assessment results are expected to be used to guide programming changes that will result in learning improvement. However, few institutions have demonstrated learning improvement (Banta & Blaich, 2011; Jankowski, et al., 2018). In turn, assessment practitioners have considered strategies to address this predicament and increase learning improvement (e.g., Fulcher & Prendergast, 2019; Fulcher, et al., 2017; Smith, et al., 2018).

The outcomes-related question "Does the programming work?" muffles the equally important questions of "Why should this programming work?" and "What programming was implemented?"

We believe expanding the traditional assessment process will increase the likelihood of student learning and development on campuses (see Figure 1). We echo the need for additional evidence: "Assessment that is truly focused on improving students' educational experiences means putting a premium on evidence. It also means being smart about what constitutes evidence and how to use it effectively" (Hutchings, et al., 2015, p. 3). We call for the thoughtful combination of three types of evidence. Specifically, faculty and student affairs professionals should *articulate program theory using existing evidence* (Bickman, 1987; Pope, et al., 2019) and *collect implementation fidelity evidence* (O'Donnell, 2008; Smith, et al., 2019) to *effectively use outcomes evidence* to identify what programming requires adjustments to achieve student learning outcomes and to efficiently use diminishing resources.

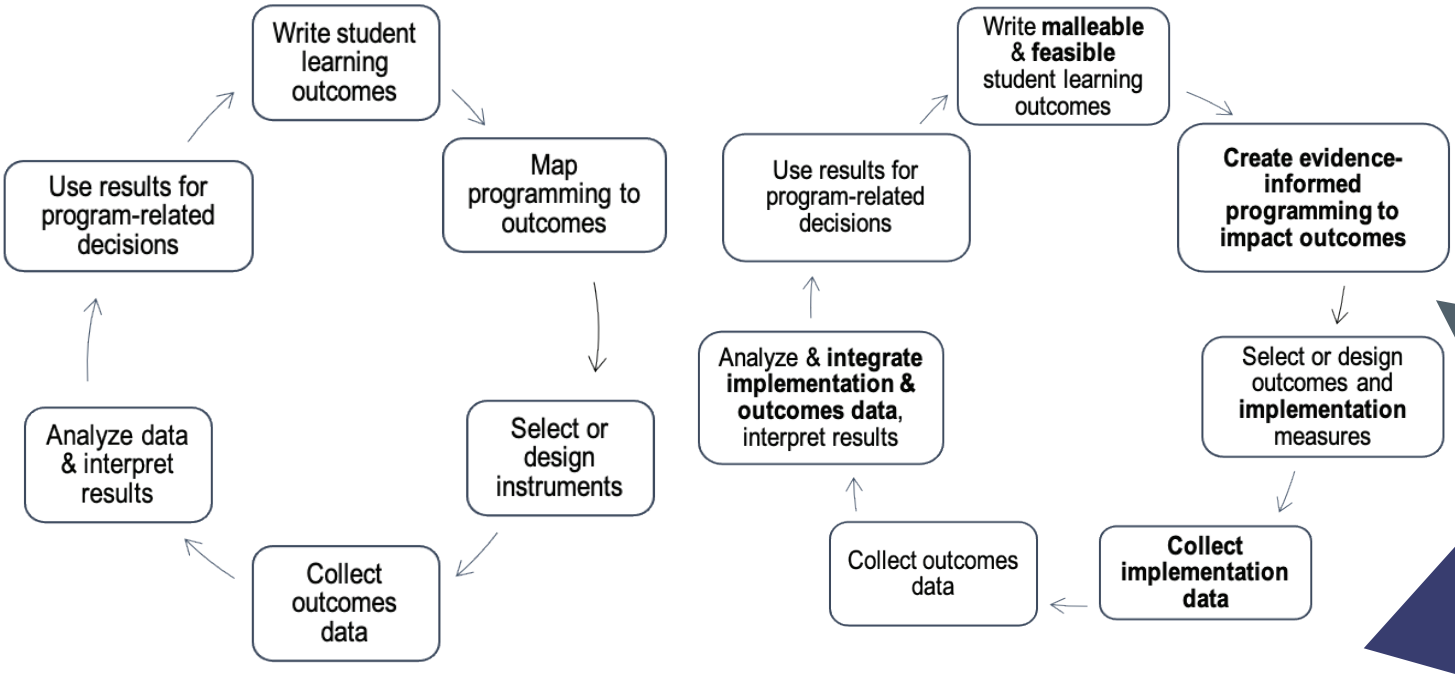


Figure 1. Typical (left) versus expanded (right) outcomes assessment process.

Unfortunately, discussions regarding the importance and process of gathering evidence to inform program theory and implementation fidelity is often stifled by the emphasis on gathering outcomes assessment evidence. That is, the outcomes-related question "Does the programming work?" muffles the equally important questions of "Why should this programming work?" and "What programming was implemented?" Answering why programming should be effective and what programming students received are prerequisites to successful use of outcomes assessment results for learning improvement. Thus, we situate program theory and implementation fidelity in the outcomes assessment process and recommend that professional standards and practice be updated to give each equal footing with outcomes data.

Articulating Program Theory to Answer "Why Should the Programming Work?"

Stakeholders (e.g., parents, students) are entitled to know if educational programming was intentionally created to achieve desired outcomes. Clear intentions are particularly important for vaguely described student affairs and co-curricular programs.

> While they [students, faculty, parents, politicians] understand that students do change and grow emotionally and socially during college, they do not attribute the change to anything other than natural maturation and some vague notion about the college experience. The idea that students might be learning outside of class is frequently regarded with skepticism and is even a bit unsettling—who is directing this surreptitious learning and what are their goals? (Carpenter, 2012, p., vii)

Program theory allows stakeholders to understand *what* programming is implemented and *why*, making obvious the links between programming and intended outcomes. By making the rationale of programming explicit, it can be interrogated, assessed, and improved.

Program theory is defined as "the construction of a plausible and sensible model of how a program is supposed to work" (Bickman, 1987, p. 5). Furthermore, it "clarifies the set of cause-and-effect relationships" believed to connect the things students do (i.e., programming) to the outcomes they are expected to achieve (Bickman, 1987, p. 5). Faculty and student affairs professionals need to move beyond the simple input-output approach (i.e., no program theory, top of Figure 2) and instead explicitly state *how* they expect programming to work, thereby making their implicit assumptions explicit (i.e., strong program theory, bottom of Figure 2).

It is helpful to distinguish between weak and strong program theory. Weak program theory is often based on limited personal experiences, assumptions, or hunches (middle of Figure 2). Strong program theory is research- or theory-based, providing evidence-based links between program activities and student learning outcomes (Pope et al., 2019). Strong program theory is not simply a mapping of programming components to outcomes (Jankowski & Baker, 2020) or a logic model (Finley, 2019), although these may be initial steps. Instead, strong program theory uses research or theory to justify each arrow linking programming and outcomes (Baldwin, et al., 2004). In turn, strong program theory communicates to stakeholders that faculty and staff intentionally built programming that should be effective given existing evidence, and they can explain why and how.

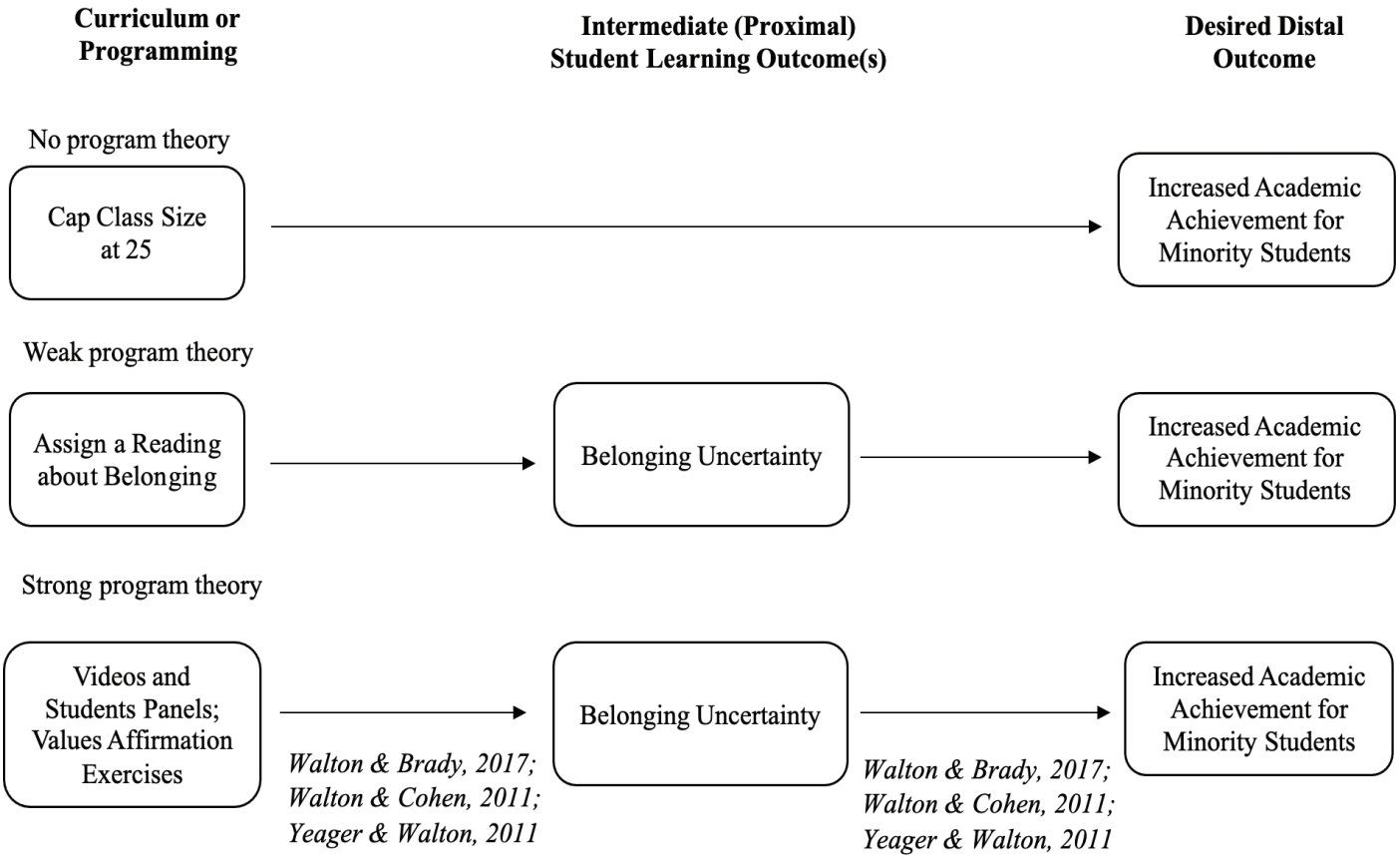


Figure 2. Logic models depicting the difference between a program with no program theory, weak program theory, and strong program theory.

Strong program theory also allows faculty and student affairs professionals to use the outcomes assessment process in a confirmatory way to test the hypothesis of program effectiveness. This evidence-informed approach can be contrasted with an unsystematic program development approach of cobbling together materials and activities. Program theory counters a happenstance or accidental approach to learning. In turn, the evidence-informed approach is more cost-effective with regard to time and resources because the programming generated is more likely to improve student learning and development than programming based on tradition, hunches, or guesses (Reinholz & Andrews, 2020). Thus, fewer iterations of the assessment process are necessary to inform changes to programming to obtain the desired outcomes.

The evidence-informed approach to program development also speaks to the responsibility of faculty and student affairs professionals. Carpenter (2001) noted this issue decades ago, when he stated that without engaging in the literature, higher education practice can become "simply random activity, bound by tradition and convention, maybe helpful, maybe not, probably suiting some students, almost certainly leaving others out" (p. 305). Understanding the research regarding the effectiveness of strategies, activities, and experiences across different student populations allows equity to be intentionally built into programming. In fact, Carpenter noted the need to keep pace with the best available

Step	Question(s) to Ask	Example
1. Articulate the Distal Outcome	• Is the distal outcome malleable? • Is the distal outcome feasible?	• The distal outcome was to increase the academic achievement of minority students.
2. Articulate Theory- or Research-Based Intermediate (Proximal) Outcomes	• What are the causes (i.e., etiology) of the distal outcome based on current research? • What do students need to know, think, or do to achieve the distal outcome?	• Reducing belonging uncertainty (concern of whether one belongs) positively affects minority students' academic achievement (Walton & Cohen, 2011). • The level of security students feel in their sense of belonging is correlated with future academic success (Yeager & Walton, 2011).
3. Develop Theory- or Research-Based Programming to Impact Intermediate Outcomes	• What programming affects the intermediate outcomes based on current theory and research? • What curricular or pedagogical strategies do research or theory suggest may be effective to influence the intermediate outcomes?	• 1-hour intervention to lessen belonging uncertainty by framing social adversity as common and transient. To internalize this framing of adversity, students created written and oral messages (e.g., student panels, student videos) for prospective students. Beyond facilitating internalization, this procedure averted the potential stigma of receiving a program/intervention, because students perceived themselves as benefactors, not beneficiaries (Walton & Brady, 2017; Walton & Cohen, 2011) • Values affirmation exercise (e.g., activities where students affirm something of importance to them) to expand sense of self, promote positive relationships, reveal commonalities with peers, and ultimately improve belonging in a setting that may feel threatening (Walton & Brady, 2017). • Key aspect of programming: brief, social-psychological exercises that focused on students' beliefs without students being told they were receiving a program/intervention (Yeager & Walton, 2011).
4. Collect Implementation Fidelity Data	• To what extent did delivered programming differ from intended programming?	• Because of the short and covert nature of the programming, auditors (e.g., colleagues), not implementers or students, rated programming for adherence, duration, and quality.

Table 1. Steps to Articulate Program Theory and Evaluate Fidelity of Implementation.

evidence in a domain: "Any student affairs professional not reading the literature, not becoming knowledgeable of research and theory, is not acting ethically. Students have a right to expect that student affairs professionals are knowledgeable of appropriate theories, current research, and proven best practices" (p. 311). This expectation also applies to faculty. According to Suskie (2018), "an effective curriculum uses research-informed strategies to help students learn and succeed" (p. 69). Articulating program theory makes explicit the responsible use of current best available evidence to offer should-be-effective strategies. When interrogating the existing research, faculty and student affairs professionals may conclude that the best available evidence is weak or not credible (e.g., limited or no research, excluded student populations, weak research design). In these cases, we recommend using theory to articulate how the program should work, acknowledging that alternative theories in the domain will afford the creation of several alternative models of program theory.

We advocate for a three-step process when articulating strong program theory: 1) identify a feasible and malleable distal outcome; 2) specify theory- or research-based intermediate outcomes; and 3) create intentional, theory- or research-based programming (Pope et al., 2019). In Table 1, we present questions faculty and staff should answer when working through this process. The resulting logic model visually represents "how" programming should result in the distal outcome (bottom of Figure 2). Thus, stakeholders have a concise answer to the question "Why should this programming work?" We walk through the steps, providing an example of articulating program theory for the distal outcome of increasing academic achievement for minority students.

Articulating program theory begins by stating an ultimate, distal outcome hoped to be achieved via programming (e.g., ethical reasoning, civic engagement, healthy drinking behavior, written communication, leadership skills, intercultural competence). This distal outcome may emanate from professional standards, student need, or faculty/staff consensus. For programming effort to be well-spent, the distal outcome must be malleable and feasible. If the outcome is not malleable but rather has trait-like stability, then attempts to develop programming to influence this outcome will be futile. Thus, the malleability of the distal outcome must be researched before further steps in articulating program theory are pursued. Given a malleable distal outcome, the feasibility of influencing the outcome given time, resources, and other practical constraints is determined. For example, it may be possible to decrease academic entitlement in college students (i.e., outcome is malleable). However, if research suggests changes in entitlement require resource-intensive programming that spans several years, it may be infeasible to target this outcome if programming must be limited to short or one-off experiences. In Table 1 and Figure 2, the distal outcome is increased academic achievement for minority students. Research indicates that academic achievement is malleable and sensitive to intentional programming (Ambrose et al., 2010; Walton & Cohen, 2011; Yeager & Walton, 2011). It is a feasible outcome given higher education's commitment to equity and student success.

Research and theory are not regularly used to articulate program theory. We call for more published examples of articulating strong program theory to supplement what currently exists.

Once a malleable and feasible distal outcome is stated, the underlying causes or antecedent conditions of the distal outcome are articulated (Renger & Hurley, 2006). This step requires consulting empirical research and relevant theory to understand the etiology of the distal outcome. Research and theory inform what specific knowledge, skills, behaviors, and/or attitudes (i.e., intermediate outcomes) should be cultivated through programming to achieve the distal outcome (Renger & Titcomb, 2002). These more proximal, intermediate outcomes are assessed to inform program effectiveness inferences and program revisions (Renger & Hurley, 2006). Some distal outcomes cannot be easily assessed because they manifest at different times, sometimes long after programming (e.g., after students leave the institution), or they require a real-life context to demonstrate the outcome (e.g., bystander intervention, voting, engagement in anti-racist acts, contraction of STD, ethical behavior, healthy drinking). However, the research-based links between intermediate and distal outcomes allow for a defensible argument that observed change in the intermediate outcome(s) should lead to change in the (unmeasured) distal outcome (Renger & Hurley, 2006). For the distal outcome of academic achievement of minority students, research indicates that belonging uncertainty undermines minorities' performance (e.g., Walton & Brady, 2017; Walton & Cohen, 2011; Yeager & Walton, 2011).

Given the articulation of the intermediate outcomes, programming is developed to influence these outcomes. Programming encompasses content (e.g., materials, activities) and delivery (e.g., pedagogy). Just as research guided the specification of the intermediate outcomes that influence the distal outcome, research guides how to achieve intermediate outcomes through educationally purposeful programming. To build research-informed programming, it is helpful to search for intervention studies that evaluate the effectiveness of programming and if effectiveness is differential for different student populations (Brousselle & Champagne, 2011). Repositories that house systemic reviews of effectiveness studies (e.g., What Works Clearinghouse, Campbell Collaboration) can be incredibly useful. Articles that synthesize empirical research on how students learn (e.g., Halpern & Hakel, 2003) or how attitudes and behaviors are changed (e.g., Yeager & Walton, 2011) can also inform program content and delivery. Additionally, books on evidence-based

pedagogical techniques exist (e.g., Ambrose, et al., 2010). If the best available research evidence is weak or not credible (e.g., limited research, excluded student populations, weak design), we recommend using theory to design programming that should work. For our intermediate outcome of belonging uncertainty, short activities involving students sharing that social adversity is common and temporary (Walton & Brady, 2017; Walton & Cohen, 2011) and engaging in values affirmation (i.e., affirm something of deep importance) in a setting where they may feel threatened (Yeager & Walton, 2011) have been shown to decrease concerns about belonging (and ultimately increase academic achievement for minority students).

Research and theory are not regularly used to articulate program theory (e.g., Brousselle & Champagne, 2011). We call for more published examples of articulating strong program theory to supplement the following that exist in higher education: university-wide programming to increase ethical reasoning (Smith & Finney, 2020), college access programming to increase admittance to and graduation from college (Millett, et al., 2018), emotional intelligence training to decrease stress in pre-service teachers (Vesely-Maillefer, 2015), semester-long programming to increase student success and retention (Pope, Finney, & Crewe, in press), trauma-informed resilience activities to help students manage stress (Oehme, et al., 2020), and education reform efforts in undergraduate STEM programs (Reinholz & Andrews, 2020).

Assessing Implementation Fidelity to Answer "Was Programming Executed as Planned?"

Specification of strong program theory and creation of research-informed programming is not sufficient to achieve the intermediate and distal outcomes. To benefit from the should-be-effective programming, students must *experience* the programming as intended or designed. The influence of effective programming on student outcomes is moderated by implementation fidelity. If programming is not implemented as planned, we should not expect evidence of program effectiveness (Fisher et al., 2014).

Implementation fidelity data determine the extent to which programming *as designed* differs from programming *as delivered* (Gerstner & Finney, 2013; O'Donnell, 2008). Deviations from planned programming include eliminating critical curriculum, shortening activities or sessions, changing mode of delivery, or adding extraneous information. Unfortunately, faculty and staff may not know if planned programming was implemented with high fidelity, especially when activities are implemented by multiple individuals in multiple settings. There may be little record of what was implemented. Thus, advertised program descriptions may be inaccurate.

Moreover, lack of implementation fidelity evidence negatively impacts the quality of inferences about program effectiveness, which significantly hinders improvement efforts (Finney & Smith, 2016; Mathers, et al., 2018). If faculty and staff do not know what programming was implemented, then they do not know what programming was assessed. In absence of fidelity data, faculty and staff are assessing the effectiveness of an unknown program ("black box"). If implemented programming is a "black box", outcomes data cannot be linked to programming to inform effectiveness claims (Gerstner & Finney, 2013). Thus, faculty and staff cannot modify programming for improvement if they are not aware of what programming was experienced.

Additionally, it is extremely difficult to engage in a curricular approach to programming (i.e., intentional continuity and sequencing of learning experiences to achieve desired outcomes) in academic (e.g., Lattuca & Stark, 2009) or student affairs programs (e.g., Kerr et al., 2017) if accurate descriptions of implemented programming are not available. Faculty and staff are asked to trust that colleagues assigned to prerequisite experiences offer programming necessary to build prerequisite knowledge and skills. If faculty and staff gather implementation fidelity evidence, it can guard against potential disconnect between sequenced learning experiences. Students then benefit from the intended coherent programming experience.

After specifying the theory- or research-based programming aligned with the intermediate outcomes (Table 1 Step 3), faculty and staff can create a fidelity checklist to evaluate if programming was implemented as planned (Swain et al. 2013). The checklist can capture data on four aspects of implementation fidelity: 1) whether each programming feature was delivered, 2) the quality with which each feature was delivered, 3) the exposure of all students to the full "dose" of programming, and 4) student engagement during programming features (Table 2). Creating the checklist is another opportunity to clarify programming.

Component	Definition	Evaluation
Program Differentiation	The theory- or research-based programming features (e.g., activities, materials, demonstrations, assignments) that should impact the intermediate outcome according to program theory.	Not evaluated via implementation fidelity checklist, but rather the programming features are listed on the checklist so the other components can be evaluated for each program feature.
Adherence	Whether or not the programming feature was implemented (i.e., opportunity to learn).	Indicate "yes" or "no" for each programming feature.
Quality	How well the programming feature was implemented.	Rating (e.g., 1 = Low, 5 = High) of the quality of implementation, coupled with some descriptive text.
Exposure	Extent to which all students received the complete programming feature.	Proportion of students who received the programming feature and duration of the programming feature.
Responsiveness	Receptivity of students to programming.	Students or implementers rate level of student engagement (e.g., 1 = Not engaged, 5 = Very engaged).

Table 2. Implementation fidelity components, definitions, and evaluation.

There are four methods of gathering implementation fidelity evidence via the checklist: auditors of the "live" programming, videos of programming, facilitators of the programming, and students engaged in the programming. Auditors experience the programming as "students." Using auditors is resource-intensive, especially for long programs, but auditors provide a real-time, authentic evaluation of program implementation.

Fidelity Results	Outcomes Results	Inferences that can be made from coupling data
High	Good	Programming was implemented as planned and the outcomes were met; thus, the planned programming may be effective. That is, the planned programming may be contributing to meeting intended outcomes.
Low	Poor	No claims can be made about the planned programming, because the planned program was not implemented. Moreover, the intended outcomes were not observed. A new study should be conducted with increased implementation fidelity to assess the effectiveness of the planned programming. Do not claim the planned program was ineffective.
High	Poor	Programming was implemented as planned, but the intended outcomes were not observed. Thus, low implementation fidelity can be ruled out as the reason for poor outcomes. Outcome assessment results should contribute to informed changes to the planned programming.
Low	Good	Programming was not implemented as planned. Thus, the planned programming cannot be credited with contributing to students meeting the outcomes. One should not claim the planned program was effective.

Table 3. Coupling implementation fidelity and outcomes assessment data.

Video recording the program allows for a large number of raters, including independent auditors and program facilitators. However, recordings may not allow for an authentic representation of the program, especially if programming involves a large number of people in a large setting. Also, the presence of a camera may influence those facilitating or engaging in programming. When program facilitators gather implementation fidelity data, this activity serves as a reminder of the agreed upon, research-informed programming features. The simple act of asking facilitators to review programming features and then indicate whether they implemented those features communicates the importance of executing the program as planned. In addition, the process of gathering facilitator ratings may reduce time needed to retrain facilitators (Durlak & DuPre, 2008). Gathering fidelity data from both program facilitators and auditors provides an opportunity to assess inter-rater reliability. If facilitators and auditors provide similar ratings, then auditors are not needed to gather this data. Collecting implementation fidelity data from students may be especially helpful for understanding student responsiveness and has the added benefit of communicating to students the intentionality and commitment to high-quality programming that is implemented equitably.

Pairing implementation fidelity and outcomes data provides insights that neither set of data could provide independently (O'Donnell, 2008). If planned programming was not implemented, outcomes data indicate nothing about this programming (Table 3). If planned programming was implemented with high fidelity, then outcomes data provide

insight into its effectiveness. This coupling of evidence guides changes to programming and resource allocation.

Given the utility of implementation fidelity data to reflect programming experienced by students, we hope it is gathered for most programming in higher education. Possible reasons why implementation fidelity is not commonly gathered include lack of knowledge of implementation fidelity, lack of guidelines on procedures to collect data, and lack of requirements to collect data (Gerstner & Finney, 2013). To address lack of knowledge and procedures, we call for more published examples of gathering and using implementation fidelity evidence in higher education. Additional exemplars would expand upon published examples in the domains of communication studies courses (Meixner, et al., 2020), orientation (Gerstner & Finney, 2013), and ethical reasoning curriculum (Smith, et al. 2017, 2019).

Recommendations for All Higher Education Professionals

Professionals can take immediate actions to integrate program theory and implementation fidelity into practice.

1. Consider why existing programming is provided, especially if outcomes have not been assessed. Is programming grounded in previous research showing its effectiveness?
2. Ask yourself and others involved with programming to answer, "Why *should* this programming be effective and for *whom* is it appropriate?" This exercise typically prompts necessary conversations.
3. Stay current on research in the outcome domain and in learning, cognition, and motivation to inform program theory, specifically the intermediate outcomes that will be assessed.
4. Articulate program theory via logic models and tables when developing new programming and reviewing existing programs. These efficient visuals coherently explain program logic.
5. Consider implementation fidelity and how to assess it, especially for programs implemented by multiple individuals. Do not assume programming is implemented well or equitably.
6. Create an implementation fidelity checklist, data collection strategy, and plan for coupling implementation and outcomes data to inform suggestions for program improvement.
7. Build a library of evidence-based successful practices to borrow from and adopt.
8. Integrate program theory and implementation fidelity throughout program review.

Recommended Updates to CAS Standards to Emphasize Program Theory and Implementation Fidelity

To address the lack of formal requirements for program theory and implementation fidelity data, we recommend expanding professional standards to include both. We believe the CAS standards provide an opportunity to speak to the importance of program theory and implementation fidelity, which we outline below. We encourage other associations (e.g., ACPA, NASPA) to examine their standards and professional competencies for opportunities to integrate program theory and implementation fidelity.

For over 40 years, the Council for the Advancement of Standards in Higher Education (CAS) has developed standards of good practice. CAS is a consortium of over 40 higher education associations representing over 115,000 professionals. CAS "promotes the use of its professional standards for the development, assessment, and improvement of quality student learning, programs, and services" (CAS Mission Statement, 2015, para. 2). The standards are guided by five principles derived from theories and conceptual models that inform the work of higher education professionals: students and their environments; advocating for diverse, equitable, and inclusive communities; organization, leadership, and human resources; ethical considerations; and learning-conducive structures, resources, and systems (CAS, 2019). The general standards, embedded within all 47 functional area standards, state that functional areas **must** be intentionally designed using theories of learning, development, and success and that professionals **must** remain current regarding research and theories that affect programs and services (CAS, 2019). Although these statements allude to program theory and implementation fidelity, no direct and explicit reference to either exists.

With a bedrock in higher education evaluation, it seems fitting that CAS serve an active role in promoting program theory and implementation fidelity in the assessment of programs. We believe there are four areas where CAS can more explicitly include program theory and implementation fidelity within the standards.

First, the contextual statements should include program theory and implementation fidelity. Each set of standards includes a contextual statement, which offers background and perspective on the functional area. The contextual statement introduces the nature, foundational principles, and current issues of the functional area. Each statement includes a historical perspective, important tenets, current issues, and references. Given program theory explains why a program is expected to work (i.e., why activities lead to outcomes), the contextual statement should outline this rationale. When CAS revises its standards, we recommend that the contextual statements include information to inform program theory. Although it is not realistic that contextual statements summarize all relevant research, it is reasonable for each contextual statement to provide a solid introduction to applicable theory and research to guide professionals in implementing informed interventions. Examples of evidence-informed interventions can be included in contextual statements as examples of program theory emerge in each functional area.

Second, the general standards should include program theory and implementation fidelity. According to CAS, all functional areas have identifiable commonalities. Thus, CAS incorporated common criteria, known as the general standards. Because the general standards appear verbatim in each set of functional area standards, we recommend the next revision include standards on program theory and implementation fidelity. Specifically, "Part 2: Program" and "Part 4: Assessment" should be revised to include program theory and implementation fidelity standards.

Third, the three sets of cross-functional frameworks (e.g., First Year Experiences) should include program theory and implementation fidelity. CAS defines their cross-functional approach as addressing issues or topics using a multi- and inter-disciplinary perspective by employing teams of higher education professionals from different fields. Cross-functional frameworks are organized into six parts. We recommend program theory be integrated into "Part 3. Strategy, Approach, and Processes" and implementation fidelity into "Part 6. Assessment".

Fourth, the CAS Self-Assessment Guides (SAGs) should include program theory and implementation fidelity. In addition to the CAS standards being tools to design new programs, direct assessment efforts, and guide staff development, CAS standards are used for program review. CAS provides a SAG for each functional area that includes a comprehensive self-study process. Each SAG contains: 1) instructions for conducting self-assessment using the SAG, 2) an overview of the process, and 3) the instrument comprised of criterion statements, rating scales, overview questions, evaluation forms, and work forms. If a revision of the general standards includes program theory and implementation fidelity, the SAG criterion statements should be revised to rate the functional area's effectiveness. In addition, each SAG section includes overview questions that enhance the review, provide narrative detail, and evaluate the program at a more holistic level. A revision of these overview questions is necessary to ensure important questions posed by program theory and implementation fidelity are included. Perhaps more important, the SAG instructions should include a description of program theory and implementation fidelity to assist internal and/or external reviewers of the program throughout the process. Reviewers must understand these concepts to conduct high-quality reviews.

In closing, mandates to gather outcomes assessment data to assess program effectiveness can divert attention from the equally important expectation of building programming informed by research and implementing it well. Creating evidence-informed programming and gathering implementation fidelity data can be challenging. Nonetheless, both types of evidence can direct and motivate high-quality programming and efficient outcomes assessment efforts.

Mandates to gather outcomes assessment data to assess program effectiveness can divert attention from the equally important expectation of building programming informed by research and implementing it well. Both types of evidence can direct and motivate high-quality programming and efficient outcomes assessment efforts.

References

Ambrose, A., Bridges, M., DiPietro, M., Lovett, M., & Norman, M. (2010). *How learning works: Seven research-based principles for smart teaching.* San Francisco, CA: Jossey-Bass.

Baldwin, C., Hutchinson, S., & Magnuson, D. (2004). Program theory: A framework for theory-driven programming and evaluation. *Therapeutic Recreation Journal, 38,* 16-31.

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

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

Brousselle, A., & Champagne, F. (2011). Program theory evaluation: Logic analysis. *Evaluation and Program Planning, 34,* 69-78.

Carpenter, D. (2012). Foreword. In K. Collins & D. Roberts (Eds.), *Learning in not a sprint: Assessing and documenting student leader learning in cocurricular involvement* (pp. vii – ix). Washington, DC: NASPA-Student Affairs Administrators in Higher Education.

Carpenter, D. (2001). Student affairs scholarship (re?)considered: Toward a scholarship of practice. *Journal of College Student Development, 42,* 301–318.

Council for the Advancement of Standards in Higher Education. (2019). *CAS professional standards for higher education* (10th ed.). Washington, DC: Author.

Council for the Advancement of Standards in Higher Education. (2015). *Mission statement.* https://www.cas.edu/mission

Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41,* 327–350.

Finley, A. (2019). *A comprehensive approach to assessment of high-impact practices.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

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

Finney, S., & Horst, S. (2019b). The status of assessment, evaluation, and research in student affairs. In V. L. Wise & Z. Davenport (Eds.), *Student affairs assessment, evaluation, and research: A guidebook for graduate students and new professionals* (pp. 3-19). Springfield, IL: Charles Thomas Publisher.

Finney, S., & Smith, K. (2016). *Ignorance is not bliss: Implementation fidelity and learning improvement.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Fisher, R., Smith, K. L., Finney, S., & Pinder, K. (2014). The importance of implementation fidelity data for evaluating program effectiveness. *About Campus, 19,* 28-32.

Fulcher, K., & Prendergast, C. (2019). Lots of assessment, little improvement? How to fix a broken system. In S. Hundley & S. Kahn (Eds.), *Trends in assessment: Ideas, opportunities, and issues in higher education* (pp. 157-174). Sterling, VA: Stylus.

Fulcher K., Smith, K., Sanchez, E., Ames, A., & Meixner, C. (2017). Return of the pig: Standards for learning improvement. *Research & Practice in Assessment, 11,* 10-27.

Gerstner, J., & Finney, S. (2013). Measuring the implementation fidelity of student affairs programs: A critical component of the outcomes assessment cycle. *Research & Practice in Assessment, 8,* 15-28.

Halpern, D., & Hakel, M. (2003). Applying the science of learning. *Change: The Magazine of Higher Learning, 35*, 36-41.

Hutchings, P., Kinzie, J., & Kuh, G. (2015). *Evidence of student learning: What counts and what matters for improvement.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Jankowski, N., & Baker, G. (2020). *Mapping and assessing student learning in student affairs.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Jankowski, N., Timmer, J., Kinzie, J., & Kuh, G. (2018). *Assessment that matters: Trending toward practices that document authentic student learning.* Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment.

Kerr, K., Tweedy, J., Edwards, K., & Kimmel, D. (2017), Shifting to curricular approaches to learning beyond the classroom. *About Campus, 22,* 22-31.

Lattuca, L., & Stark, J. (2009). *Shaping the college curriculum: Academic plans in context.* San Francisco, CA: Jossey-Bass.

Mathers, C., Finney, S., & Hathcoat, J. (2018). Student learning in higher education: A longitudinal analysis and faculty discussion. *Assessment and Evaluation in Higher Education, 43,* 1211-1227.

Meixner, C., Pope, A., & Horst, S. (2020). Implementation fidelity in the classroom: Exploring the alignment of pedagogy to learning outcomes. *The Journal of Faculty Development, 34,* 15-22.

Millett, C., Saunders, S., & Kevelson, M. (2018). T*he logic underlying a research-based college access program: Depicting the theory of change of the Princeton University Preparatory Program* (Research Report No. RR-18-07). Princeton, NJ: Educational Testing Service.

O'Donnell, C. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78,* 33-84.

Oehme, K., Perko, A., Altemus, M., Ray, E., Arpan, L., & Clark, J. (2020). Lessons from a student resilience project. *Journal of College Student Development 61*, 396-399.

Pope, A., Finney, S., & Bare, A. (2019). The essential role of program theory: Fostering theory-driven practice and high-quality outcomes assessment in student affairs. *Research & Practice in Assessment, 14,* 5-17.

Pope, A., Finney, S. & Crewe, M. (in press). Evaluating the effectiveness of an academic success program: Showcasing the importance of theory to practice. *Journal of Student Affairs Inquiry.*

Renger, R., & Hurley, C. (2006). From theory to practice: Lessons learned in the application of the ATM approach to developing logic models. *Evaluation and Program Planning, 29,* 106-119.

Renger, R., & Titcomb, A. (2002). A three-step approach to teaching logic models. *American Journal of Evaluation, 23,* 493-503.

Reinholz, D., & Andrews, T. (2020). Change theory and theory of change: What's the difference anyway? *International Journal of SEM Education, 7*, 1-12.

Smith, K., & Finney, S. (2020) Elevating program theory and implementation fidelity in higher education: Modeling the process via an ethical reasoning curriculum. *Research & Practice in Assessment, 15,* 1-13.

Smith, K., Finney, S., & Fulcher, K. (2017). Actionable steps for engaging assessment practitioners and faculty in implementation fidelity research. *Research & Practice in Assessment, 12,* 71-86.

Smith, K., Finney, S., & Fulcher, K. (2019). Connecting assessment practices with curricula and pedagogy via implementation fidelity data. *Assessment and Evaluation in Higher Education, 44,* 263–282.

Smith, K., Good, M., & Jankowski, N. (2018). Considerations and resources for the learning improvement facilitator. *Research & Practice in Assessment, 13,* 20-26.

Swain, M., Finney, S., & Gerstner, J. (2013). A practical approach to assessing implementation fidelity. *Assessment Update, 25*(1), p 5-7, 13.

Suskie, L. (2018). *Assessing student learning: A common sense guide.* San Francisco, CA: Jossey-Bass.

Vesely-Maillefer, A. (2015). *Striving for teaching success: Enhancing emotional intelligence in pre-service teachers.* Electronic Thesis and Dissertation Repository, The University of Western Ontario, Paper 3371.

U.S. Department of Education. (2006). *A test of leadership: Charting the future of U.S. higher education* (ED Pubs No. ED002591P).

Walton, G., & Brady, S. (2017). The many questions of belonging. In A. Elliot, C. Dweck, & D. Yeager (Eds.). *Handbook of Competence and Motivation: Theory and Application* (2nd Edition), (pp. 272-293). New York, NY: Guilford Press.

Walton, G., & Cohen, G. (2011). A brief social-belonging intervention improves academic and health outcomes among minority students. *Science, 331,* 1447–1451.

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

About the Authors

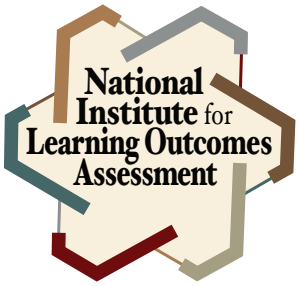
Sara Finney, PhD is Professor for the Assessment and Measurement PhD program at James Madison University. She is also Associate Director of the Center for Assessment and Research Studies, providing support to colleagues assessing learning and development outcomes. Her research focuses on the merger of implementation fidelity data and learning outcomes data when assessing program effectiveness, test-taking motivation and emotions, and the use of structural equation modeling to gather validity evidence for score interpretations.

Jennifer Wells, Ph.D is an Assistant Professor of Higher Education in the Department of Educational Leadership at Kennesaw State University. Jen is the current publications editor for the Council for the Advancement of Standards in Higher Education (CAS). Her current research focus is on the effectiveness of the CAS standards, the importance of continuous improvement efforts, and the psychosocial development of students.

Gavin Henning, PhD is Professor of Higher Education at New England College, where he directs the Doctorate of Education and Master of Science in Higher Education Administration programs. Gavin is Past-President of CAS as well as a Past-President of ACPA. His scholarship includes two books on student affairs assessment as well as numerous articles on the topic.

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.



Sign up to receive our monthly NILOA Newsletter and stay up to date with our research and publications.

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