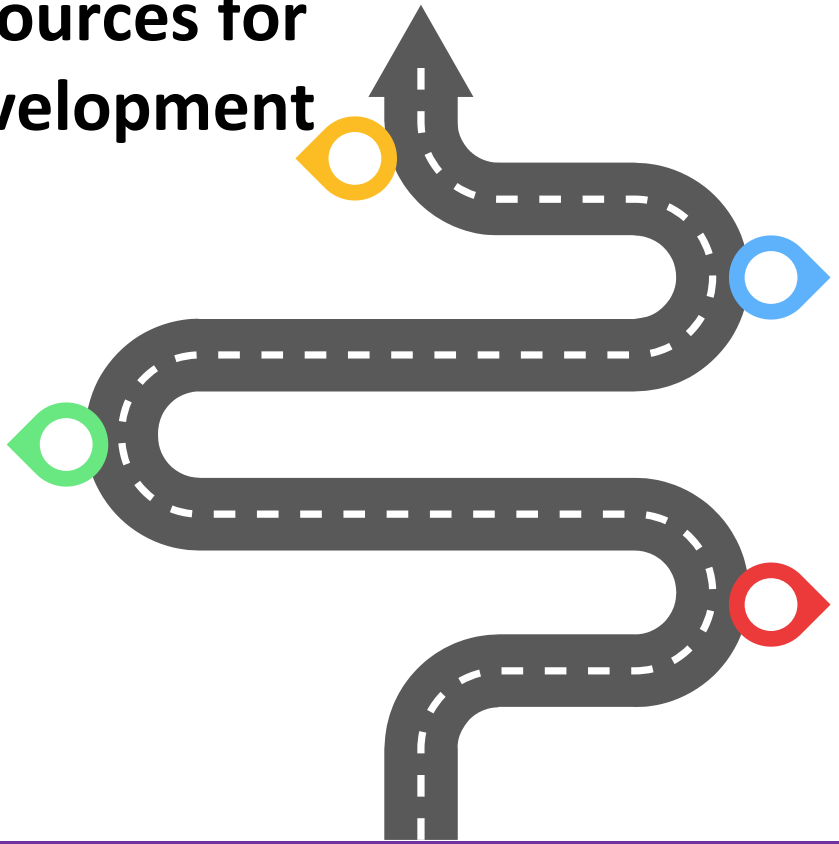
Design Smarter, Assess Better: A Roadmap to our Suite of Resources for Evidence-Based Program Development and Outcomes Assessment

Mara McFadden, Sara Finney,
Jada Willse, Jonathan Stewart



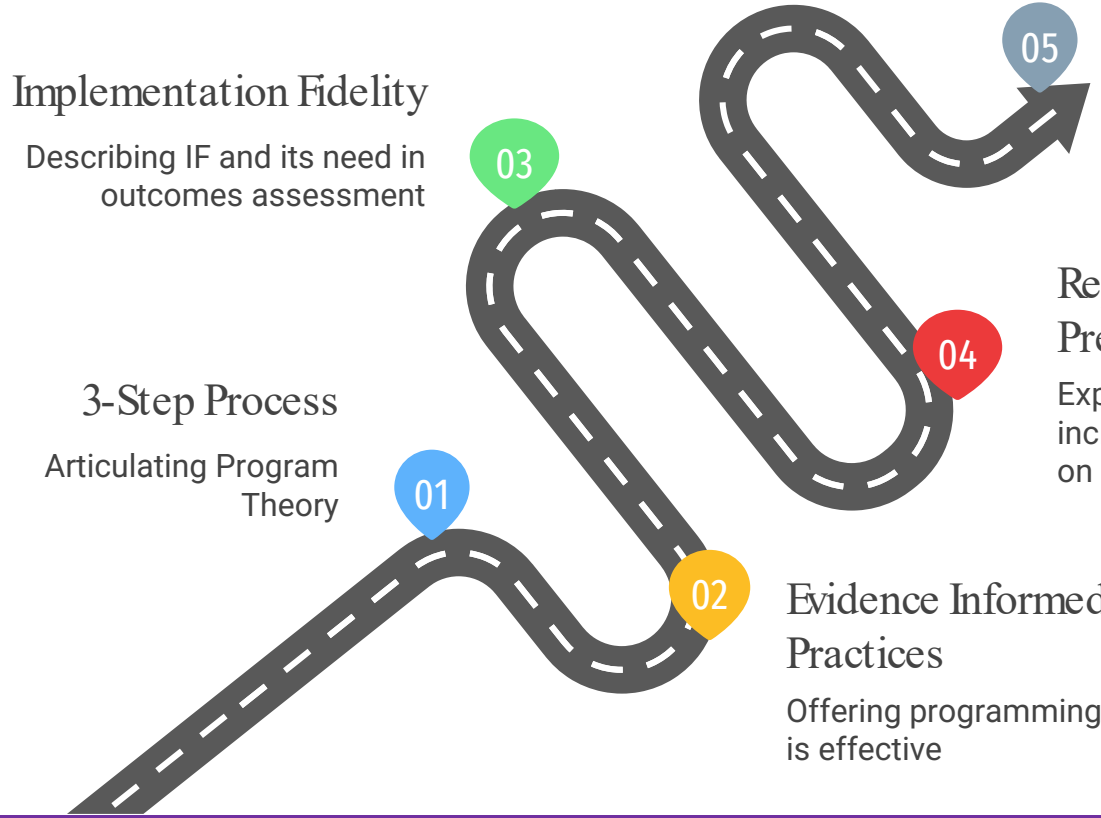
Land Acknowledgement

We invite you to recognize the written histories of the Shenandoah Valley, the city of Harrisonburg, and our university's namesake, James Madison, as fractured.

Let us acknowledge that we are currently on the land of the Indigenous Siouan, Algonquian, and Haudenosaunee communities who lived here for many generations and who continue to be systematically erased by policies and practices that remove their histories from this place.

Let us honor the enslaved people who built the wealth and foundation of James Madison. Let us recognize the histories of Virginia and the United States as complicit with the racism of white supremacy.

We recognize that these difficult histories persist in present-day racial realities and privileges at this university. We commit to dismantling racism in spaces of our work. We invite you to work beside us to create change.



3-Step Process

Articulating Program Theory

Implementation Fidelity

Describing IF and its need in outcomes assessment

Evidence Informed Practices

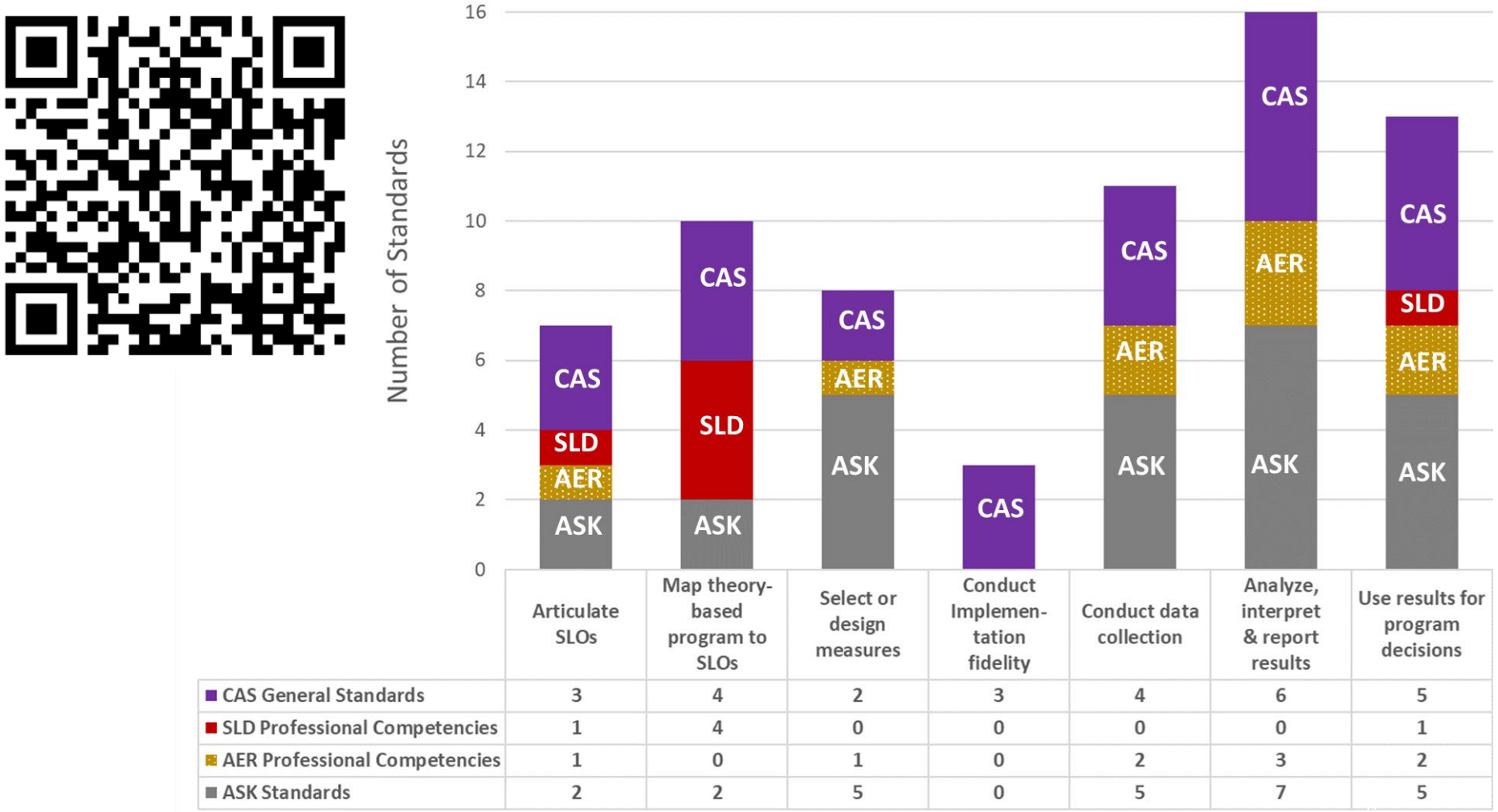
Offering programming that is effective

Repositories of Pre-existing Measures

Expert review of measures, including recommendations on whether it should be used

Meta-Assessment Rubric

Makes the assessment process transparent to all involved



	Articulate SLOs	Map theory-based program to SLOs	Select or design measures	Conduct Implemen-tation fidelity	Conduct data collection	Analyze, interpret & report results	Use results for program decisions
CAS General Standards	3	4	2	3	4	6	5
SLD Professional Competencies	1	4	0	0	0	0	1
AER Professional Competencies	1	0	1	0	2	3	2
ASK Standards	2	2	5	0	5	7	5

Expected Training in CSPA/HESA Programs

CAS (2023) standards related to graduate programs: **“Master’s Level Higher Education and Student Affairs Professional Preparation Programs.”**

Subpart 5b.5 - Assessment, Evaluation, and Research (CAS, 2023, p. 720)
Program curriculum **must** include the study of assessment, evaluation, and research that centers on evidence-based practice to further accountability and continuous improvement. Content must include assessment planning and design; outcome development; qualitative, quantitative, mixed methods, and critical data collection and analysis methods; measurement of learning processes and outcomes; assessment of environments and organizations; measurement of program and environment effectiveness; effective reporting; critiques of published studies; integration of social justice; and assessment and change management strategies….

Gaps in Training in CSPA/HESA Programs

Upon reading these "**must**" statements expected of graduate curriculum, one would believe SA educators would have skills in evidence-based programming, outcomes assessment, and use of results for improvement, all with an equity frame.

- However, a review of 111 syllabi from assessment/evaluation courses within student affairs graduate programs found "limited" alignment with professional standards (ACPA et al., 2024).
- Only 1.8% of syllabi had a primary theme of equity-centered assessment, indicating a clear discrepancy between expectations set by professional standards & current practice in graduate-level programs.

Gaps in Training →Impact Job Performance

Concerns are further fueled because many SA educators enter the workforce lacking assessment skills (Cooper et al., 2016; Hoffman, 2015).

- In a mixed-methods study examining perceptions of students in a SA graduate program, ~31% of students reported they did *not* learn how to "evaluate research," a skill needed to inform programming decisions (Wright-Mair et al., 2022).
- Additionally, 47.83% of respondents indicated they did *not* learn how to "understand statistics," & 34.78% indicated they did *not* learn how to "communicate results," both necessary skills to use assessment data for improvement.
- 61% of leadership personnel strongly agreed that they expected new hires to have skills in assessment; however, nearly 1/3 of respondents indicated that new hires were *not* adequately prepared with these skills (Dean & Langham, 2022).

Because SA graduate programs inconsistently train educators regarding assessment, "on the job" professional development is needed to ensure proper training.

We are here to offer our FREE resources.

Beyond Student Learning Outcomes…

We all have things we want students to know, think, value, and be able to do via the experience they have on our college campuses

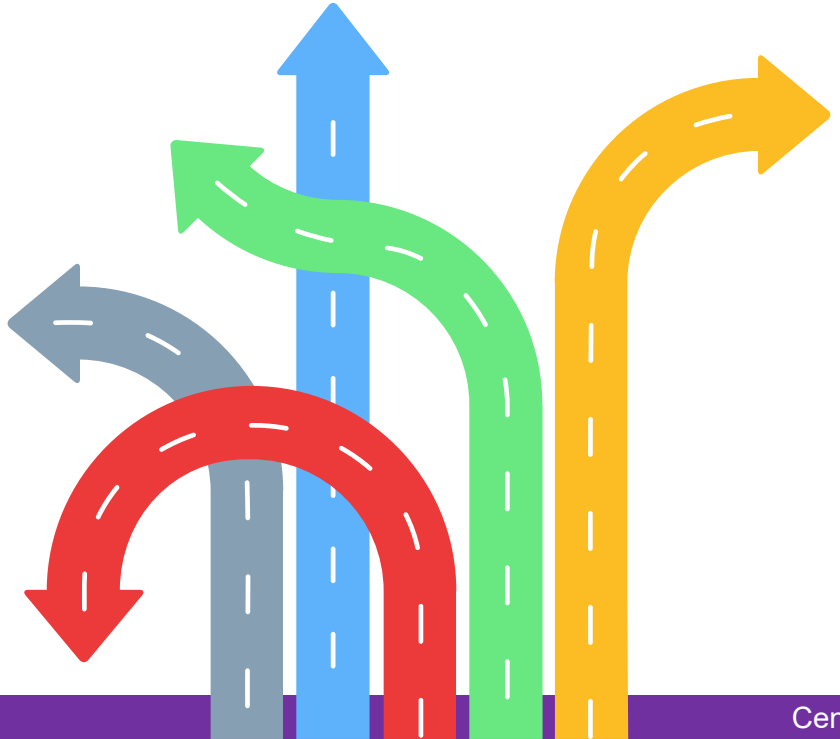
How do you build programming that will result in those outcomes?

How do you evaluate the quality of your assessment process?

How do you measure those outcomes?

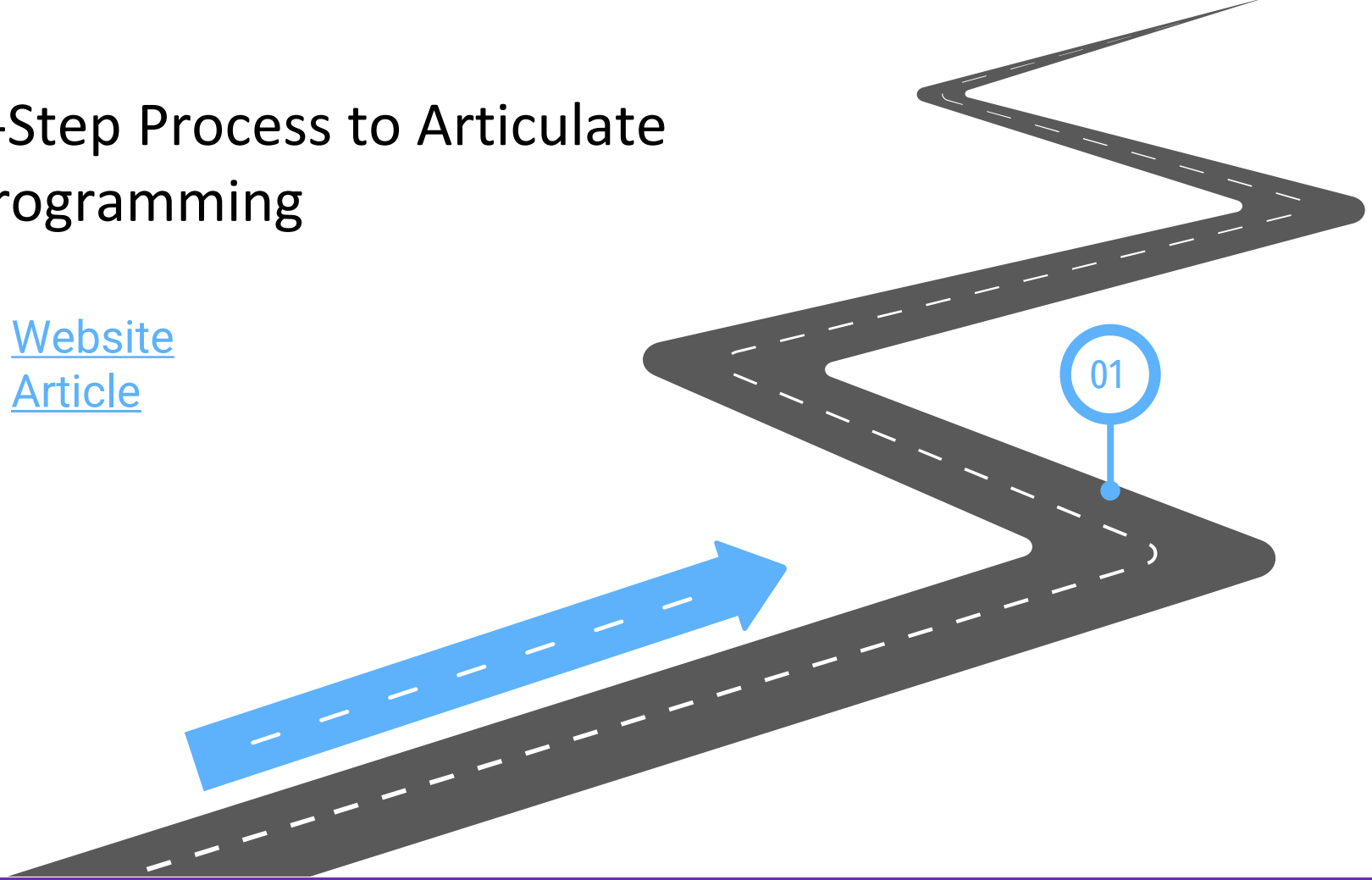
I think I need to build a new program

How do I start?



3-Step Process to Articulate Programming

- Website
- Article



Website on Program Creation and Mapping

JMU > Assessment & Research Studies > Student Affairs Assessment Support Services

Program Creation & Mapping

On This Page...

The second step in the Assessment Cycle is **programming with purpose**. Now that you have specified your student learning and development outcomes (SLOs), you need to 1) determine *how* to achieve those outcomes (i.e., develop a program theory) and 2) clearly articulate the connection between program components and SLOs (i.e., develop a logic model and/or program-to-outcome map). **During this step of the process, you will need to consider questions such as:**

- Which components of your program are hypothesized to lead to which SLOs?
- *Why* do you believe these program components should lead to the stated SLOs?
- What do relevant theories say about how to impact your SLOs?
- *What evidence-based practices* have been shown to impact your SLOs?

3-Step Process to Articulate Logic of Programming

Abstract

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

LEARNING IMPROVEMENT

AUTHORS

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

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

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

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

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

3-Step Process to Articulate Logic of Programming

Pope, A., Finney, S.J., & 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.

Step 1: State Appropriate/Feasible Distal Outcome

- What is the distal outcome?
- What do you ultimately hope to achieve?

Step 2: ??

Step 3: ??

Step 1: State the Distal Outcome

Distal Outcome

How Do You Choose the Distal Outcome?

CAS Standards Outcomes

Exhibits behaviors of a leader (Leadership)
Engages in behaviors that promote health (Health/Wellness)
Seeks involvement with people different from self (Appreciation of Differences, Diversity)

Focus in Division or at the University

Civically Engaged
Demonstrates a Global Perspective
Ethical Behavior

3-Step Process to Articulate Logic of Programming

Step 1: State Appropriate/Feasible Distal Outcome

- What is the distal outcome?
- What do you ultimately hope to achieve?

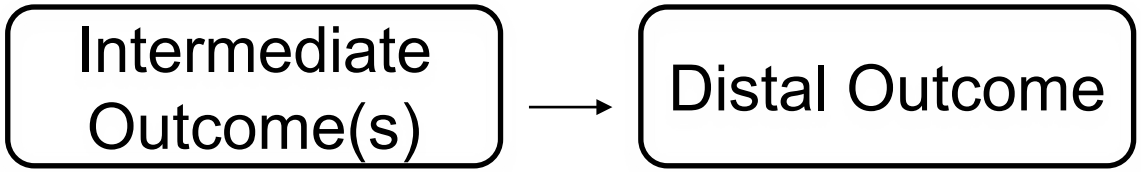
Step 2: Specify Intermediate Outcomes

Explain their role in specifying program theory.

- Specify attitudes, skills, & behaviors that ***influence*** the distal outcome
- What is the *etiology* (cause, reason, origin) of the distal problem, behavior, skill?

Step 3: ??

Step 2. Specify *Intermediate* Student Learning Outcomes



How Do You Specify the Intermediate Outcomes?

Ask yourself: What knowledge, attitudes, skills, and/or behaviors will the program need to cultivate to achieve the distal outcome?

Then Go Find the Answers--Read the Research

Exhibits behaviors of a leader (Leadership)

Research has shown students need to know X, think/value/perceive Y, do Z to exhibit leadership behaviors

Engages in behaviors that promote health (Health/Wellness)

Research has shown students need to know X, think/value/perceive Y, do Z to engage in health behaviors

Seeks involvement with people different from self (Appreciation of Differences, Diversity)

Research has shown students need to know X, think/value/perceive Y, do Z to engage with people different from them

3-Step Process to Articulate Logic of Programming

Step 1: State Appropriate/Feasible Distal Outcome

- What is the distal outcome?
- What do you ultimately hope to achieve?

Step 2: Specify Intermediate (More Proximal) Outcomes

- How do you achieve your distal outcome?
  - Specify the attitudes, skills, & behaviors that influence the distal outcome
- What is the etiology (cause, reason, origin) of the distal problem, behavior, skills?

Step 3: Develop Program Components

- Knowing the intermediate outcomes helps you develop **theory- or research-based program components** to help achieve *the intermediate outcomes*, leading to the achievement of *your distal outcome*

I've articulated the logic of my programming

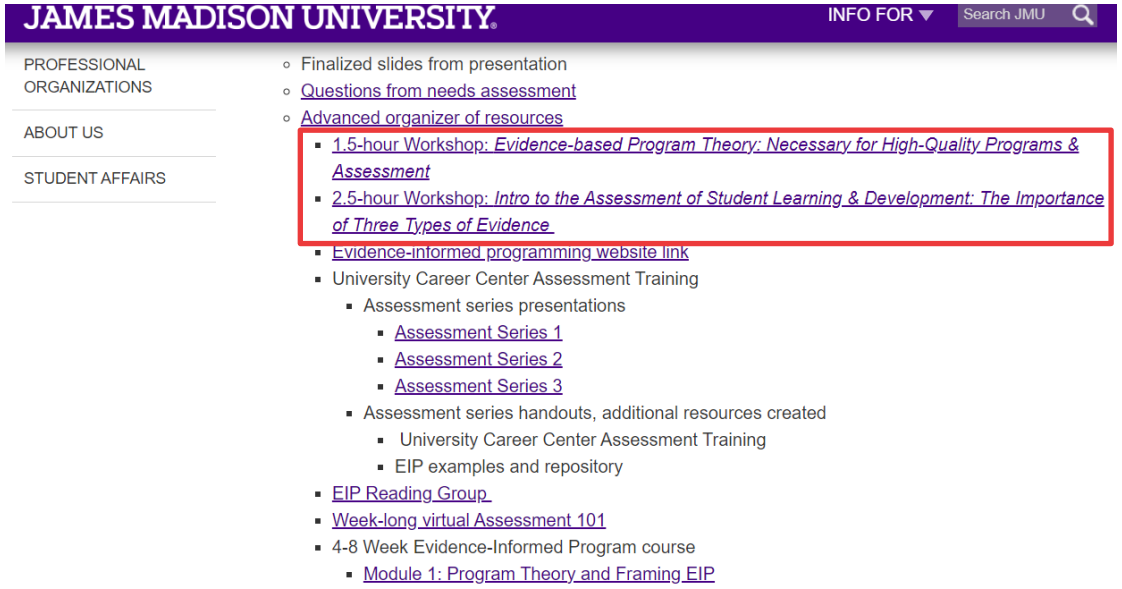
How do I select programming that will be the most effective?

Resources on Evidence Informed Practices

- Website - EIP Professional Development Resources
- Website – Best Available Evidence

Website of EIP Learning Resources

- Organized in order of **ease of implementation**, from easiest to most difficult
- Ease of implementation takes into consideration a number of factors, including time commitment & amount of self-led learning required



EIP & Outcomes Assessment: Assessment 101

1.5 hour workshop on Program Theory

- "*Evidence-based Program Theory: Necessary for High-Quality Programs & Assessment*"
- Emphasis on using evidence for selection of outcomes (malleable, feasible) & programming, where to find credible evidence, how to address equity considerations

2.5 hour Intro to Assessment in SA with focus on EIP

- *"Intro to the Assessment of Student Learning & Development: Importance of Three Types of Evidence"*
- To engage in program improvement efforts efficiently & effectively 3 types of evidence are necessary: what has been ***shown*** to be effective, what programming did students ***experience*** & what were the ***outcomes*** in in this context with these students

Developing Competency & Confidence in Evidence-Informed Practice via Professional Development Experiences

- Finalized slides from presentation
- Questions from needs assessment
- Advanced organizer of resources
  - 1.5-hour Workshop: *Evidence-based Program Theory: Necessary for High-Quality Programs & Assessment*
  - 2.5-hour Workshop: *Intro to the Assessment of Student Learning & Development: The Importance of Three Types of Evidence*
  - Evidence-informed programming website link
  - University Career Center Assessment Training
    - Assessment series presentations
      - Assessment Series 1
      - Assessment Series 2
      - Assessment Series 3
  - EIP Reading Group
  - Week-long virtual Assessment 101
  - 4-8 Week Evidence-Informed Program course
    - Module 1: Program Theory and Framing EIP
    - Module 2: Value of EIP
    - Module 3: Finding Credible Evidence
    - Module 4: Evaluating Evidence
    - Capstone Project: EIP Cover Letter and Interview
  - Curricular Approach
    - Professional development consultant
    - Reading group questions

4 Week EIP Course

- Four modules & a capstone project that align directly with results of needs assessment
  1. Introduction: Program Theory & Framing EIP
  2. The Value of EIP
  3. Finding Credible Evidence
  4. Evaluating Evidence
  5. Capstone Project: EIP Focused Cover Letter and Mock Interview
- Variety of readings with questions, PowerPoints, & activities
- Canvas course made available to all SA employees
- Limited asynchronous use from SA Professionals
- Considering annual workshops using the materials
- Modules taught within CSPA course "Professional Issues in Higher Education"

Evidence-Based Programming Website

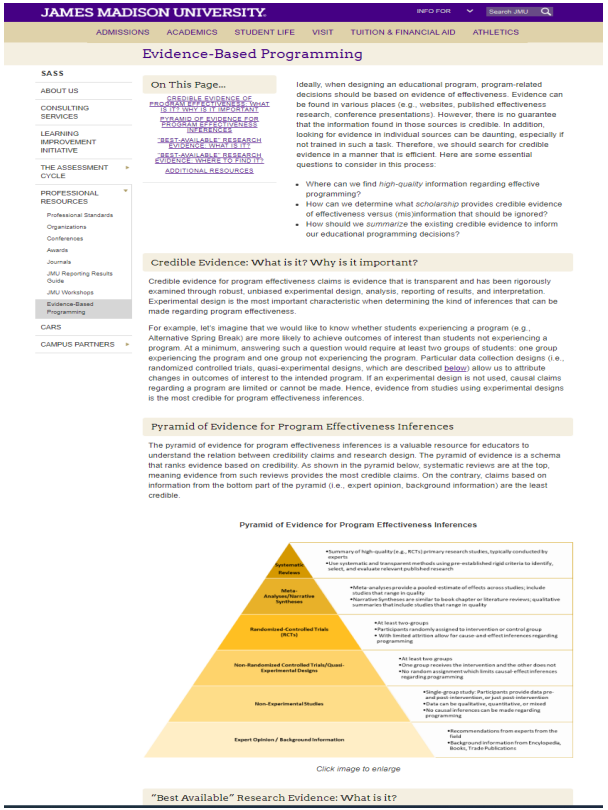
WHAT: website that helps student affairs educators answer following questions:

- *Where* can we find high-quality information regarding effective programming?
- *How* can we determine what scholarship provides credible evidence of effectiveness vs (mis)information that should be ignored?
- *How* should we summarize existing credible evidence to inform educational programming decisions?

WHY: Needs assessment indicated lack of competence & confidence to find, evaluate, & use evidence to inform practice; everyone can't engage in semester-long reading group

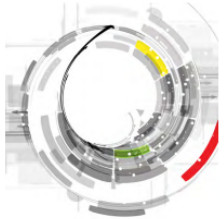
WHO: Available to anyone or any office

TIME: Can be completed in 1 week



Repositories of Effectiveness Studies

RESEARCH & PRACTICE IN ASSESSMENT



Abstract

Identifying evidence-informed programming (e.g., strategies, activities, pedagogies) facilitates both the intentional offering of programming that should "work" and the use of the outcomes assessment process to evaluate program effectiveness. Evidence-informed programming is more efficient than unsupported programming because the programming is more likely to improve learning and development. Thus, faculty and student affairs professionals require fewer iterations of the assessment cycle to inform programming changes in order to achieve desired outcomes. To help locate evidence-informed programming, we describe systematic review repositories (e.g., *Campbell Collaboration*, *What Works Clearinghouse*) that synthesize high-quality research to identify "what works". We share a tool we created that organizes relevant systematic review repositories and other collections of evidence of effectiveness, providing numerous examples of evidence-informed programming pertinent to higher education. These resources aid faculty and student affairs professionals in achieving their ethical obligation to engage students in effective learning and development experiences.

AUTHORS

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

Holly A. Buchanan, DNP, ANP-BC
James Madison University

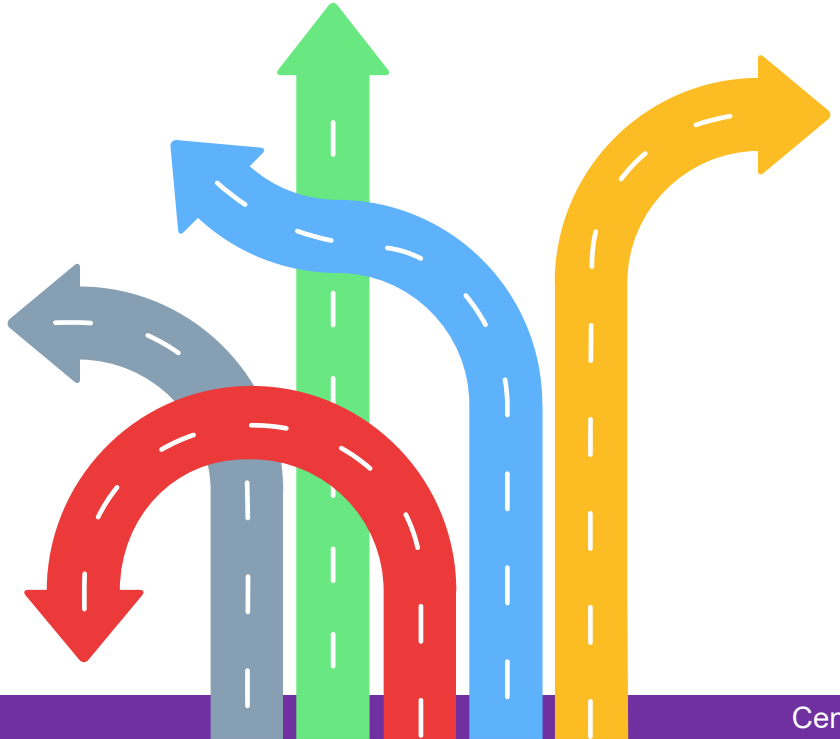
A More Efficient Path to Learning Improvement: Using Repositories of Effectiveness Studies to Guide Evidence-Informed Programming

Repository	Description	Examples Relevant to Higher Education
Campbell Collaboration	Exists to help people make well-informed decisions about **social & behavioral interventions**. Provides systematic reviews of programs or interventions using rigorous review & synthesis processes of high-quality (RCTs or quasi-experimental designs) primary research. Some research designs have such weak internal validity that they are unacceptable in reviews to inform effective claims (e.g., simple before-after programming studies without comparison groups).	• Bystander Intervention • Mindfulness-based Stress Reduction • Motivational Interviewing for Substance Abuse • Exercise to Improve Self-Esteem in Young People • Advocacy Interventions to Reduce Violence & Promote Well-Being of Women who Experience Partner Abuse
What Works Clearinghouse	A trusted source of scientific evidence on **education programs, practices, & policies**. WWC reviews research, determines which studies meet rigorous standards (RCTs, quasi-experimental designs), summarizes findings, and provides practice guides.	• Using Technology To Support Postsecondary Learning • Linked Learning Communities • Organizing Instruction & Study to Improve Learning • First Year Experience Courses • Strategies for Postsecondary Students in Developmental Education
Cochrane Library	Provides short plain language summaries of their longer systematic reviews of empirical research that focus on interventions for **health outcomes** (e.g., alcohol, STIs). Indicates the quality of the studies that informed their conclusions.	• Social norms interventions are not effective enough on their own to reduce alcohol misuse among college students • Self-help & Guided Self-help for Eating Disorders • Prevention of Suicide in University Settings

Note. RCTs = Randomized Controlled Trials.

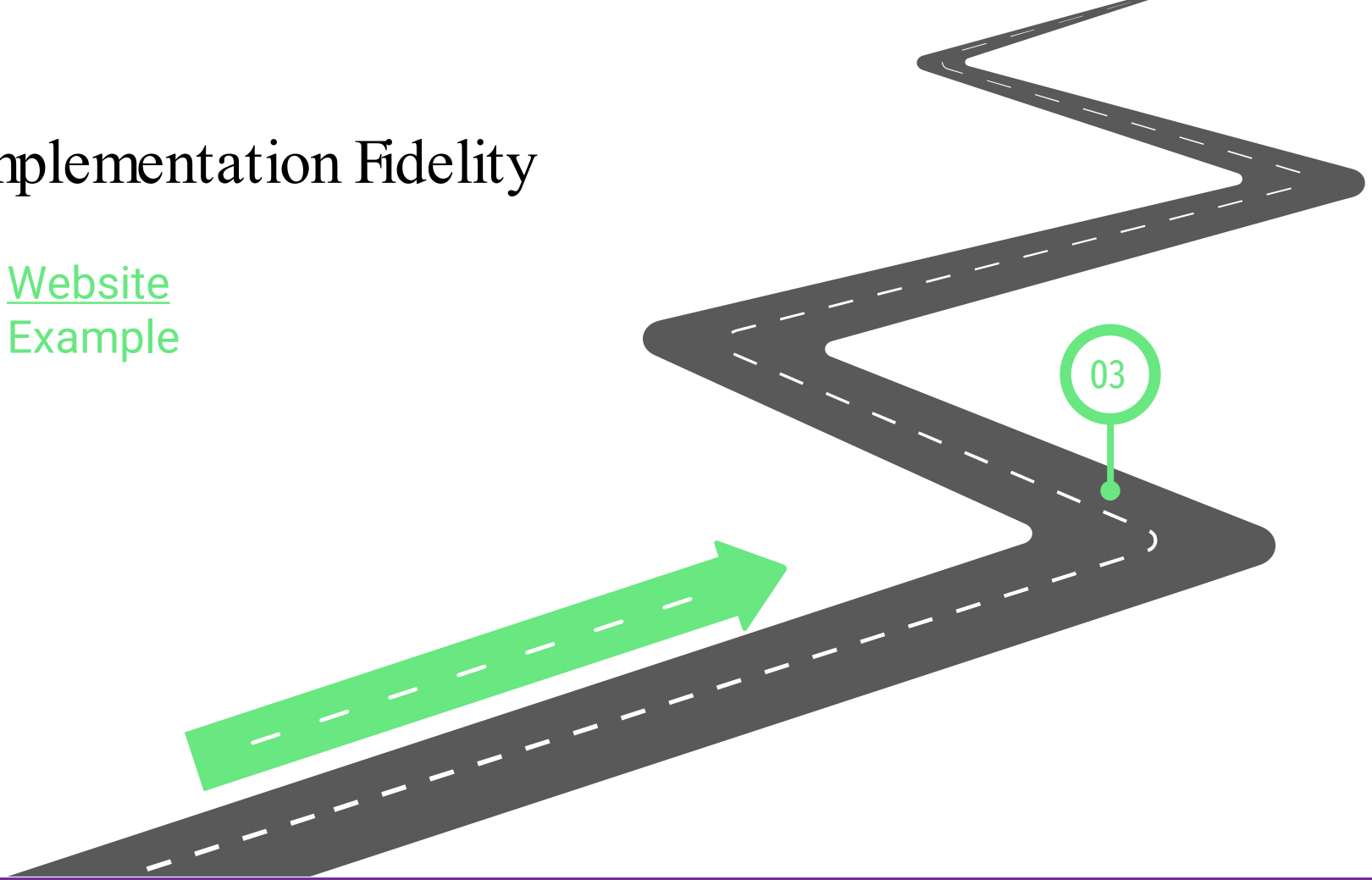
I've found sound programming components.

How do I make sure they're being implemented as intended?

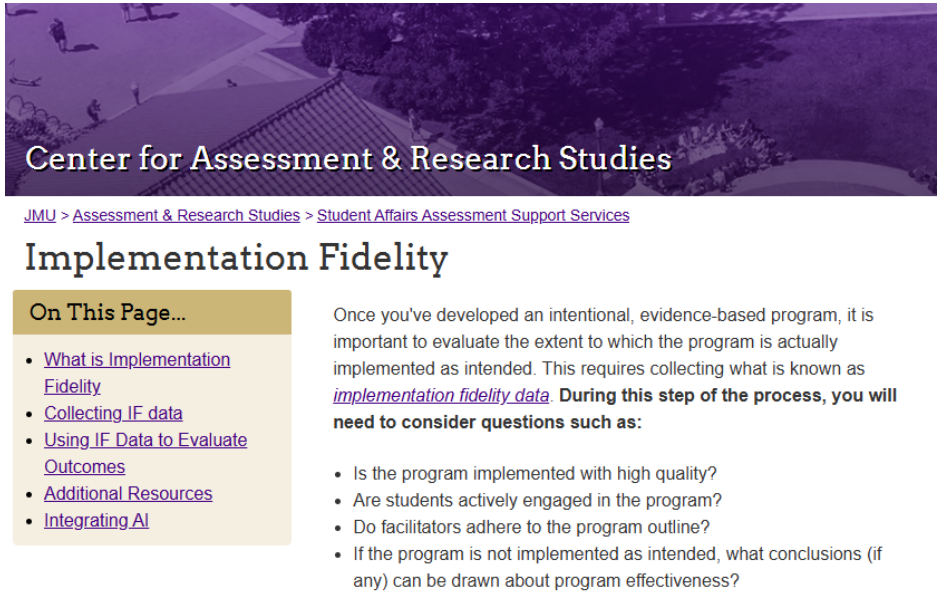


Implementation Fidelity

- Website
- Example



Implementation Fidelity Website



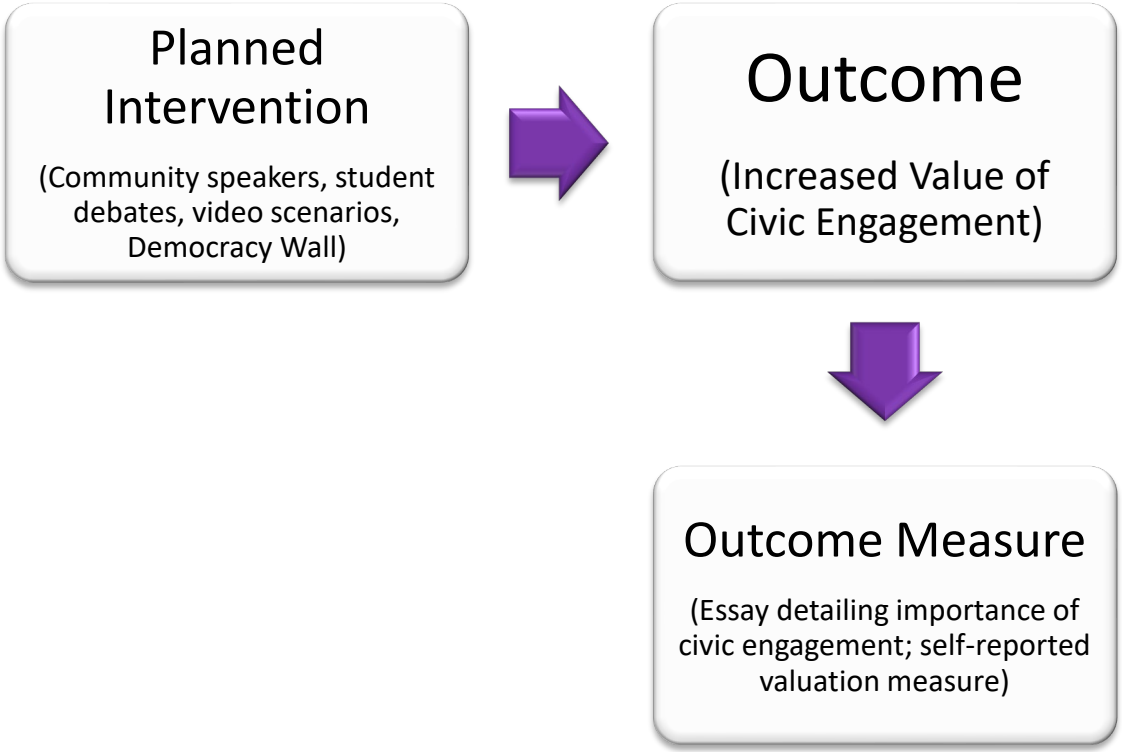
What is Implementation Fidelity?

- “The regrettable fact is that many good ideas are not well implemented or implemented fully." - Tinto, 2006, p. 8

- "The bridge between a promising idea and the impact on students is implementation, but innovations are seldom implemented as intended." -Berman & McLaughlin, 1976, p. 349

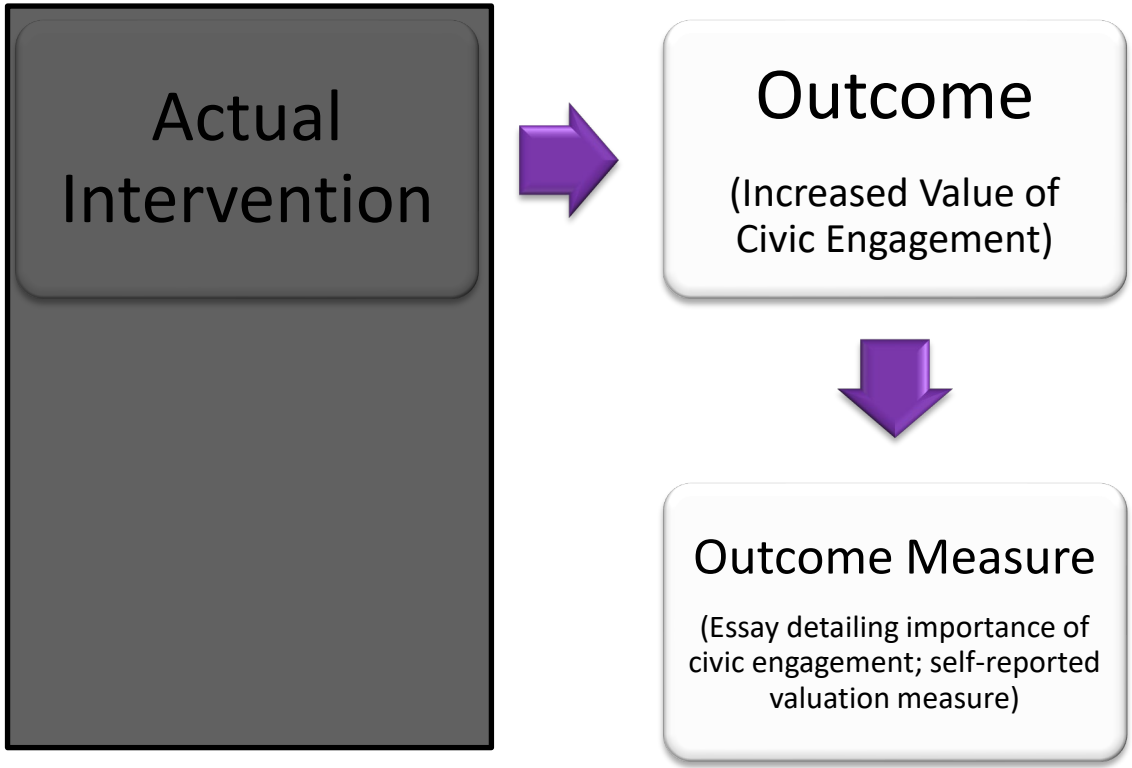
- "We owe it to our students to ensure HIPs and other innovations intended to enhance the quality of undergraduate education are implemented equitably and with fidelity so that students realize the promised benefits." - Kuh & Kinzie, 2018

“Planned” Civic Engagement Intervention



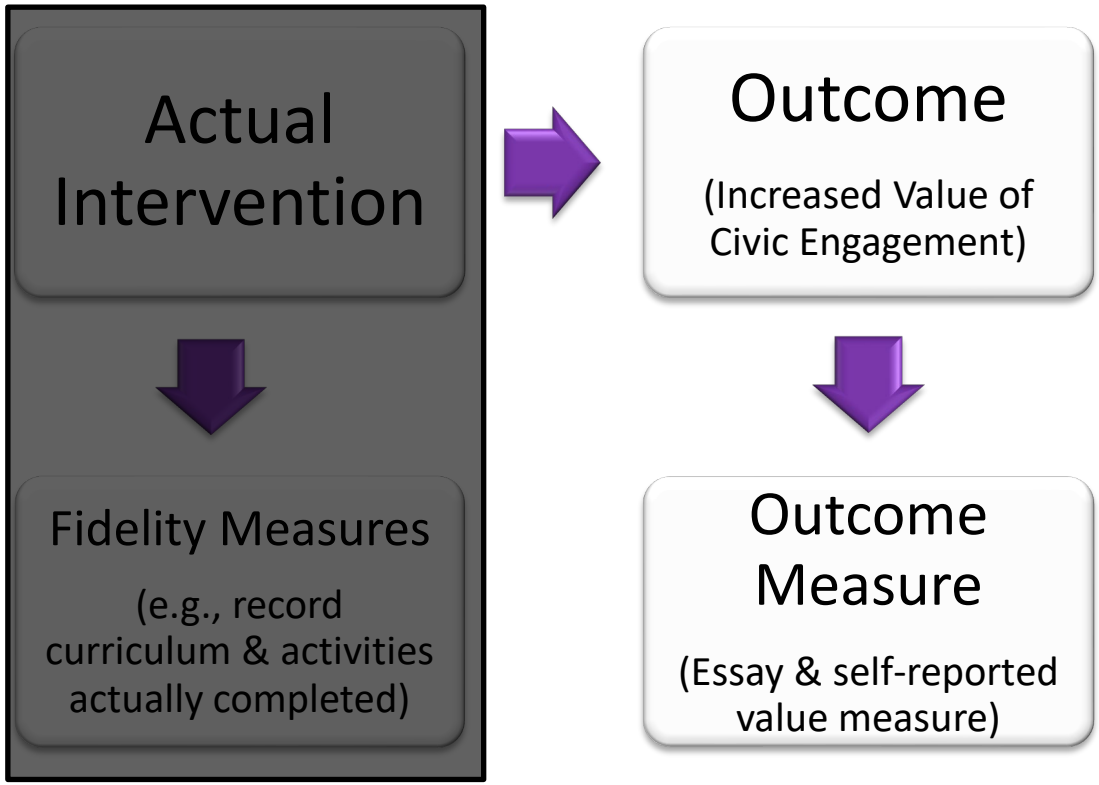
The Actual Civic Engagement Intervention is a "Black Box"

If students can't articulate the value of civic engagement or their self-reported value doesn't increase, does that mean the ***planned*** curriculum is ineffective?

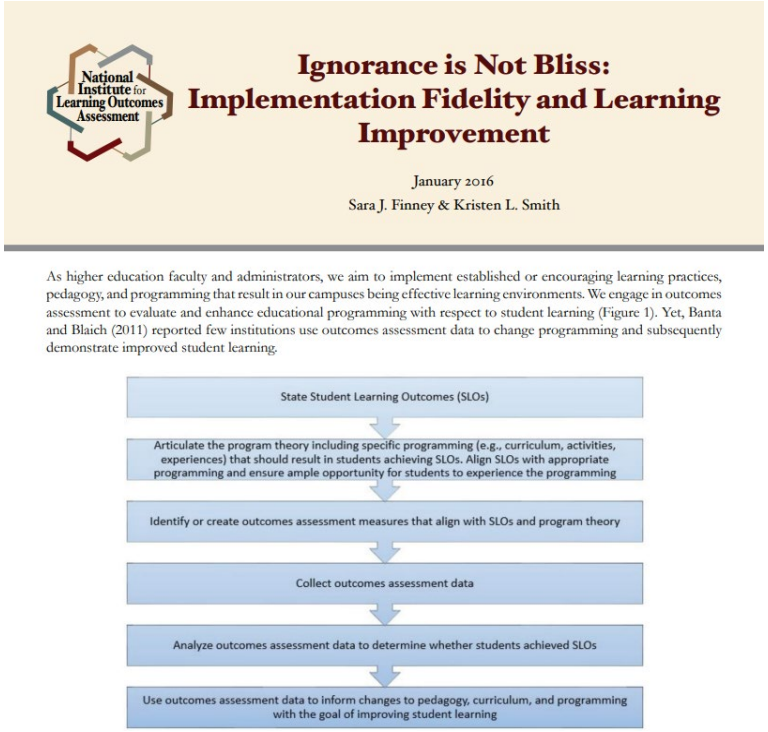


Implementation Fidelity Assessment can Open the Black Box

Imagine that only one speaker came to the university and that the students only engaged with the democracy wall, they didn't engage in debates or complete video scenarios. If our outcomes aren't met, we can't blame the ***planned*** program! We can't answer these questions without IF data



More Resources on Webpage



National Institute for Learning Outcomes Assessment

Ignorance is Not Bliss: Implementation Fidelity and Learning Improvement

January 2016

Sara J. Finney & Kristen L. Smith

As higher education faculty and administrators, we aim to implement established or encouraging learning practices, pedagogy, and programming that result in our campuses being effective learning environments. We engage in outcomes assessment to evaluate and enhance educational programming with respect to student learning (Figure 1). Yet, Banta and Blaich (2011) reported few institutions use outcomes assessment data to change programming and subsequently demonstrate improved student learning.

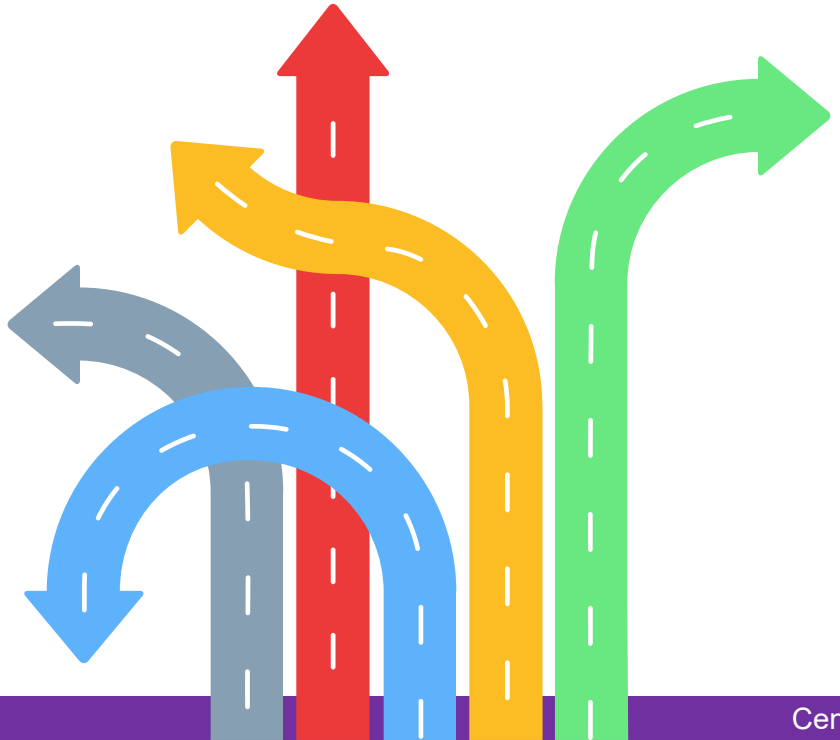
Measuring the Implementation Fidelity of Student Affairs Programs: A Critical Component of the Outcomes Assessment Cycle

What is Implementation Fidelity and Why is it Important?

Implementation fidelity has been discussed in many domains (e.g., K-12 education, health, psychology). As a result, numerous definitions of implementation fidelity exist. The general definition provided by O'Donnell (2008) is "the determination of how well an intervention is implemented in comparison with the original program design during an efficacy and/or effectiveness study" (p. 33). Specific to the student affairs context, implementation fidelity examines the extent to which the planned student affairs program matches the implemented program. That is, student affairs programs (or any educational program) should be designed thoughtfully to meet particular learning and development outcomes. However, as Berman and McLaughlin (1976) noted, "The bridge between a promising idea and the impact on students is implementation, but innovations are seldom implemented as intended" (p. 349). Importantly, research has shown that programs implemented with high fidelity have more of an impact with respect to program outcomes than those with low fidelity (e.g., Durlak & DuPre, 2008). Thus, higher education practitioners and instructors need to ask themselves, "Are students receiving the *planned* program?"

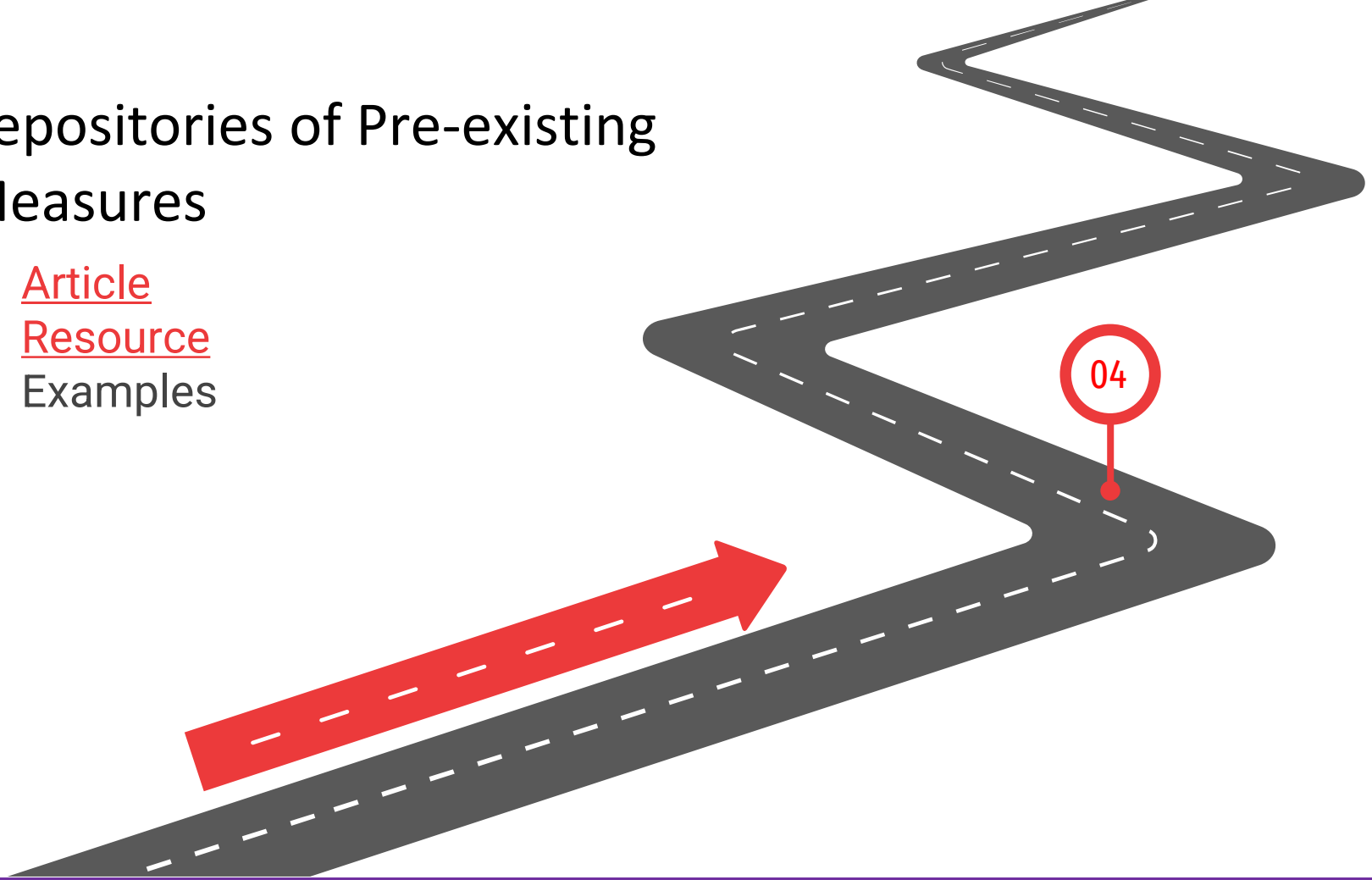
I have my programming & outcomes!

Now how do I assess those outcomes?

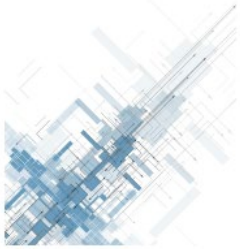


Repositories of Pre-existing Measures

1. Article
2. Resource
3. Examples



Finding Existing Measure



AUTHORS

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

Gabriel R. Gilmore, MA
James Madison University

Sarah Alahmadi, MS
James Madison University

Abstract

When engaging in outcomes assessment, higher education professionals (i.e., faculty, student affairs educators) are expected to gather reliable data and make valid inferences. Decisions about how to measure student learning and development outcomes impact inferences about the achievement of outcomes and determination of improvement efforts. Professionals may search for existing outcome measures due to lack of experience in the challenging instrument development process and/or the time required to construct a high-quality measure. To support professionals in their search, we created a tool that describes relevant repositories of measures. Given most professionals lack training in psychometrics, we purposefully categorized these repositories by the level of guidance they provide when selecting a measure. That is, in addition to identifying an existing measure and summarizing the measure's psychometric properties, some repositories provide an evaluation of the measure's quality. This resource facilitates the collection of high-quality data that informs valid inferences about student outcomes.

"What's A Good Measure Of That Outcome?" Resources To Find Existing And Psychometrically Sound Measures

Finding Existing Measure

Repositories of Pre-existing Measures

Purpose

To provide a resource for locating pre-existing measures, thus, potentially avoiding the need to create and study the properties of a newly created measure.

Organization of this Resource

We organized the repositories in this document into three tiers based on utility. Repositories are arranged in hierarchical order with those of higher utility listed first, followed by those of lower utility. When searching for measures, we recommend searching all repositories in all tiers starting at the top tier.

Tier 1. Repositories in this tier provide psychometric information (e.g., reliability, validity) for the measures, as well as their own reviews or ratings of the quality of the measures. Reviews or ratings can be in the form of a statement, number, or recommendation for use. Reviews or ratings may not be provided for every measure, but are available for the majority of measures. We consider repositories in this tier having the highest utility for the selection of evidence-informed pre-existing measures.

Tier 2. Repositories in this tier provide psychometric information (e.g., reliability, validity) for the measures, but do not provide their own reviews or ratings of the quality of the measures. Also, psychometric information may not be provided for every measure, but are available for most measures in the repository. The majority of the repositories in this document fall in this category.

Tier 3. Repositories in this tier do not provide psychometric information (e.g., reliability and validity) for the measures or their own ratings of the quality of the measures. Often, the psychometric information can be found in the linked source articles.

acquisition, construction, integration, and application; interpersonal competence; and practical competence bolded and*. Hence, in this repository, you will find measures that align with those specific student learning and development domains.

If you are not using the CAS outcome domains, but rather outcomes specified by LEAP (AAC&U), the Degree Qualifications Profile (DQP), Learning Reconsidered, or other organizations, we recommend this useful crosswalk of outcomes by organization to show their overlap.

In addition to providing a description of each repository, we listed five measures included in each repository. These five measures serve simply as examples and a mechanism to quickly access and examine the repository.

Notes

This document refers to both commercial and noncommercial measures. Commercial measures are copyrighted by the companies or organizations that created them and must be purchased for use. Noncommercial measures are publically available (e.g., published in journal articles) and do not require payment for use. Repositories that include commercial instruments or a mix of both are specified.

Some repositories in this document are books and may not be available in an online format. Links to Google previews or institutional access are provided when available. As of the publication of this document, all links to repositories and measures were active. However, many of the online repositories continually update their websites, so some of the links to the measures they house are subject to change.

If you use a measure from one of the listed

Finding Existing Measures

When selecting measures, ask yourself…

- Does this measure meet the standards of **psychometric quality**?
- Are included items **direct measures** of your outcome of interest?
- Does the measure **map to your outcome** of interest?
- Is the measure **equity centered**?

More specific evaluation of these questions can be found in the rubric

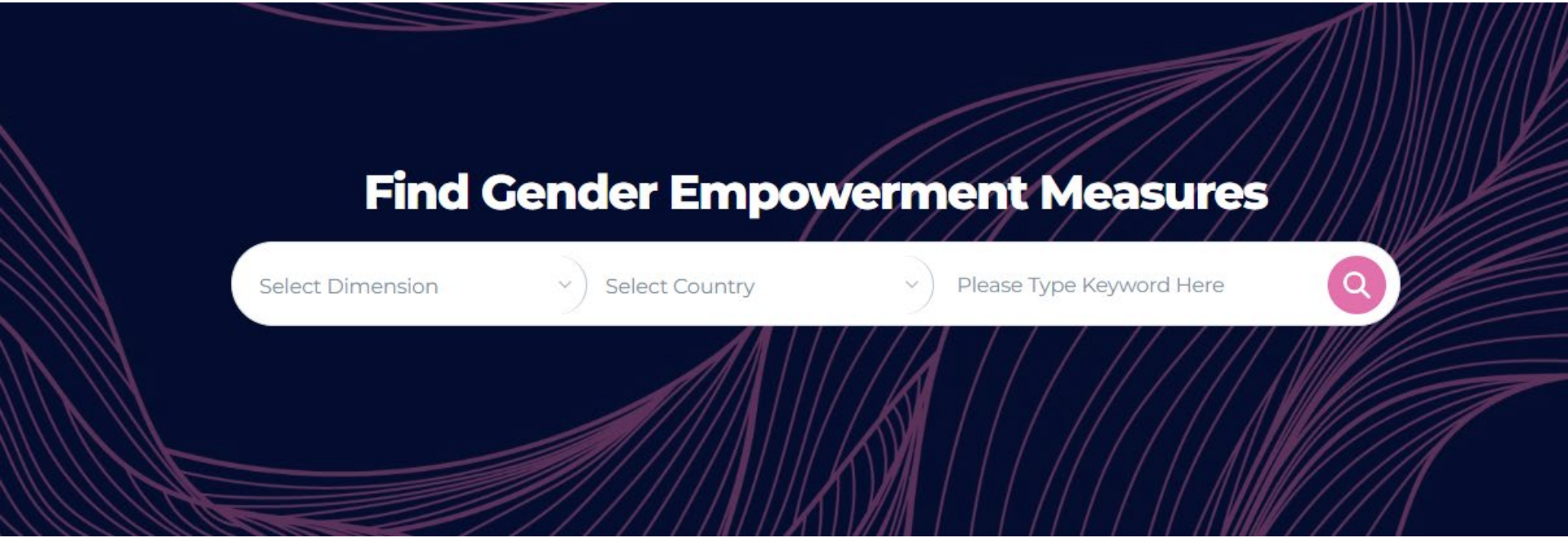
Finding Existing Measures

Repositories of Pre-existing Measures

- Three different tiers of repositories
  - Tier 1: Repositories that provide **psychometric information** and **expert ratings** of the quality of the measure
  - Tier 2: Repositories that provide **psychometric information**, but no expert ratings of the quality of the measure
  - Tier 3: Repositories that only provide **descriptive information**, no psychometric information nor expert ratings of the quality of the measure

EMERGE

Tier 1 | Provides Psychometric Information And Their Own Rating Of The Quality Of The Measure				
Source	**CAS Standards**	**Description of Resource**	**Information About the Characteristics of Measures**	**Examples of Measures**
EMERGE	**Knowledge acquisition, construction, integration, & application*** Cognitive complexity* Intrapersonal development* **Interpersonal competence*** Humanitarianism & Civic Engagement* **Practical competence***	Evidence-based Measures of Empowerment for Research on Gender Equality (EMERGE) is a project "focused on gender equality and empowerment measures used in India and other multi-country settings to monitor and evaluate health programs and to track progress on UN sustainable Development Goal (SDG) 5." One aim of EMERGE is to evaluate quantitative measures of gender equality and empowerment. They compiled a collection of measures that fall under the following domains: Social, Psychological, Education, Economic, Legal, Political, Health, Household and Intrafamilial Relations, Environment and Sustainability, and Time-Poverty. Although the focus of EMERGE is on India and other multi-country settings, the measures it houses do not appear to be culturally specific and could be applied to the U.S. EMERGE allows you to search for specific measures within its database, or you can select a dimension of gender equality and empowerment and it will populate all measures associated with that dimension.	EMERGE provides a description of each scale, including the number of items, subscales, response options, scoring procedures, primary citation, and more. For each measure, the entire scale is provided, with items broken down by subscales (if any). Each measure undergoes extensive psychometric evaluation. The details of this process can be found here. Measures are given a psychometric score between 1-7, 1-8, or 1-10 based on the type of applicable psychometric information, and labeled as low, medium, high, or no data. This score is broken down by reliability, validity, and formative research. These aspects are broken down further by a color-coded system of how many "points" were assigned based on the quality of the specific information available (e.g., full 1 point for test-retest reliability). Citation frequency is also scored as either no data, low (less than 20 citations), medium (20-49 citations), and high (50 or more citations). EMERGE also published an article on how to quantitatively measure gender equality and empowerment.	1. Self-Esteem Stability Scale 2. The Multidimensional Scale of Perceived Social Support 3. Condom Use Self-Efficacy Measure 4. Resiliency Attitudes and Skills Profile 5. Optimism/Pessimism Scale


Find Gender Empowerment Measures
Select Dimension
Select Country
Please Type Keyword Here

All of the measures on our website are open access and freely available for your use. You may use the measure as long as you cite the original citation as listed on the bottom of the respective measure page.

Select Dimension | Select Country | resilience scale

Brief Resilience Scale (BRS)

The Brief Resilience Scale (BRS) is a 6-item, self-reported measure of an individual's ability to bounce back or recover from stress. The items assess an individual's ability to bounce back quickly after a stressful event or life's set-backs.

High Psychometric Score

Connor-Davidson Resilience Scale

The Connor-Davidson Resilience Scale is a 25-item self-reported measure of personal resilience. Items were based on Kobasa, 1979, Rutter, 1985, and Lyons, 1991. Different aspects of resilience captured by this scale include one's...

High Psychometric Score

Widowhood Resilience Scale (WRS)

Widowhood Resilience Scale (WRS) is a 25-item measure of the attributes of resilience that impact positive adaptation after conjugal loss. This scale covers 6 domains of resilience: social support, integration, living in the...

Brief Resilience Scale (BRS)

The Brief Resilience Scale (BRS) is a 6-item, self-reported measure of an indivi
items assess an individual's ability to bounce back quickly after a stressful eve

See Google Citation

High Psychometric Score

Psychometric Score

HIGH

Ease of Use Score

NO DATA

Categories

Geographies Tested: United States of America

Populations Included: Female, Male

Age Range: Adolescents, Adults

Items:

1. I tend to bounce back quickly after hard times
2. I have a hard time making it through stressful events *(R)
3. It does not take me long to recover from a stressful event
4. It is hard for me to snap back when something bad happens *(R)
5. I usually come through difficult times with little trouble
6. I tend to take a long time to get over set-backs in my life *(R)

Response Options:

Strongly disagree - 1
Disagree - 2
Neutral - 3
Agree - 4
Strongly agree - 5

Scoring Breakdown

Formative Research

- Qualitative Research
- Existing Literature/Theoretical Framework
- Field Expert Input
- Cognitive Interviews / Pilot Testing

Ease Of Use

- Readability
- Scoring Clarity
- Length

Reliability

- Internal
- Test-retest
- Interrater

Validity

- Content
- Face
- Criterion (gold-standard)
- Construct

KEY

- Full Points
- Partial Points
- Not Assessed
- Not Applicable

COSMIN

Tier 2 – Provides Psychometric Information, But Not Their Own Rating Of The Quality Of The Measure				
Source	**CAS Standards**	**Description of Resource**	**Information About the Characteristics of Measures**	**Examples of Measures**
COSMIN	**Knowledge acquisition, construction, integration, & application*** **Cognitive complexity*** Intrapersonal development* **Interpersonal competence*** Humanitarianism & Civic Engagement* Practical competence*	COSMIN is an initiative of an international multidisciplinary team of researchers who aim to improve the selection of outcome measures both in research and in clinical practice. They develop methodology and practical tools for selecting the most suitable outcome measure. The database consists of 1,142 systematic reviews, last updated December 2018 that pertain to outcome measures. Goals • Advance the science and application of health outcome measurement • Develop new and update existing methodology and practical tools for the selection and use of outcome measures for research and clinical practice • Monitor and maintain the scientific quality of COSMIN tools • Encourage widespread adoption of the COMSIN methodology • Call for standardization of outcomes and outcome measures by developing Core Outcome Sets (COS) and COS methodology COSMIN also provide resources and guidelines for conducting systematic reviews, selecting instruments for a Core Outcome Set and a manual for using the database.	COSMIN collects and provides summaries of systematic reviews of outcome measures. Links are provided to the full text and the specific measures. These systematic reviews rate the quality of the outcome measures. They have developed a checklist for assessing study quality as well as checklists for assessing potential bias and best reporting practices. To be included in the database, studies must meet the following criteria: 1. The review should be a systematic review; 2. The aim of the review should be to identify all outcome measures of interest and to summarize the evidence of their measurement properties; 3. The construct of interest of the review should be aspects of health, defined as (a) biological and physiological processes, (b) symptoms, (c) physical functioning, (d) social/psychological functioning, (e) general health perceptions, or (f) health-related quality of life; 4. The population studied can be general population or patients; 5. The measures of interest should be outcome measures, defined as measures which can be/are applied in longitudinal studies to monitor changes in health over time; 6. The review should evaluate and report on at least one or more measurement properties of the included measures. Studies are excluded if they do NOT include all commonly used measures of a construct for a particular population.	1. A methodological review of resilience measurement scales 2. A review of the validity and reliability of smokeless tobacco dependence measures 3. Evaluation of the psychometric properties of self-reported measures of alcohol consumption: a COSMIN systematic review 4. A review of social inclusion measures 5. Systematic review of body image measures

Database of systematic reviews of outcome measurement instruments

All Fields | Search... | Search | More options

Limit your search

- Biological and physiological variables
- Symptom status
- Functional status
- General health perceptions / HRQoL
- Overall quality of life
- Age
- Disease

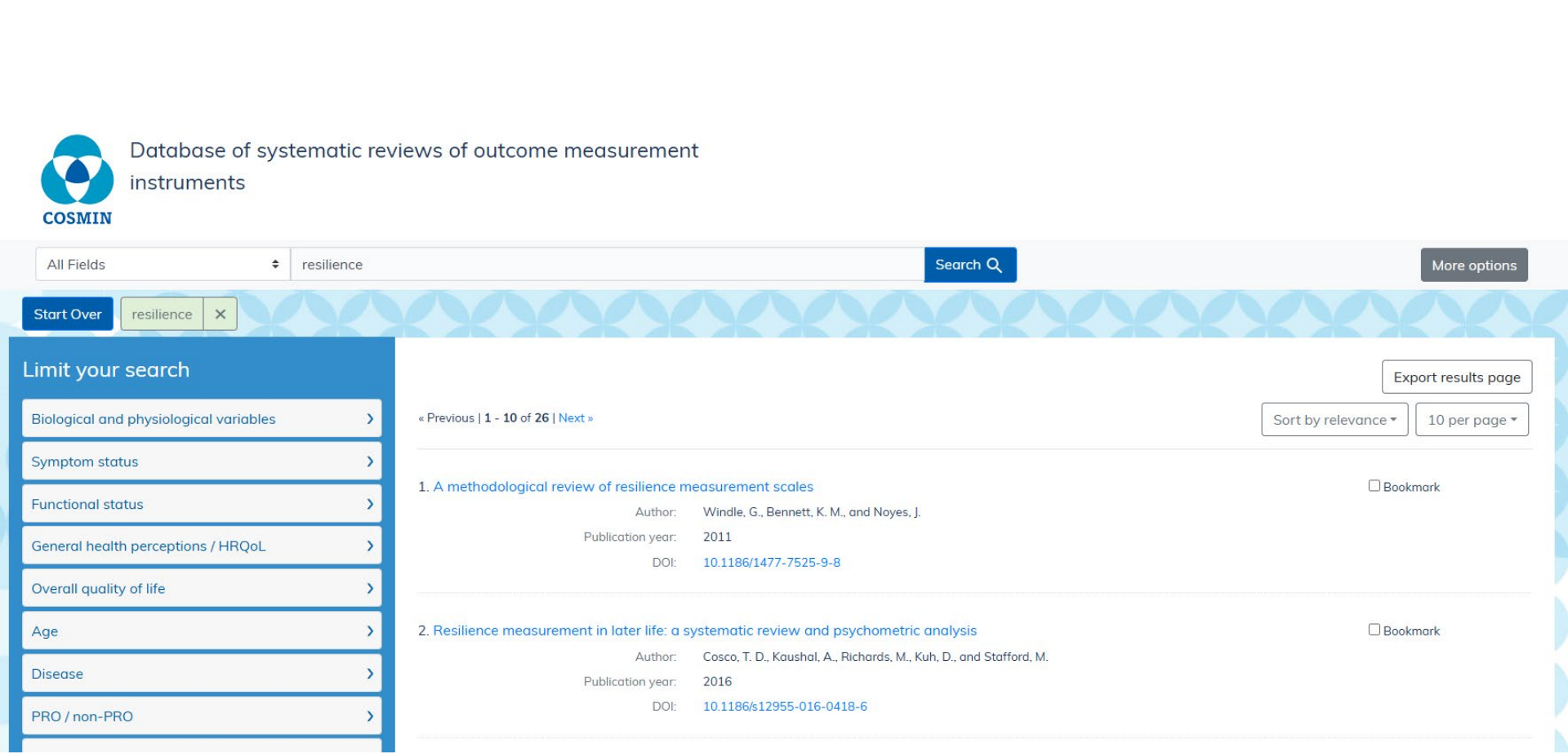
COSMIN database of systematic reviews of outcome measurement instruments

The COSMIN initiative aims to improve the selection of outcome measurement instruments for research and clinical practice by developing methodology and practical tools for selecting the most suitable outcome measurement instrument. As part of this initiative, COSMIN systematically collects systematic reviews of outcome measurement instruments. Such systematic reviews are important tools for the selection of outcome measurement instruments for research and clinical practice and for identifying gaps in knowledge on the quality measurement properties of outcome measurement instruments.

The COSMIN database of systematic reviews is now connected to the PROQOLID database of measurement instruments.

A systematic review is included in this database when it fulfills the following criteria:

1. The review should be a systematic review (i.e. a search in at least one electronic database was performed);
2. The aim of the review should be to identify all outcome measurement instruments of interest and to summarize the evidence on their measurement properties;
3. The construct of interest of the review should be (aspects of) health status, defined as (a) biological and physiological processes, OR (b) symptoms, OR (c) physical functioning, OR (d) social/psychological functioning, OR (e) general health perceptions, OR (f) health-related quality of life (based on the model of Wilson & Cleary, JAMA 1995);
4. The population of included studies should contain humans (patients or general population);
5. The instruments of interest should be outcome measurement instruments, defined as instruments which can be/are applied in longitudinal studies to monitor changes in health over time (the outcome measure is the dependent variable);
6. The review should evaluate and report on at least one or more measurement properties of the included instruments.


COSMIN
Database of systematic reviews of outcome measurement instruments
All Fields
resilience
Search
More options
Start Over
resilience
Limit your search
Biological and physiological variables
Symptom status
Functional status
General health perceptions / HRQoL
Overall quality of life
Age
Disease
PRO / non-PRO
Export results page
« Previous | 1 - 10 of 26 | Next »
Sort by relevance
10 per page
1. A methodological review of resilience measurement scales
Bookmark
Author: Windle, G., Bennett, K. M., and Noyes, J.
Publication year: 2011
DOI: 10.1186/1477-7525-9-8
2. Resilience measurement in later life: a systematic review and psychometric analysis
Bookmark
Author: Cosco, T. D., Kaushal, A., Richards, M., Kuh, D., and Stafford, M.
Publication year: 2016
DOI: 10.1186/s12955-016-0418-6

Resilience measurement in later life: a systematic review and psychometric analysis

Authors:	Cosco, T. D., Kaushal, A., Richards, M., Kuh, D., and Stafford, M.
Abstract:	**OBJECTIVES:** To systematically review and examine the psychometric properties of established resilience scales in older adults, i.e. >/=60 years. **METHODS:** A systematic review of Scopus and Web of Science databases was undertaken using the search strategy "resilience" AND (ageing OR aging)". Independent title/abstract and fulltext screening were undertaken, identifying original peer-reviewed English articles that conducted psychometric validation studies of resilience metrics in samples aged >/=60 years. Data on the reliability/validity of the included metrics were extracted from primary studies. **RESULTS:** Five thousand five hundred nine studies were identified by the database search, 426 used resilience psychometrics, and six psychometric analysis studies were included in the final analysis. These studies conducted analyses of the Connor Davidson Resilience Scale (CD-RISC) and its shortened 10-item version (CD-RISC10), the Resilience Scale (RS) and its shortened 5- (RS-5) and 11- (RS-11) item versions, and the Brief Resilient Coping Scale (BRCS). All scales demonstrated acceptable levels of internal consistency, convergent/discriminant validity and theoretical construct validity. Factor structures for the RS, RS-11 and CD-RISC diverged from the structures in the original studies. **CONCLUSION:** The RS, RS-5, RS-11, CD-RISC, CD-RISC10 and BRCS demonstrate psychometric robustness adequate for continued use in older populations. However, results from the current study and pre-existing theoretical construct validity studies most strongly support the use of the RS, with modest and preliminary support for the CD-RISC and BRCS, respectively. Future studies assessing the validity of these metrics in older populations, particularly with respect to factor structure, would further strengthen the case for the use of these scales.
DOI:	10.1186/s12955-016-0418-6
URL:	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4730639/pdf/12955_2016_Article_418.pdf
Journal:	Health Qual Life Outcomes
issn:	1477-7525 (Electronic); 1477-7525 (Linking)
Publication year:	2016
pages:	16
Symptom status:	Cognitive/mental state Emotional state
Functional status:	Cognitive/mental functioning
Age:	Adults (18-65) Seniors (65+)
Disease:	Factors influencing health status and contact with health services Mental and behavioural disorders and related symptoms
PRO / non-PRO:	Patient Reported Outcome
Type of measurement instrument:	1 - Questionnaires
Instrument:	BRCS - Brief Resilient Coping Scale PROQOLID CD-RISC - Connor-Davidson Resilience Scale (versions: CD-RISC-10) PROQOLID RS - Wagnild&Young's Resilience Scale (versions: RS-5; RS-11) PROQOLID

Finding Existing Measures

When selecting measures, ask yourself…

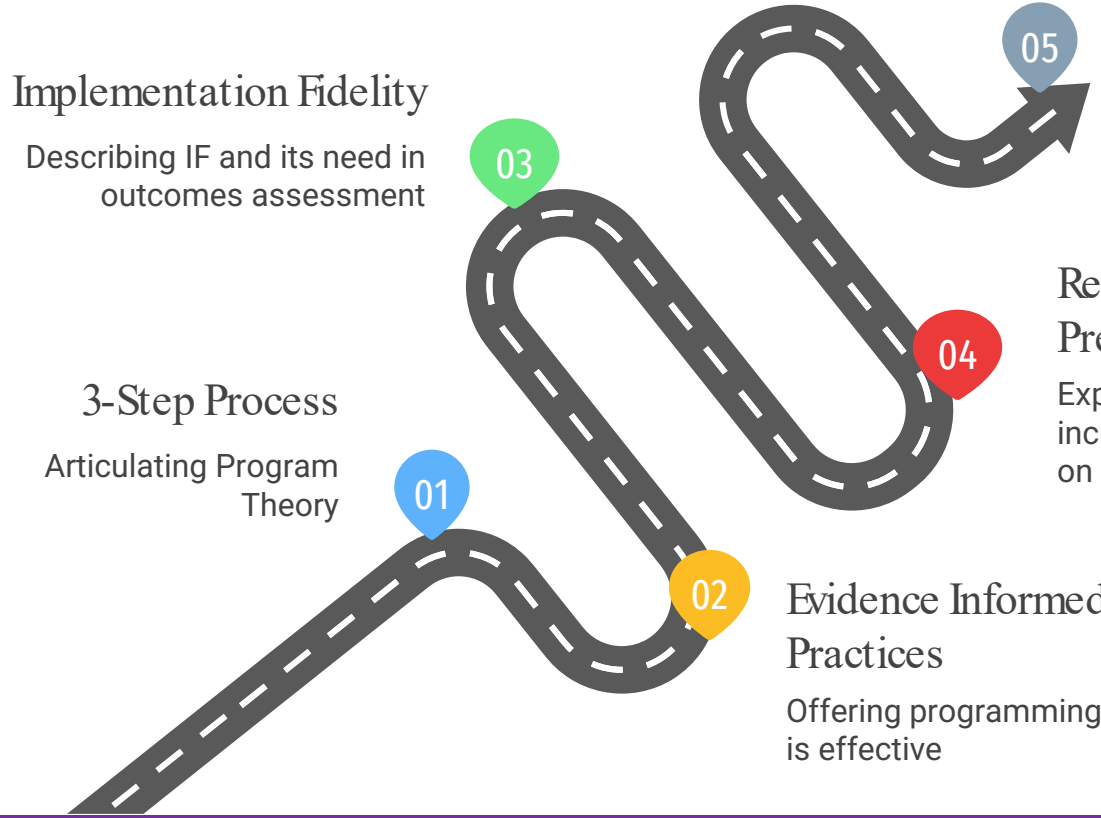
- Does this measure meet the standards of **psychometric quality**?
- Are included items **direct measures** of your outcome of interest?
- Does the measure **map to your outcome** of interest?
- Is the measure **equity centered**?

More specific evaluation of these questions can be found in the rubric

Selecting or Creating Measures of the Outcome	A description of how outcomes are operationalized is provided. Often discussed as "tools", "measures" or "instruments" and include self-report measures, tests, rubrics (to rate writing, performances, portfolios, products, presentations, etc.), or observation tools (e.g., did a student engage in particular behaviors?). A direct link should be made between specific measures and stated outcomes.			
Sub-Criteria	**Exemplary (3)**	**Proficient (2)**	**Developing (1)**	**Missing (0)**
Measure Selection	Pre-existing measures with **high-quality** psychometric properties have been chosen from the literature and those properties are **reported**.	Pre-existing measures with **adequate** psychometric properties have been chosen from the literature and those properties are **reported**.	Pre-existing measures with **poor** psychometric properties have been chosen from the literature and those properties are **reported**.	Pre-existing measures were chosen but **no** psychometric properties presented.
Measure Development	Rigorous measure development process is detailed, and validity and reliability **evidence is reported.**	Rigorous measure development process is detailed, but **no evidence** of psychometric properties.	**Unclear** how the measure was developed.	
Use of Direct Measures	**All** outcomes are assessed with direct measures of the outcome (measures reflect what students know, value, and can do).	**Most** outcomes are assessed with direct measures of the outcome.	**Some** outcomes are assessed with direct measures of the outcome.	**No** outcomes are assessed with direct measures of the outcome (e.g., satisfaction or attendance used to infer learning or development).
Measures-to-Outcomes Map	**All** outcomes are mapped to measures that represent the outcome and the match between the two is explained in detail (e.g., reviewed several existing measures and the selected measure aligns best with outcome; designed a measure and had others engage in backwards translation to confirm match between measure & outcome).	**Most** outcomes are mapped to measures with their match being explained in detail; for other outcomes, the details are brief or vague regarding the match.	**Some** outcomes are mapped to measures with their match being explained in detail; for other outcomes, the details are brief or vague regarding the match.	**Superficial** match between measures and outcomes.
Equity Centered	The measure, whether developed or selected, **produces** scores that allow for equally trustworthy inferences about the outcome, regardless of student population. **Multiple forms of evidence** are provided to support this claim. **Student feedback is solicited** regarding clarity of the measures used.	The measure, whether developed or selected, ***should*** **produce** scores that allow for equally trustworthy inferences about the outcome, regardless of student population. **Limited evidence** is available or provided to support this claim.	There is **no evidence** that the measure functions equivalently across different student populations and because of this, there is **language cautioning** the use of the scores to make inferences.	**No mention** of how the measure functions across different student populations.
Comments:				Average Score:

Okay I completed a round of assessment!

How do I evaluate my assessment practices?



3-Step Process

Articulating Program Theory

Implementation Fidelity

Describing IF and its need in outcomes assessment

Evidence Informed Practices

Offering programming that is effective

Repositories of Pre-existing Measures

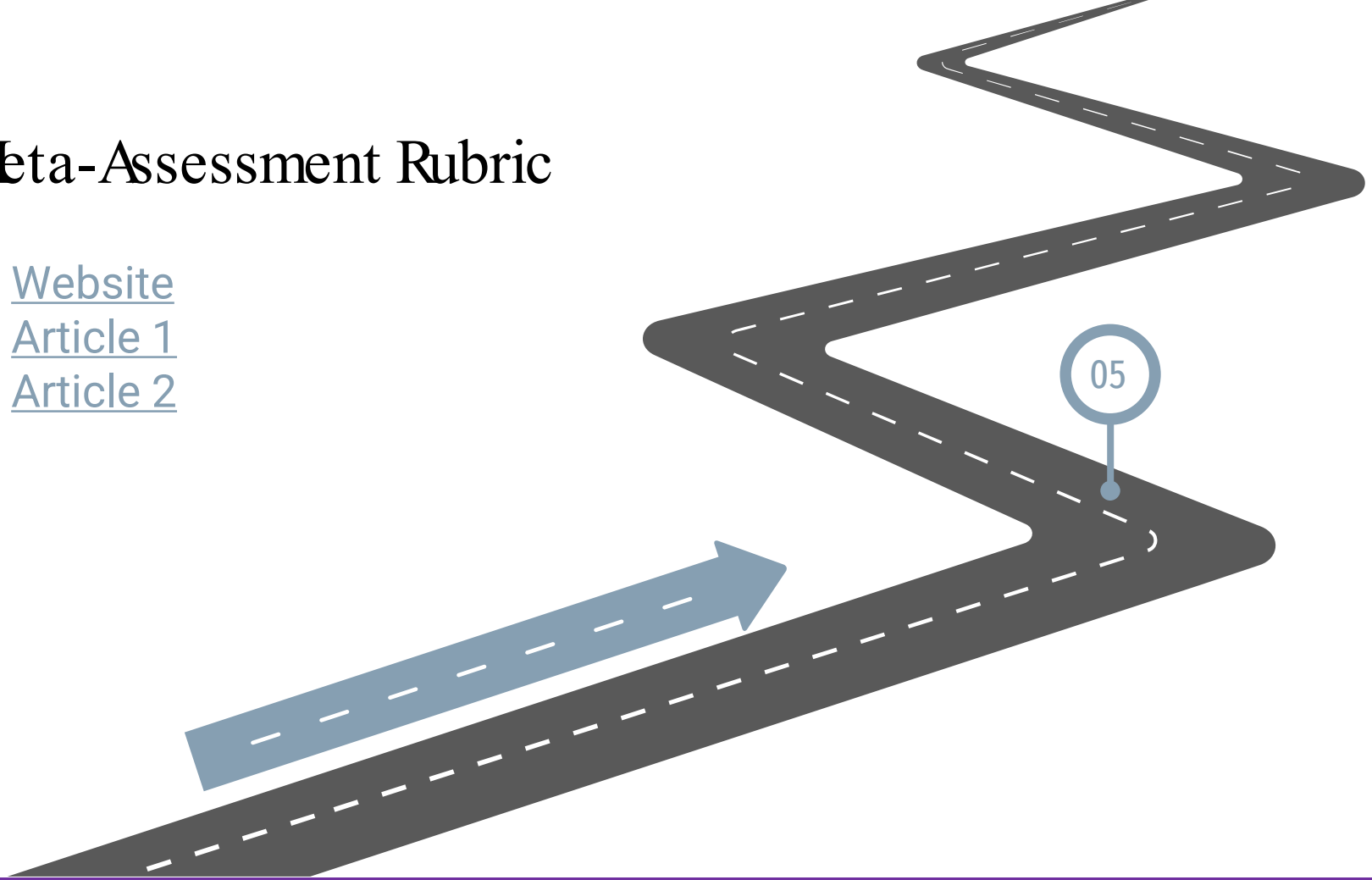
Expert review of measures, including recommendations on whether it should be used

Meta-Assessment Rubric

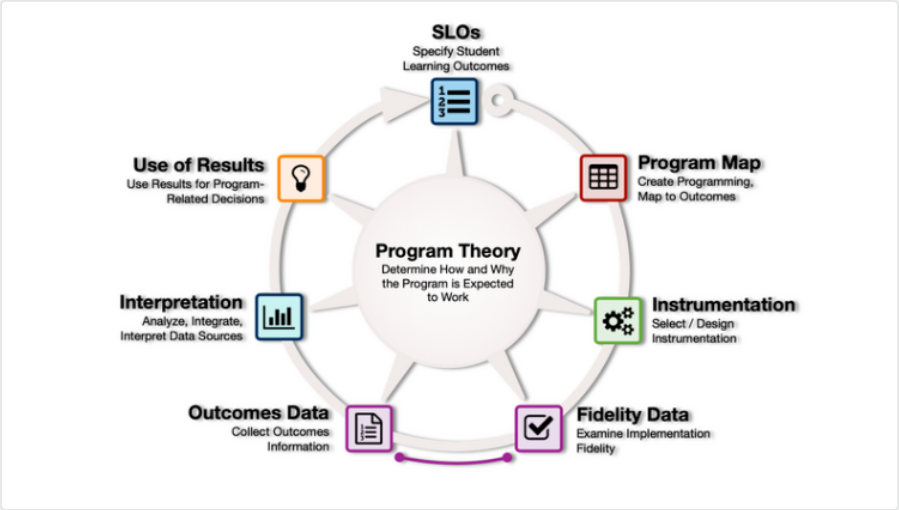
Makes the assessment process transparent to all involved

Meta-Assessment Rubric

- Website
- Article 1
- Article 2



Fostering Equity in Outcomes Assessment: A Rubric, Mock Reports, and Training Support Materials



Overview

In this collection, we are sharing a meta-assessment rubric, mock reports, and training support materials created by a team at James Madison University to foster equity in educational programming and outcomes assessment. The intention of sharing these materials under creative commons license is to provide other organizations or institutions a starting place to expose, model, and train others on equity considerations when designing and evaluating programming. This content was developed in partnership with the Center for Assessment and Research Studies and the Division of Student Affairs at James Madison University.

Overview of Resources

Student View Save Report

Details Resource Library

DOCX **Equity In Assessment Resource Document_Showcasing Examples of Levels of Equity**
Download

DOCX **JMU Assessment Improvement Rubric (Feb 2024)**
Download

DOCX **Mock Report_Advising Program**
Download

DOCX **Mock Report_Alcohol Program (PRIME for Life)**
Download

DOCX **Mock Report_Community Service Learning**
Download

Using Theory and Research to Create and/or Map Programming to the Outcomes	A synopsis of the specific theory and research used in the creation and mapping of each programming element is provided.			
Sub-Criteria	**Exemplary (3)**	**Proficient (2)**	**Developing (1)**	**Missing (0)**
Outcome-Programming Mapping	**All** program elements are mapped to their respective outcomes.	**Most** program elements are mapped to their respective outcomes.	**Some** program elements are mapped to their respective outcomes.	**No** program elements are mapped to respective outcomes or elements are **not** clearly identified.
Program Theory	A theoretical framework is identified, cited, and applied to **all** the elements of the program (mapping between programming and outcomes is **clear** and **justified**).	A theoretical framework is identified, cited, and applied to **most** of the elements of the program.	A theoretical framework is identified, cited, and applied to **some** of the elements of the program.	A theoretical framework is **not** identified, cited, and applied to any element of the program.
Evidence-Based Practice	Evidence-based practices have been articulated and provide rationale for **all** program elements.	Evidence-based practices have been articulated and provide rationale for **most** of program elements.	Evidence-based practices have been articulated and provide rationale for **some** of the program elements.	Evidence-based practices have **not** been articulated and **do not** provide rationale for any program elements.
Reasonable	Outcomes are **reasonable** given length and strength of the program and reasonableness is **extensively supported by evidence** from previous research.	Outcomes are **reasonable** given length and strength of the program and reasonableness is **minimally supported by evidence** from previous research.	It is **unclear** if the outcomes are reasonable given the information provided in the report.	Outcomes are **not** reasonable given the length and strength of the program.
Equity Centered	**Clear** articulation about if and how the etiology of outcomes differs across student populations. Intersectionality of students' identities is **considered** and **addressed**. **Student feedback is solicited** on the clarity of the applied program theory.	**Clear** articulation about if and how etiology differs across broad student populations.	**Unclear** or **limited** articulation around the differences in etiology across student populations. **Acknowledgement** that this assessment process may be the first step in identifying these differences.	**No** discussion of potential differences in etiology

Implementation Fidelity Data	A description of the alignment between the planned programming (e.g., curriculum, pedagogy, activities, strategies) and the implemented programming (i.e., the programming the student experienced).			
Sub-Criteria	**Exemplary (3)**	**Proficient (2)**	**Developing (1)**	**Missing (0)**
Inclusion of Core Aspects of Implementation Fidelity	The implementation fidelity checklist **includes and specifies *all*** core aspects: program differentiation, adherence, exposure, quality, and responsiveness.	The implementation fidelity checklist **includes but does *not* clearly specify** all core aspects: program differentiation, adherence, exposure, quality, and responsiveness.	The implementation fidelity checklist **includes some, but *not* all,** core aspects.	The implementation fidelity checklist **does *not* include any** core aspects of implementation fidelity.
Clarity of Description of Implementation Fidelity Methods	Provides **clear descriptions** for methods used to collect implementation fidelity (from who, with what, and when). Someone can **replicate** this data collection.	Provides **descriptions** for the methods used to collect implementation fidelity with some indication of who, what, or when.	Provides a **vague/unclear description** for the methods used to collect implementation fidelity data. **Cannot be replicated.**	Provides **no descriptions** for the methods used to collect implementation fidelity data.
Description of Implementation Fidelity Results	For **all** outcomes, clearly specifies if all programming was implemented as planned, most programming was implemented as planned (noting which programming was not implemented as planned), some programming was implemented as planned (noting which programming), or no programming was implemented as planned.	For **some** outcomes, clearly specifies if all programming was implemented as planned, most programming was implemented as planned (noting which programming was not implemented as planned), some programming was implemented as planned (noting which programming), or no programming was implemented as planned.	**General statements** across outcomes regarding the implementation of programming.	**No discussion** of the implementation fidelity results.
Use of Implementation Fidelity Data	Implementation fidelity data **accurately** informs the evaluation of the programming's effectiveness **and recommendations** for its improvement.	Implementation fidelity data **accurately** informs the evaluation of the programming's effectiveness.	There was an attempt to use implementation fidelity data to inform the evaluation of the programming's effectiveness, but it was **not accurately and/or clearly** interpreted.	Implementation fidelity data was **not** used to inform the evaluation of the programming's effectiveness.
Equity Centered	Examined whether program components were implemented with equal fidelity across diverse student populations and findings are **clearly articulated** in the report. Plans to address implementation discrepancies are **included, if applicable**. Students are **involved** in gathering or interpreting implementation fidelity data.	Examined whether the program components were implemented with equal fidelity across diverse student populations and findings are **clearly articulated** in the report.	**Crude** data collection is used to make claims about whether the program components were implemented with equal fidelity across diverse student populations.	**No mention** of whether the program components were implemented with equal fidelity across diverse student populations.

Abstract

Meta-assessment, or the assessment of the outcomes assessment process, is a useful strategy to communicate, guide, document, and provide feedback on assessment practices. We share the creation of a rubric that aligns with assessment processes aimed to improve student learning and development in higher education. More specifically, the rubric was created to align with the student affairs professional standards and explicitly evaluates equity-related aspects of each component of the outcomes assessment process. Additionally, we highlight the rubric as a necessary step in a broader change management effort. We then share procedures of the initial use of the rubric to evaluate assessment reports. After analyzing rubric scores (i.e., G-study) and qualitative feedback from several raters, we then describe additional support resources intentionally created to facilitate more reliable ratings. We freely share the rubric, training reports, and support materials to enable a culture of improvement in higher education institutions.

A Meta-Assessment Rubric to Guide Professional Development and Practice in Equitable Outcomes Assessment

AUTHORS

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

Jonathan P. Stewart, Ph.D.
James Madison University

Autumn N. Wild, M.A.
James Madison University

Riley K. Herr, M.A.
James Madison University

Katarina E. Schaefer, M.A.
James Madison University

JSAIII

Journal of Student Affairs Inquiry, Improvement, and Impact | Volume 7, Issue 1

Professional Development in Equity-Centered Outcomes Assessment

The Utility of a Meta-Assessment Rubric

Autumn N. Wild
James Madison University
Doctoral Assistant,
Center for Assessment & Research Studies (CARS)
autumn.wild.1@gmail.com | LinkedIn

Riley K. Herr
James Madison University
Doctoral Assistant, CARS
rileykherr@gmail.com | LinkedIn

Sara J. Finney, Ph.D.
James Madison University
Associate Director, CARS
Professor of Graduate Psychology
finneysj@jmu.edu | LinkedIn

Abstract: Core skillsets for student affairs educators have been articulated in several professional standards. However, the expected assessment skillset is not always addressed in graduate programs; in turn, many student affairs educators feel unprepared to engage in outcomes assessment. Our study showcases the utility of a new equity-centered meta-assessment rubric to provide needed assessment training to student affairs educators. Results support the use of this rubric to advance equitable assessment and programming on college campuses.

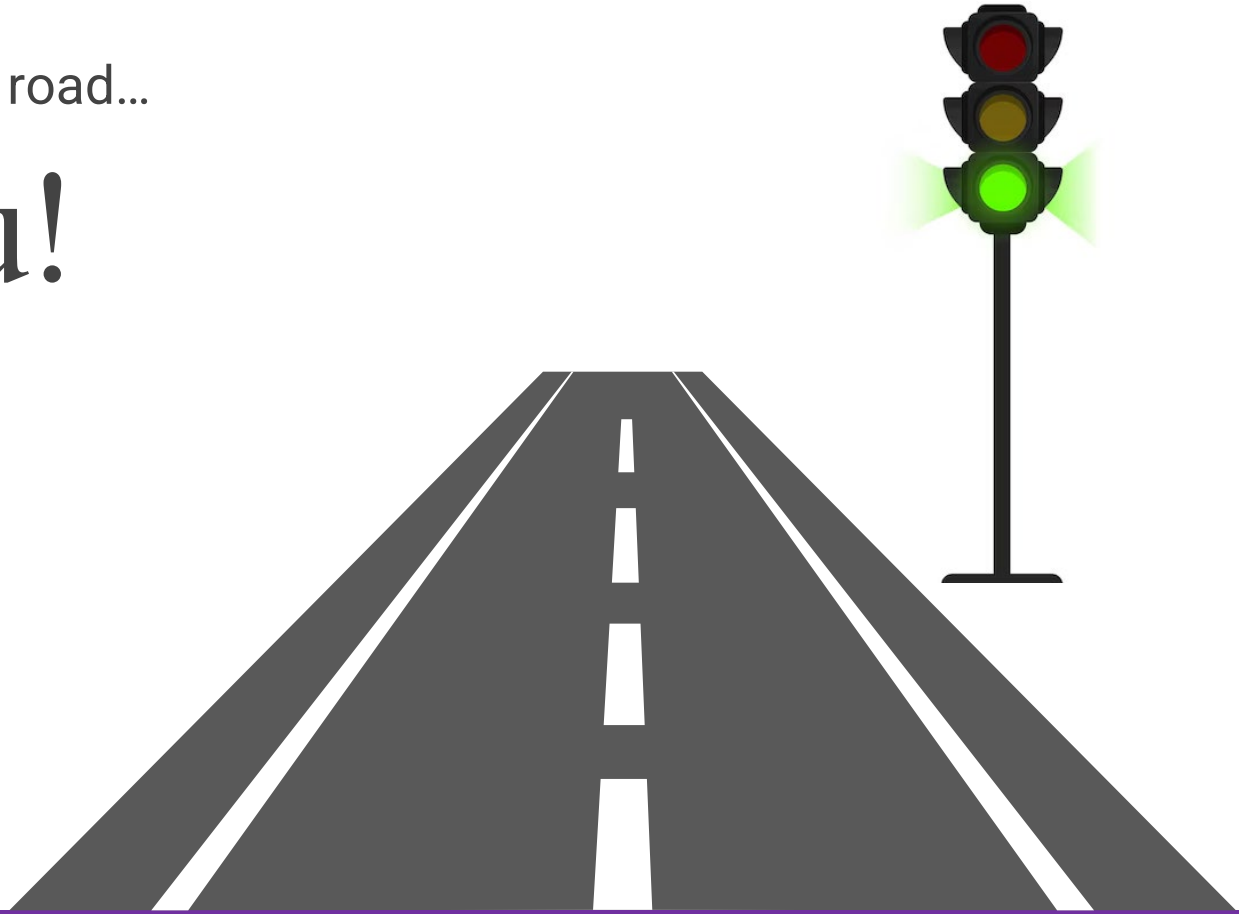
Keywords: outcomes assessment, equity in assessment, professional development in assessment

Now you're all set to hit the road…

Thank you!

Contact us:
mcfad2me@jmu.edu
finneysj@jmu.edu
willsejg@jmu.edu
stewarjp@jmu.edu

Additional Resources:
JMU SASS Website



Resource Links!