Curated Resources for High-Quality Assessment Practice

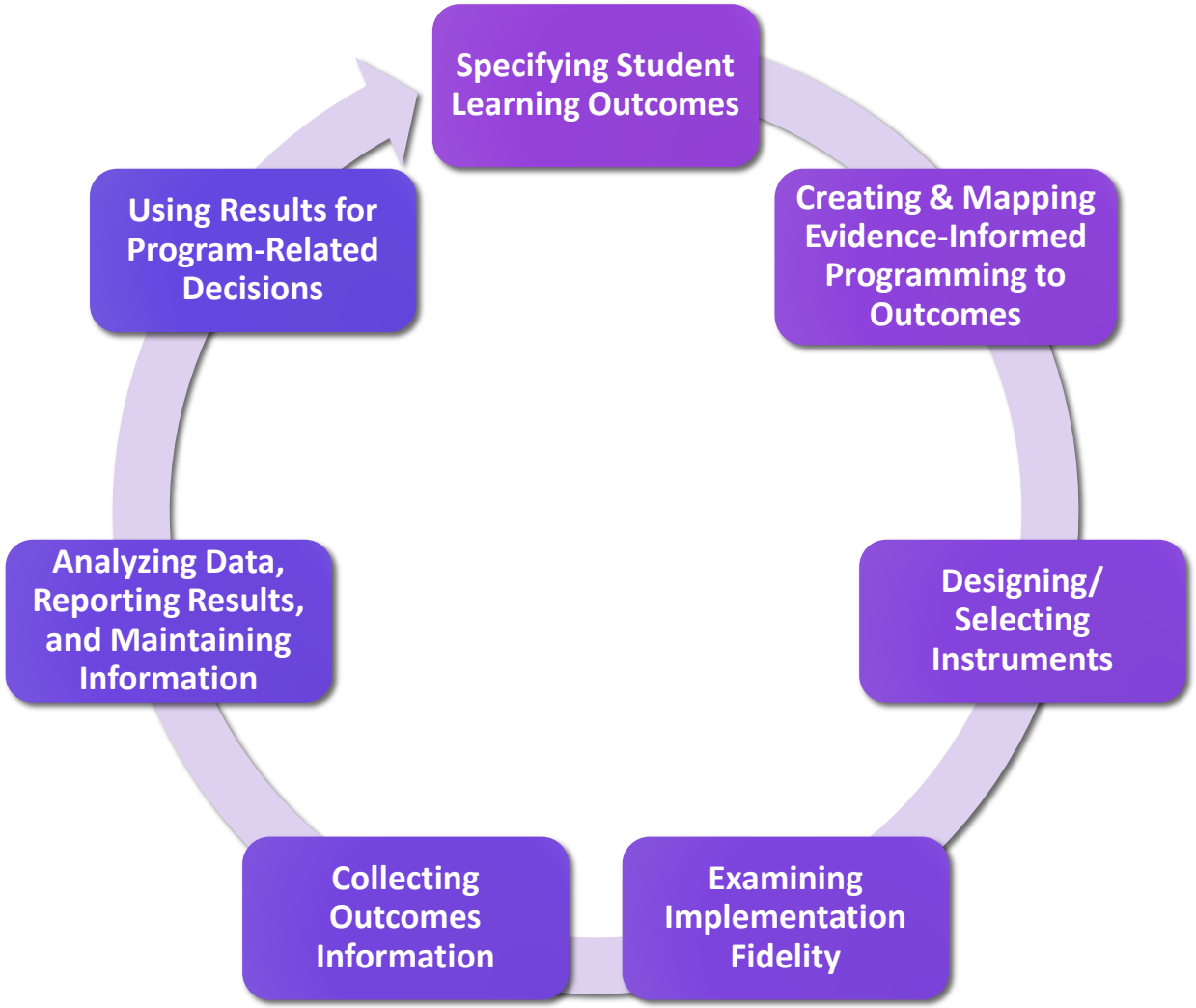
SARA FINNEY, PHD

SARAH ALAHMADI, MS

HOLLY BUCHANAN, DNP, MS, ANP-BC

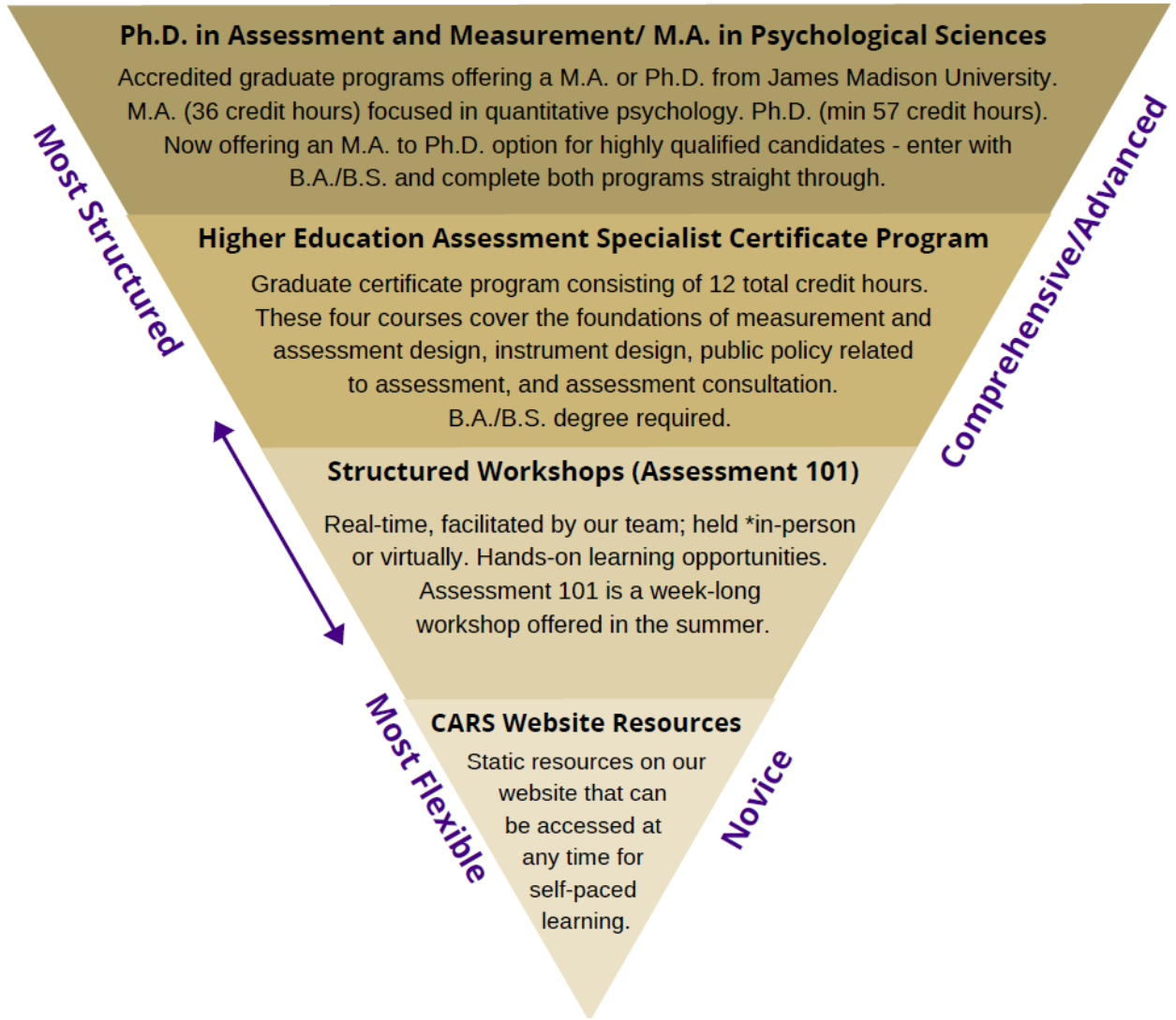
CHRIS PATTERSON, MA

The Outcomes Assessment Cycle



Professional Development in Assessment at James Madison University

Meeting your professional development needs in assessment: from beginner workshops to a PhD and everything in between.



New Website of Organized Resources!

https://www.jmu.edu/assessment/pdia/index.shtml

This webpage is a resource for those who would like to *learn about or teach others* about assessment. The page is structured using a taxonomy of assessement knowledge, skills, and attitudes described in the *Assessment Skills Framework*. Click on any of the Skill Areas below for professional development resources on the topic.

Link to 101

Building Blocks

Foundational Assessment Knowledge and Skills

Skill Area 1

Specify Student Learning Outcomes

Skill Area 2

Create and Map Programming to Outcomes

Link to Certificate

Skill Area 3

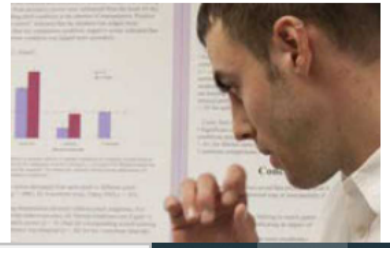
Select and Design Instruments

Skill Area 4

Examine Implementation Fidelity

Skill Area 5

Collect Outcomes Information



Skill Area 6

Analyze Data, Interpret and Report Results

Skill Area 7

Use Results to Improve Student Learning

Skill Area 8

Assessment in Practice - Additional Skills

Skill Area 1: Student Learning Outcomes

Expand All

Specify Student Learning Outcomes

Developing student learning outcomes

Novice

Assessment Quickies - Student Learning Outcomes

Assessment Quickies presents a series of short videos on how to assess student learning. The following two videos introduce us to the concepts and development of Student Learning Outcomes.

Step 1: What Are Student Learning Outcomes?

Step 2: Writing Student Learning Outcomes

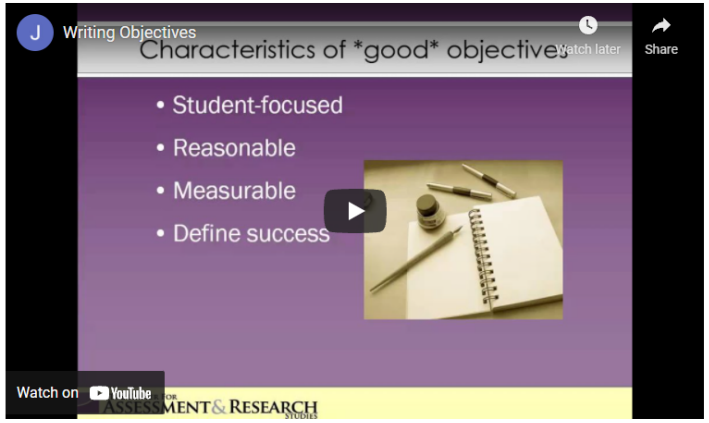
Revised Bloom's Taxonomy

Bloom's Taxonomy is a quintessential tool in writing measurable SLOs. This resource from Iowa State presents the taxonomy alongside the knowledge dimensions professionals are typically interested in assessing to facilitate better verb choice at various levels.

SASS - Writing Student Learning Objectives Introduction

Student Affairs Assessment Suport Services (SASS) at JMU present an introductory video and online guide to writing student learning objectives, detailing what characterizes a "good" objective and some go-to dos and don'ts for the process. Examples are tailored towards student affairs, but tips are general enough for all types of assessment professionals!

1. Online Guide
2. Video



CARS Consultation Materials: Objectives Made Easy

This quick handout reinforces best practices for writing objectives for the learner, with a great outline of potential verb words at different levels and domains one might be interested in assessing.

Best/Worst Practice Videos:

1. The "Weasel Word" in SLOs

Intermediate

Assessment Quickies: Levels of Student Learning

Step 3: Levels of Student Learning

This resource is a continuation of the 10 step Assessment Quickies series, which dives deeply into the levels of student learning. In conjunction with this video would be a good time to revisit Bloom's Taxonomy (https://www.celt.iastate.edu/teaching/effective-teaching-practices/revised-blooms-taxonomy/)

Quick Overview Handouts:

1. Belmont University - Correcting SLOs

This short resource explains best practices for writing SLOs with clear examples of SLOs that violate these practices and how to correct them.

2. UWA - SLO Highlights

University of Western Australia presents a 1-pager style fact sheet, introducing SLOs, highlighting taxonomies and outlining other important considerations for writing outcomes.

SASS - SLO Writing Workshop Slides

The attached slides from a previous JMU Workshop reiterate best practices for writing objectives with more examples, summary checklists of considerations, and three clearly presented common misconceptions or resistant themes that may arise when trying to explain SLOs with clarifications that address each. Handout versions of some of the resources within the PowerPoint are presented below for convenience.

1. *Checklist for Effective Outcomes*

2. *Dos and Don'ts*

Exemplars:

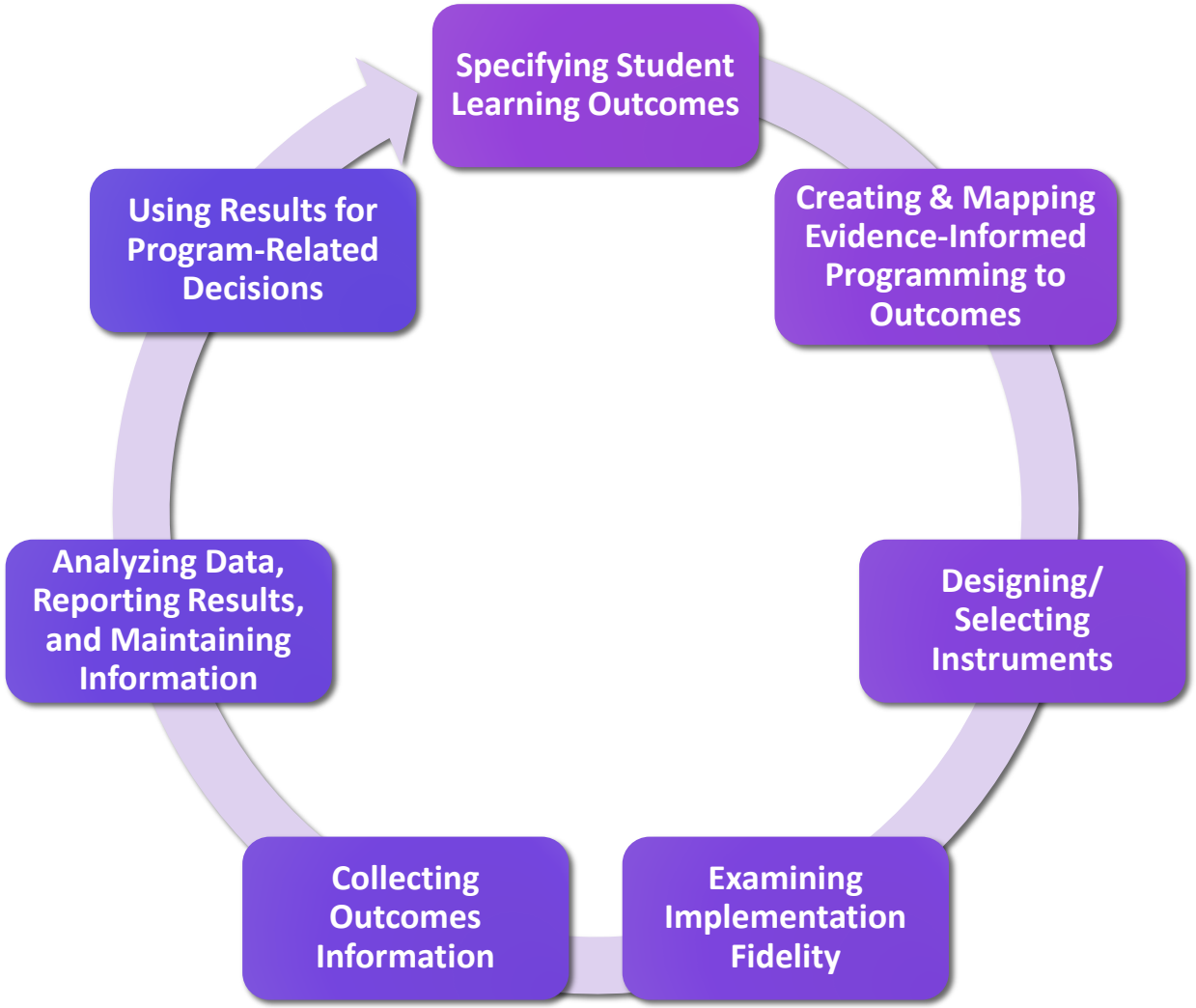
1. Las Positas Community College - SLO Handbook

Las Positas Community College provides an in-depth example of how one institution handles SLO development. Several SLO examples are provided alongside best practice information with alignment to larger institutional objectives.

2. UF - Developing Program Goals & SLOs

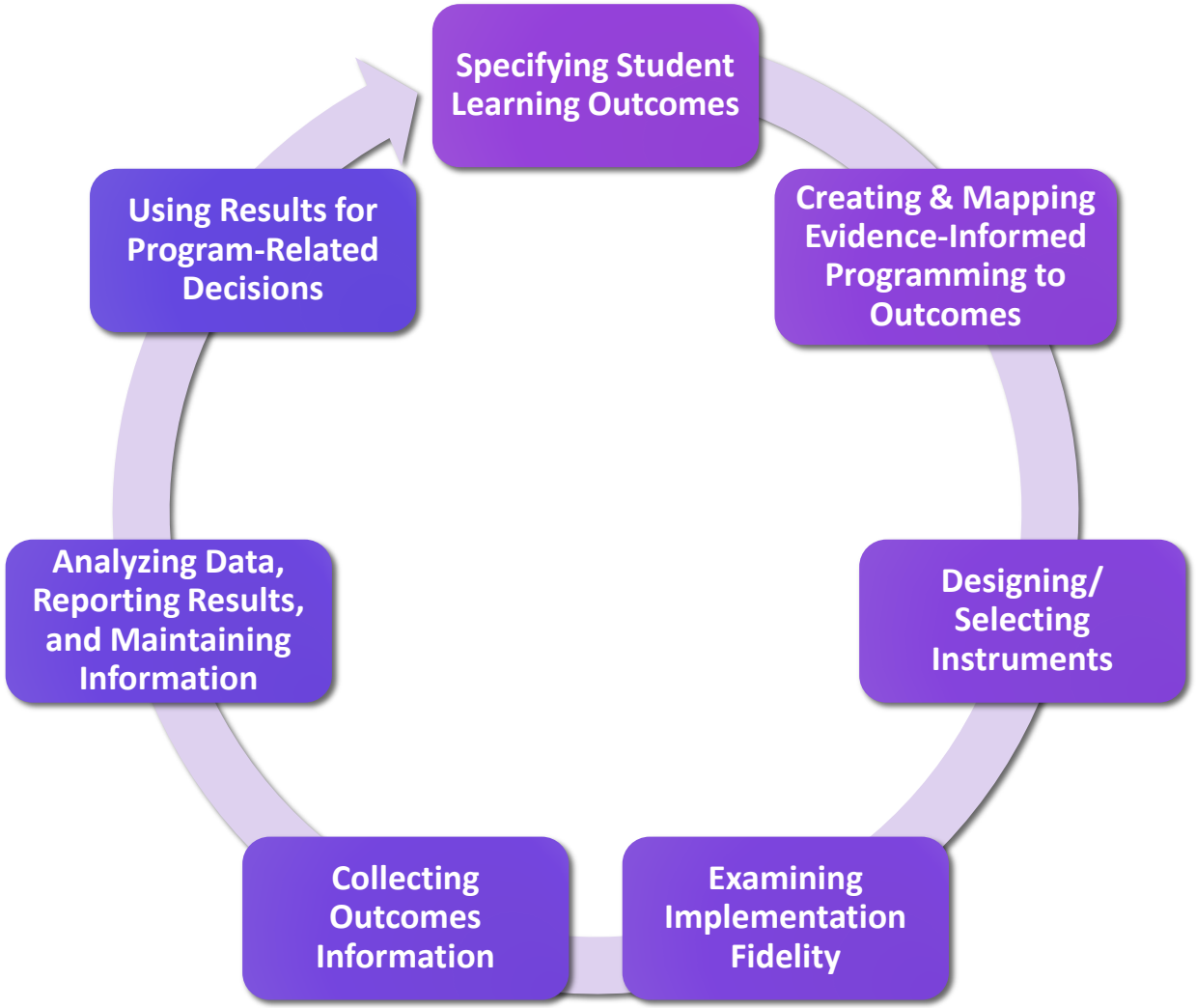
This guide from the University of Florida illustrates one university's approach to developing program goals and student learning outcomes. The guide highlights the "SMART" approach to program goals and the application of Bloom's Taxonomy for SLOs with plenty of verbs to choose

The Outcomes Assessment Cycle



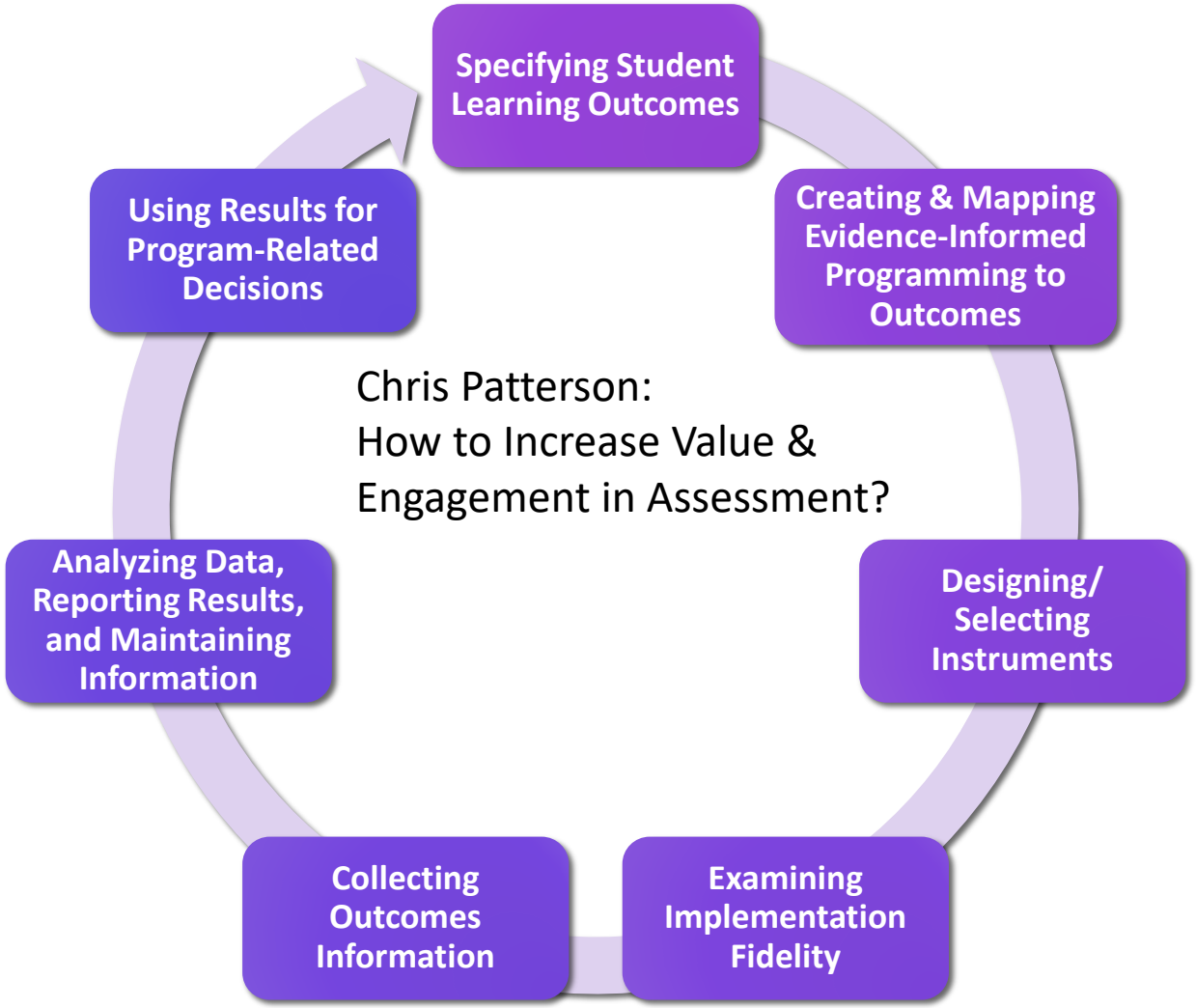
Holly Buchanan:
Where to Find Evidence-Informed Programming?

The Outcomes Assessment Cycle



Sarah Alahmadi:
Where to Find High-Quality Outcome Measures?

The Outcomes Assessment Cycle

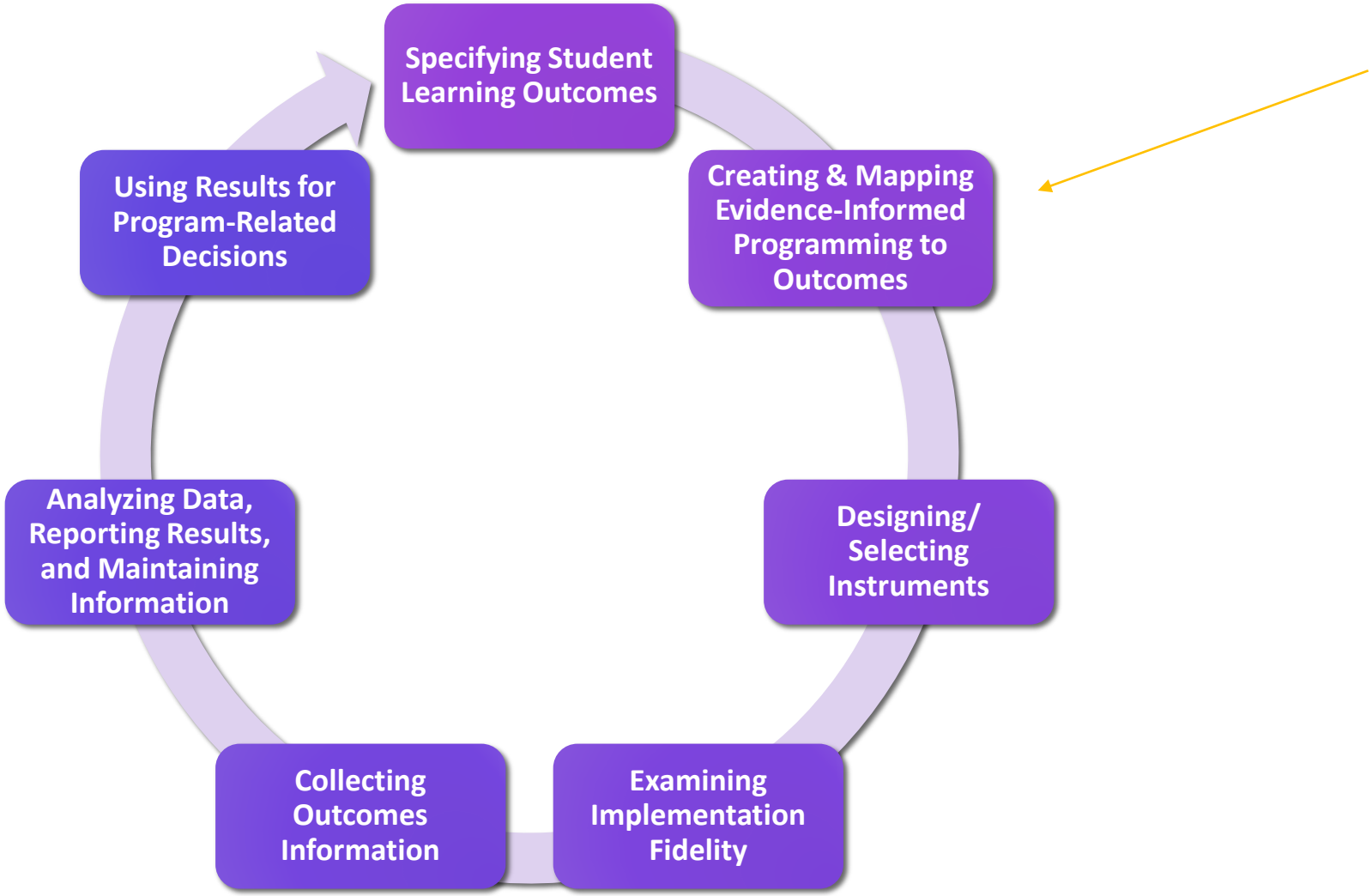


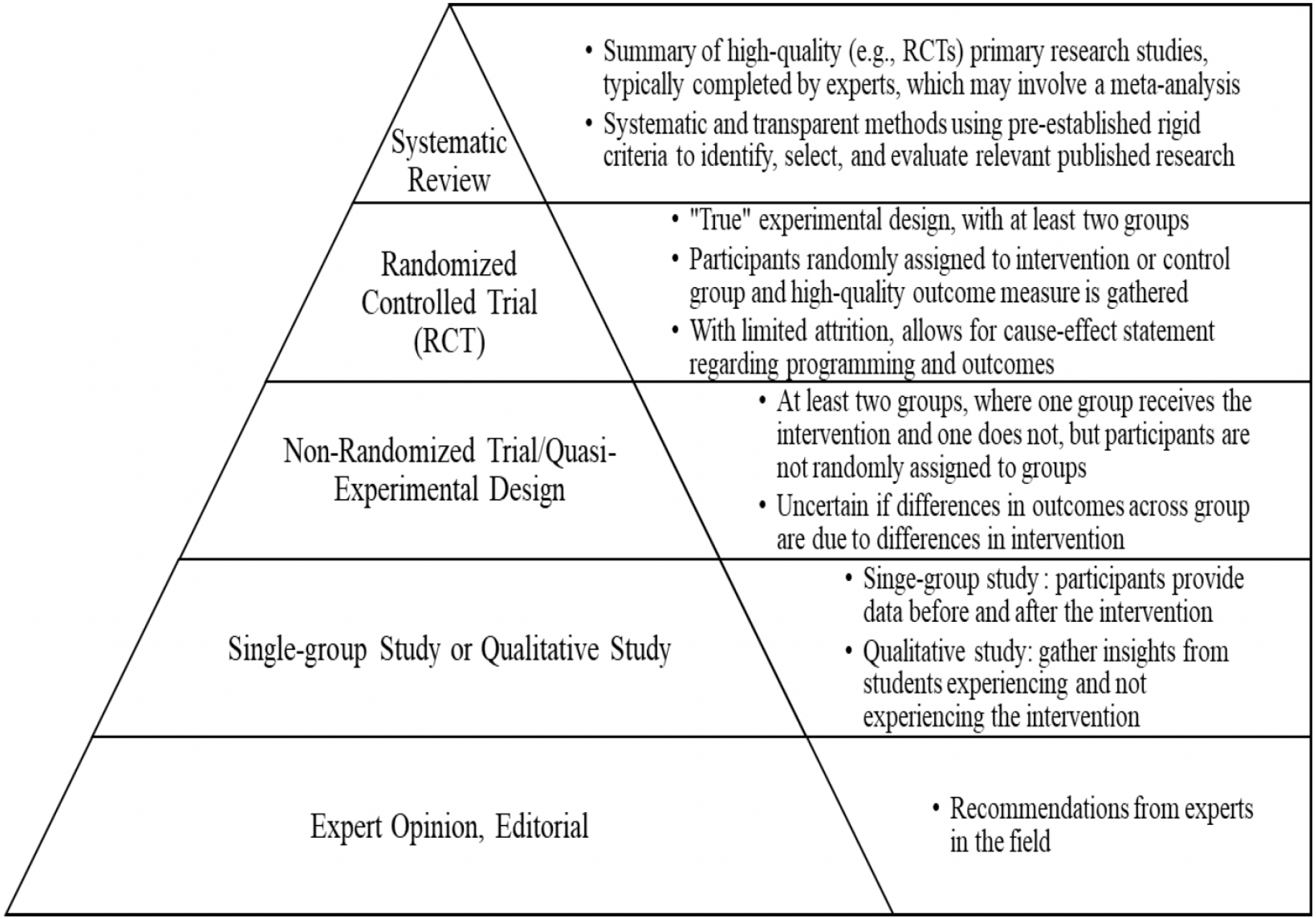
JMU The Center for Assessment and Research Studies

Systematic Review Repositories

HOLLY A BUCHANAN, DNP, MS, ANP-BC

The Outcomes Assessment Cycle





What Are Systematic Reviews?

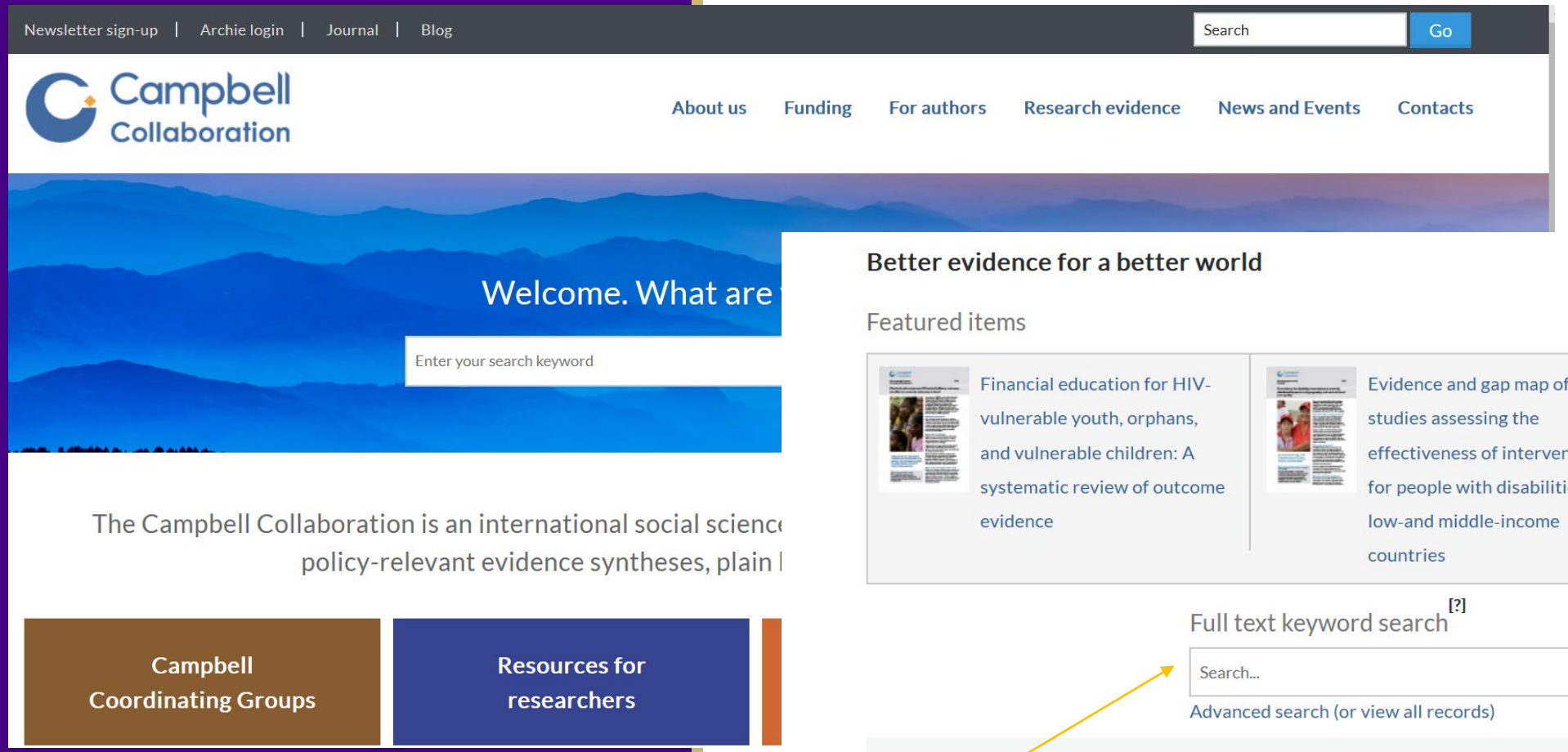
https://www.campbellcollaboration.org/explore/what-is-a-systematic-review.html

Database	Description	Examples
Campbell Collaboration	Campbell 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 QEDs) primary research.	• 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 Intimate Partner Abuse
What Works Clearinghouse	WWC is a trusted source of scientific evidence on **education programs, practices, & policies**. We review research, determine which studies meet rigorous standards (RCTs, QED), & summarize findings & provide practice guides.	• Using Technology To Support Postsecondary Student Learning • Linked Learning Communities • Organizing Instruction & Study to Improve Student Learning • First Year Experience Courses • Strategies for Postsecondary Students in Developmental Education • Teaching Students to Write Effectively Summary
Cochrane Library	Provides plain language summaries of systematic reviews focusing on interventions for a variety of **health outcomes** (e.g., mental health, alcohol, STDs). Indicate 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 • Interventions Intended to Prevent HPV
What Works for Health	Analysts review & assess research to rate effectiveness of strategies that affect **health**. Ratings assigned based on 2 analysts' assessments of strength of body of evidence (e.g., quality, # of studies, consistency of findings) as it pertains to specified outcomes. Place most weight on studies with designs that demonstrate causality (e.g., RCTs).	• Health Career Recruitment for Minority Students • Cultural Competence Training for Health Care Professionals • Condom Availability Programs • Summer Learning Programs • Technology-Enhanced Classroom Instruction • Outdoor Experiential Education & Wilderness Therapy • College-based Obesity Prevention Educational Intervention • Alcohol Brief Intervention

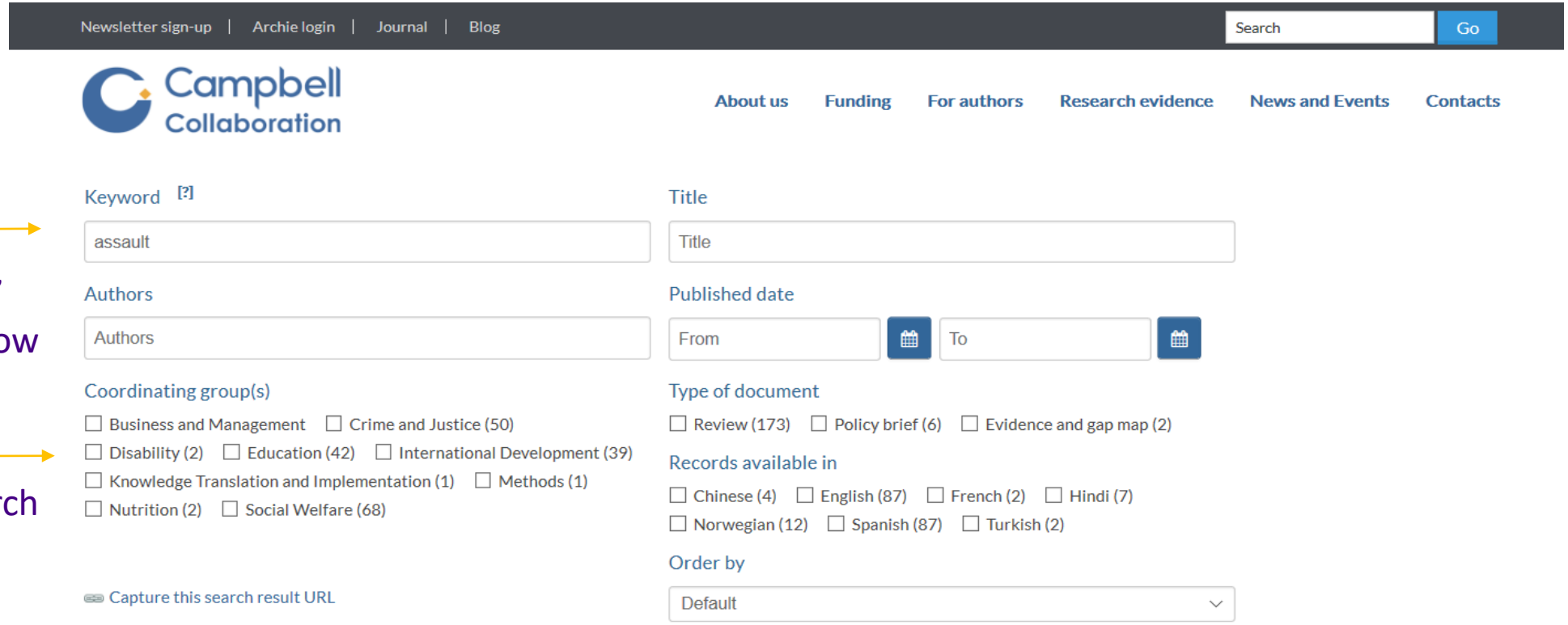
Campbell Collaboration Reviews

Provide answers for decision makers by using **rigorous** methods to synthesize evidence, including, where appropriate, statistical meta-analysis of quantitative evidence and theory-based analysis of qualitative evidence

- include a systematic search for unpublished reports, to avoid publication bias
- are usually international in scope
- protocol for the review is developed in advance & also undergoes peer review
- study inclusion & coding decisions are carried out by at least 2 reviewers who work independently & compare results
- study quality is appraised
- undergo peer review & editorial review



- Can look at all reviews in a type by selecting that type (e.g., education)
- Or can conduct a search.

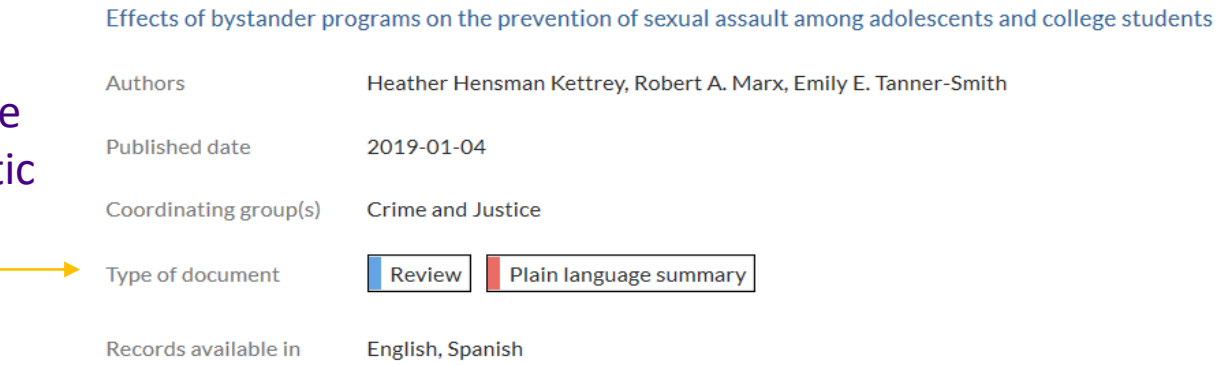


- Searched “assault”
- 8 results listed below

Could further limit search by selecting a specific “group”

Search Result: 8 Records Found

There is a Plain Language Summary in addition to the full details of the Systematic Review Process



2-page **Plain Language Summary** appears first (and PDF can be downloaded & shared).

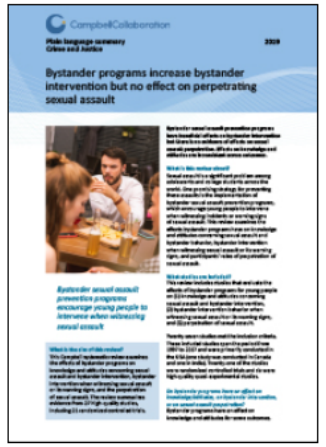
40 page **Full Report** can also be examined.
Full report contains details regarding:

- the studies gathered the quality of the studies
- the high-quality studies that were synthesized (RCTs, quasi-experimental designs with particular characteristics) vs. the low-quality studies that weren't included in synthesis
- the analyses.

Effects of bystander programs on the prevention of sexual assault among adolescents and college students

Authors: Heather Hensman Kettrey, Robert A. Marx, Emily E. Tanner-Smith
Published date: 2019-01-04
Coordinating group(s): Crime and Justice
Records available in: English, Spanish
Select language:

English



Download PDF

See the full report

Share Tweet Share

PLAIN LANGUAGE SUMMARY

Bystander programs increase bystander intervention but no effect on perpetrating sexual assault

Bystander sexual assault prevention programs have beneficial effects on bystander intervention but there is no evidence of effects on sexual assault perpetration. Effects on knowledge and attitudes are inconsistent across outcomes.

What is this review about?

Sexual assault is a significant problem among adolescents and college students across the world. One promising strategy for preventing these assaults is the implementation of bystander sexual assault prevention programs, which encourage young people to intervene when witnessing incidents or warning signs of sexual assault. This review examines the effects bystander programs have on knowledge and attitudes concerning sexual assault and bystander behavior, bystander intervention when witnessing sexual assault or its warning signs, and participants' rates of perpetration of sexual assault.

What is the aim of this review?

This Campbell systematic review examines the effects of bystander programs on knowledge and attitudes concerning sexual assault and bystander intervention, bystander intervention when witnessing sexual assault or its warning signs, and the perpetration of sexual assault. The review summarizes evidence from 27 high-quality studies, including 21 randomized controlled trials.

What studies are included?

This review includes studies that evaluate the effects of bystander programs for young people on (a) knowledge and attitudes

Types of interventions

Eligible intervention programs were those that approached participants as allies in preventing and/or alleviating sexual assault among adolescents and/or college students. Some part of the program had to focus on ways that cultivate willingness for a person to respond to others who are at risk for sexual assault. All delivery formats were eligible for inclusion (e.g., in-person training sessions, video programs, web-based training, advertising/poster campaigns). There were no intervention duration criteria for inclusion.

Eligible comparison groups must have received no intervention services targeting bystander attitudes/behavior or sexual assault.

Types of outcome measures

We included studies that measured the effects of bystander programs on at least one of the following primary outcome domains:

- General attitudes toward sexual assault and victims (e.g., victim empathy, rape myth acceptance).
- Prerequisite skills and knowledge for bystander intervention as defined by Burn (2009) (e.g., noticing sexual assault or its warning signs, identifying a situation as appropriate for intervention, taking responsibility for acting/intervening, knowing strategies for helping/intervening).
- Self-efficacy with regard to bystander intervention (e.g., respondents' confidence in their ability to intervene).
- Intentions to intervene when witnessing instances or warning signs of sexual assault.
- Actual intervention behavior when witnessing instances or warning signs of sexual assault.
- Perpetration of sexual assault (i.e., participants' rates of perpetration).

Duration of follow-up

Studies reporting follow-ups of any duration were eligible for inclusion. When studies reported outcomes at more than one follow-up wave, each wave was coded and identified by its reported duration. Follow-ups of similar durations were analyzed together.

Types of settings

The review focused on studies that examined outcomes of bystander programs that target sexual assault and were implemented with

Database	Description	Examples
Campbell Collaboration	Campbell 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 primary research.	• 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 Intimate Partner Abuse
What Works Clearinghouse	WWC is a trusted source of scientific evidence on **education programs, practices, & policies**. We review research, determine which studies meet rigorous standards, & summarize findings & provide practice guides.	• Using Technology To Support Postsecondary Student Learning • Linked Learning Communities • Organizing Instruction & Study to Improve Student Learning • First Year Experience Courses • Strategies for Postsecondary Students in Developmental Education • Teaching Students to Write Effectively Summary
Cochrane Library	Provides plain language summaries of systematic reviews focusing on interventions for a variety of **health outcomes** (e.g., mental health, alcohol, STDs). Indicate 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 • Interventions Intended to Prevent HPV
What Works for Health	Analysts review & assess research to rate effectiveness of strategies that affect **health**. Ratings assigned based on 2 analysts' assessments of strength of body of evidence (e.g., quality, # of studies, consistency of findings) as it pertains to specified outcomes. Place most weight on studies with designs that demonstrate causality.	• Health Career Recruitment for Minority Students • Cultural Competence Training for Health Care Professionals • Condom Availability Programs • Summer Learning Programs • Technology-Enhanced Classroom Instruction • Outdoor Experiential Education & Wilderness Therapy • College-based Obesity Prevention Educational Intervention • Alcohol Brief Intervention

What Works Clearinghouse

The *What Works Clearinghouse* (WWC) reviews the existing research on different programs, products, practices, & policies in education.

"Our goal is to provide educators with the information they need to make evidence-based decisions. We focus on the results from **high-quality research** to answer the question '*What works in education*?'"

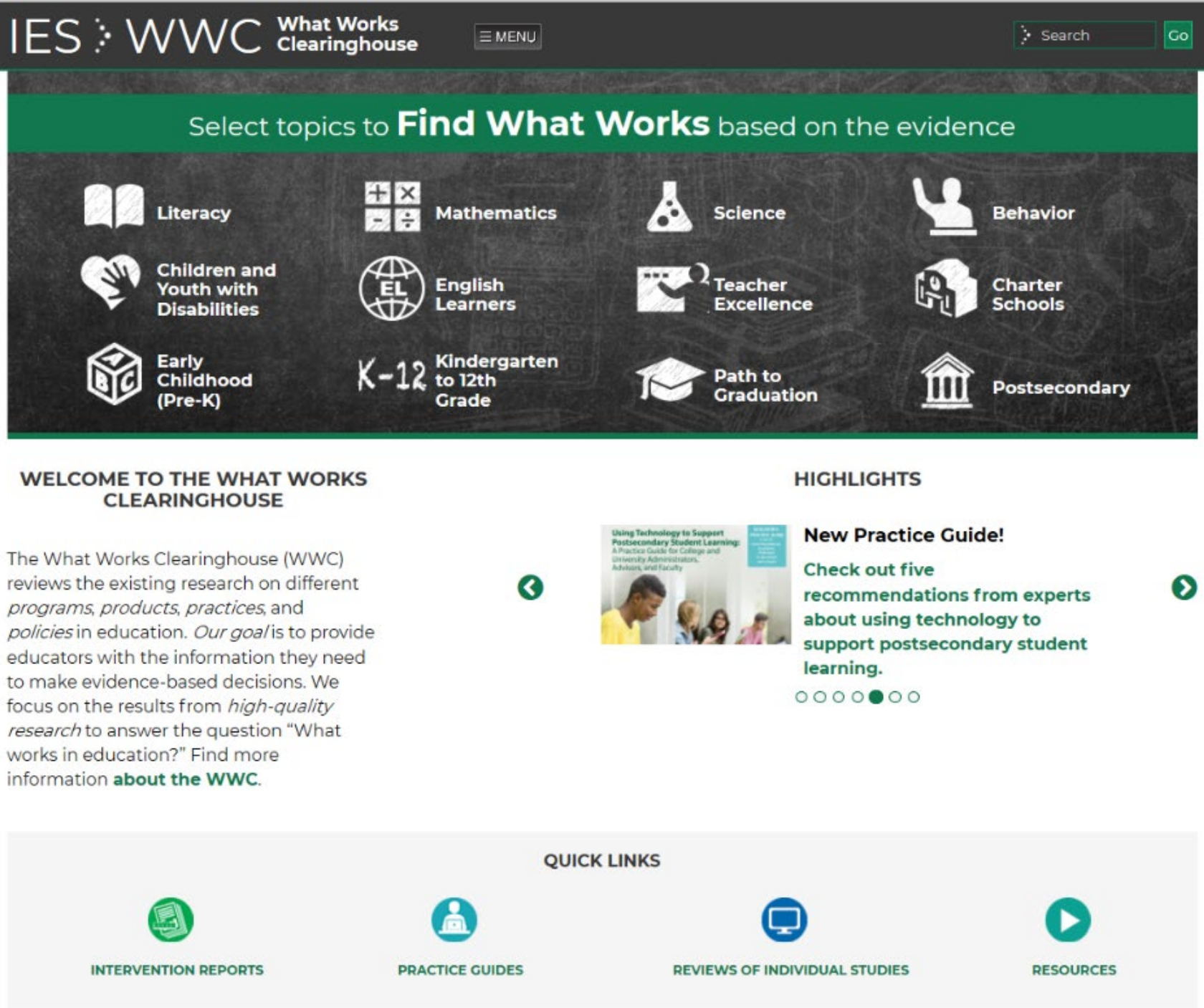
- Includes **Systematic Reviews & Practice Guides**

WWC: Practice Guides

Practice Guides give recommendations on strategies to use to meet outcomes. Each recommendation has a rating of *level of evidence* (**minimal, moderate, strong**). The level of evidence reflects how confident WWC certified reviewers & a panel of experts are that recommended practice consistently improved outcomes.

Rating of Level of Evidence: informed by the number of studies supporting the recommended practices, whether studies used a RCT or quasi-experimental design, & whether the study was examined in different contexts with different populations.

- Not uncommon for a practice to get a "minimal level of evidence" rating given these criteria.
- A minimal level of evidence does NOT mean there is no evidence supporting the recommendation.
- All recommendations are supported by at least some rigorous research.
- Minimal rating means more research is needed (examined in more settings, with more students).



- *Extremely* helpful for identifying effective evidence-informed practices
- Materials explicate *program theory, previous evidence*, & *exact intervention* (along with implementation fidelity information)
- Reviewers provide ratings of the quality of evidence & claims

IES WWC What Works Clearinghouse ≡ MENU Search Go

Find What Works based on the evidence

Print

Filter by topic

- Literacy
- Mathematics
- Science
- Behavior
- Children and Youth with Disabilities
- English Learners
- Teacher Excellence
- Charter Schools
- Early Childhood (Pre-K)
- ☐ K-12 Kindergarten to 12th Grade
- ☐ Path to Graduation
- ☑ Postsecondary

MORE FILTERS

You may also be interested in

13 Results filtered by:

Postsecondary ×

Evidence of effectiveness ▼	Intervention	Grades examined	Compare
	Dual Enrollment Programs	9-12	☐
	Accelerated Study in Associate Programs (ASAP)	PS	☐
	Summer Counseling	12-PS	☐
	Facilitating Long-term Improvements in Graduation and Higher Education for Tomorrow (FLIGHT)	7-11	☐
	First year experience courses	PS	☐
	Open Learning Initiative (OLI)	PS	☐
	InsideTrack© Coaching	PS	☐
	Summer Bridge Programs	PS	☐
	Linked Learning Communities	PS	☐
	Developmental Summer Bridge Programs	PS	☐
	First Year Experience Courses for Students in Developmental Education	PS	☐
?	ACT Aspire™		☐
?	Residential Learning Communities		☐

IES WWC What Works Clearinghouse ☰ MENU Search Go

WWC SUMMARY OF EVIDENCE FOR THIS INTERVENTION

Export Print

First year experience courses

First year experience courses, often referred to as college success courses or freshman seminars, are courses for first-year students in 2-year and 4-year colleges. The general goals of first year experience courses are to support the academic performance, social development, persistence, and degree completion of college students. Additionally, first year experience courses often aim to increase students' sense of campus community and connection to their institutions, while giving students the opportunity to interact with faculty and peers.

Reviewed Research

Number of Studies that met WWC Design Standards and provide evidence of effectiveness. Selecting an item below will display all studies that met WWC design standards in the domain. Selecting a study citation will take you to more information on that study and its findings.

Supporting Postsecondary Success

July 2016

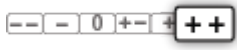
EVIDENCE SNAPSHOT INTERVENTION REPORT (814 KB) REVIEW PROTOCOL

Outcome domain	Effectiveness rating	Studies meeting standards	Grades examined	Students	Improvement index
Academic achievement	-- - 0 + ++	2 studies meet standards	PS	2,505	4 (-50 0 +50)
Attainment	-- - 0 + ++	1 study meets standards		8,290	9 (-50 0 +50)
Credit accumulation and persistence	-- - 0 + ++	3 studies meet standards	PS	11,455	9 (-50 0 +50)

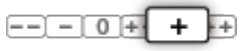
Outcome Domain: When examining details of each study, you have access to the specific outcomes: GPA, first-to-second year retention, degree attained at home institution, etc.

IMPROVEMENT INDEX: Indicator of the effect of the intervention interpreted as the expected change in percentile rank for an average comparison group student if that student had received the intervention.

WWC Effectiveness Rating

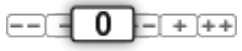


Positive: strong evidence that intervention had a positive effect on outcomes.

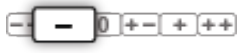


Potentially Positive: evidence that intervention had a positive effect on outcomes with no overriding contrary evidence.

Mixed: evidence that intervention's effect on outcomes is inconsistent.

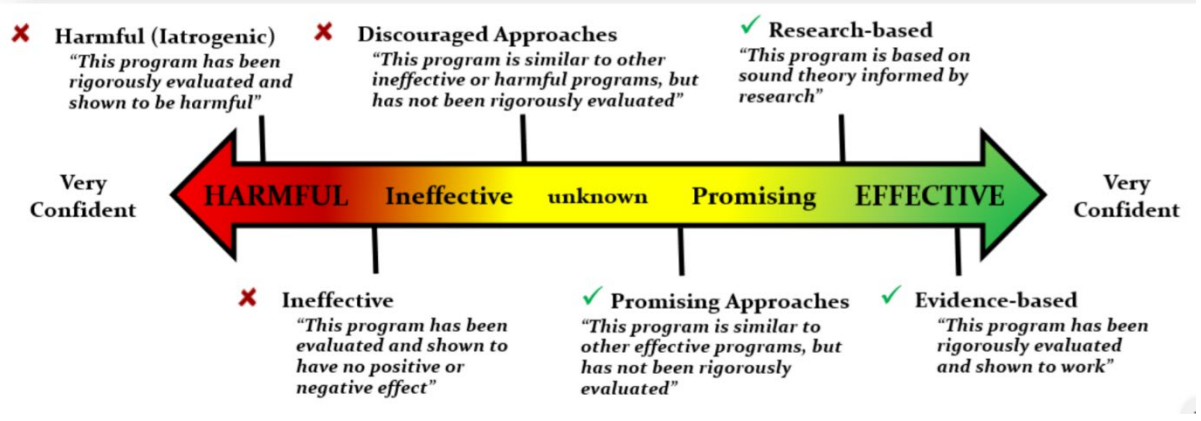


No Discernible: no evidence that intervention had an effect on outcomes.



Potentially Negative: evidence that intervention had a negative effect on outcomes with no overriding contrary evidence.

Negative: strong evidence that intervention had a negative effect on outcomes.



WWC Intervention Report

A summary of findings from a systematic review of the evidence

ies INSTITUTE OF EDUCATION SCIENCES

Supporting Postsecondary Success **July 2016***

First Year Experience Courses

Program Description[1]

Many students, even those who are academically well prepared, arrive to college without the strong study skills and coping strategies required to effectively navigate the learning and social environments in higher education. *First year experience courses*, often referred to as college success courses or freshman seminars, are courses for first-year students in 2-year and 4-year colleges. As reported by the National Resource Center for The First-Year Experience and Students in Transition, *first year experience courses* are required for all first-year students at approximately 52% of 4-year institutions.[2] Furthermore, 87% of 4-year institutions that offer freshman seminars do so for academic credit.[3] While they are not required for all students, *first year experience courses* are frequently required for special populations (e.g., academically underprepared students, students in specific majors, and students taking remedial courses). The general goals of *first year experience courses* are to support the academic performance, social development, persistence, and degree completion of college students.[4] Additionally, *first year experience courses* often aim to increase students' sense of campus community and connection to their institutions, while giving students the opportunity to interact with faculty and peers.[5] While courses vary in terms of content and focus, most *first year experience courses* are designed to introduce students to "campus resources, time management, study skills, career planning, cultural diversity, and student development issues."[6]

Research[7]

The What Works Clearinghouse (WWC) identified four studies of *first year experience courses* that both fall within

Report Contents

This intervention report presents findings from a systematic review of *first year experience courses* conducted using the WWC Procedures and Standards Handbook, version 3.0, and the Supporting Postsecondary Success review protocol, version 3.0.

Appendix A.1: Research details for Clouse (2012)

Clouse, W. A. (2012). *The effects of non-compulsory freshman seminar and core curriculum completion ratios on post-secondary persistence and baccalaureate degree attainment* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3523633)

Table A1. Summary of findings **Meets WWC group design standards with reservations**

Outcome domain	Sample size	Study findings: Average improvement index *(percentile points)*	Study findings: Statistically significant
Degree attainment (college)	8,290 students	+9	Yes
Credit accumulation	8,290 students	+8	Yes

Setting The study was conducted at a large public university in the western United States using archival data extracted from the student records central information warehouse. Data from all first-time, full-time incoming freshmen enrolled in the fall semester from 1995 to 2005 were eligible for analysis.

Study sample The author reported baseline characteristics on the full sample. The mean age was 18.69 (SD = 1.42). Of the full sample, 60% were male and 88% were White. The average amount of need-based aid received over the first year was $1,035.74 (SD = $2,471.44).

Intervention group Students in the intervention condition chose to enroll in a credit-bearing freshman seminar course designed for first-time, first-semester freshman college students. The *first year experience course* was designed to help freshmen at the university transition from high school to college and included topics in the following three areas: academic, personal, and community. Academic components included study, communication, and technology skills; personal components included a focus on developing personal goals and responsibilities; and community components focused on developing relationships and learning opportunities outside the classroom. Each section generally included 15 or fewer students and was taught by one of several faculty members and one junior teaching assistant.

Comparison group This study utilized a business-as-usual comparison condition. The comparison group consisted of those students who chose not to enroll in the freshman seminar. These students had access to typical college resources. Services received by the comparison group were not monitored.

Outcomes and measurement The author reported three dichotomous outcome variables: (1) Home degree conferred (whether a student received an undergraduate degree from their home institution within 6 years of entering); (2) External degree conferred (whether a student received an undergraduate

Additional Resource:

Systematic reviews have limitations…

- Time sensitive
- Time-consuming
- Requires high quality studies (RCTs and quasi-experimental designs)

Wise Interventions (https://www.wiseinterventions.org)…

- Curated resource by Greg Walton and colleagues
- Summarizes short, yet powerful, interventions to impact behavior, self control, health, belonging, achievement, among other outcomes

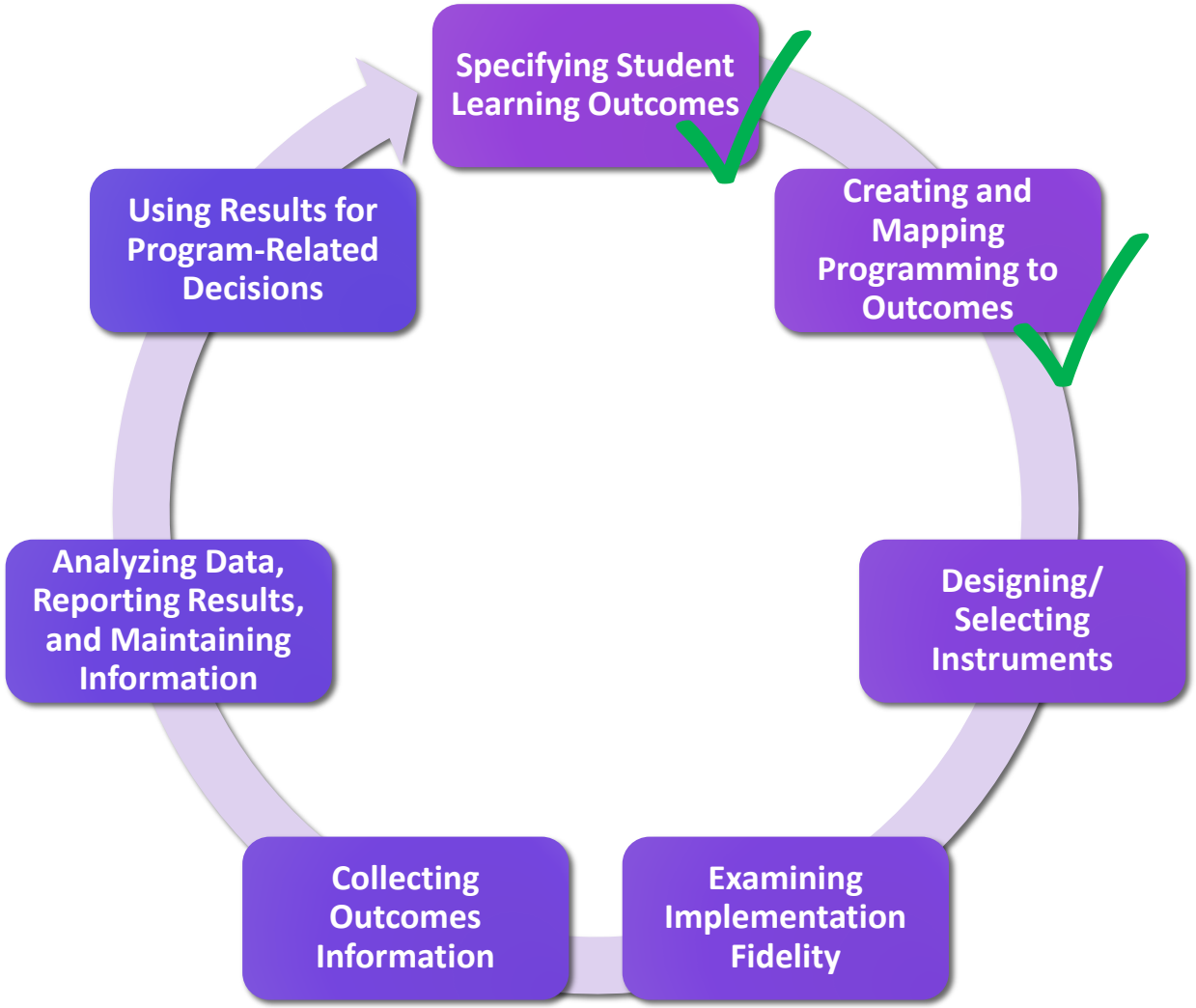
The Center for Assessment and Research Studies

Repositories of Existing Measures

SARAH ALAHMADI, M.S.

DOCTORAL ASSISTANT
CENTER FOR ASSESSMENT AND RESEARCH STUDIES
JAMES MADISON UNIVERSITY

The Assessment Cycle



Where to find existing measures?

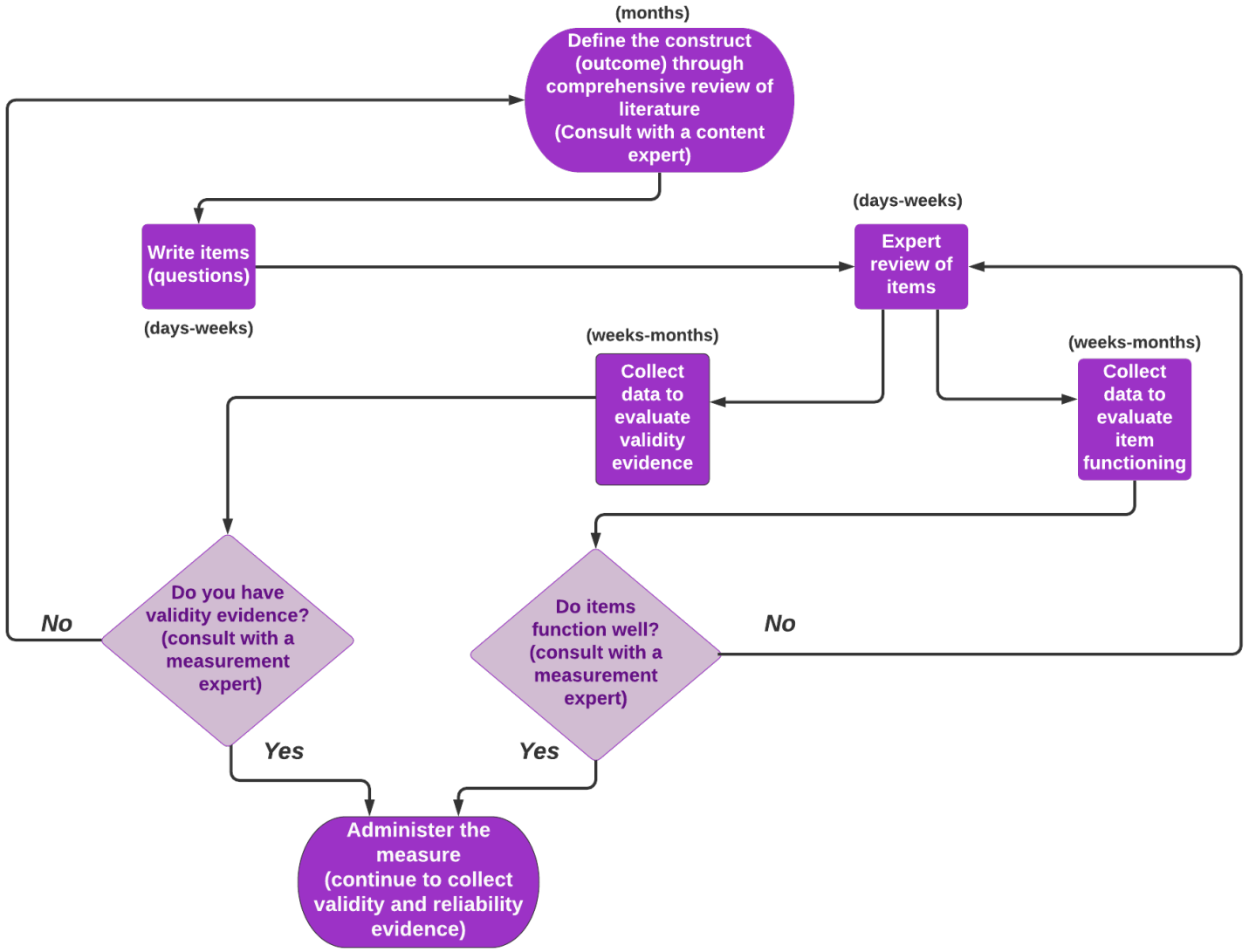
"What's a good measure of that outcome?"

Resources to find existing and psychometrically-sound measures

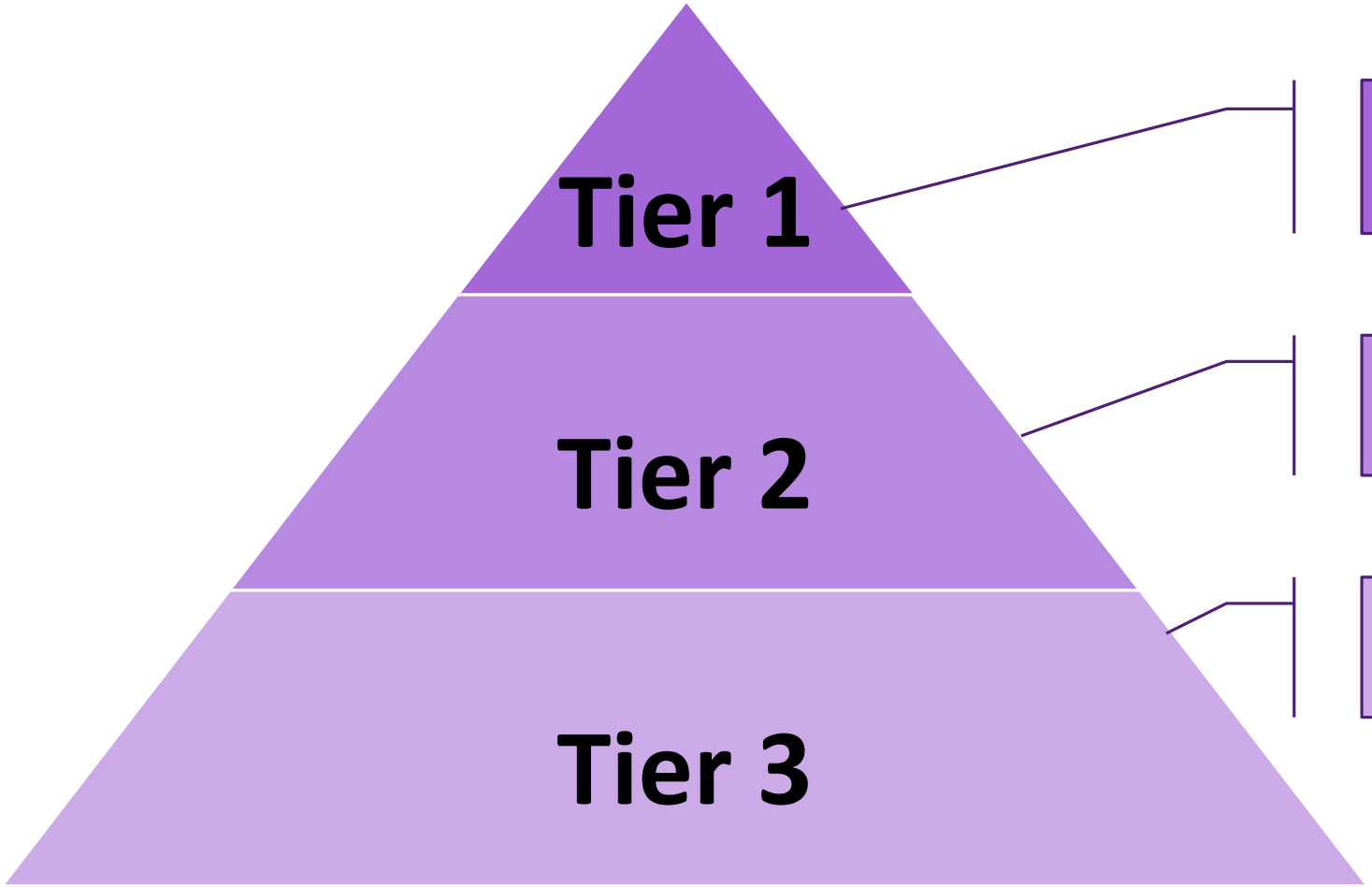
Sara J. Finney, PhD, Gabriel R. Gilmore, MA, & Sarah Alahmadi, MS

James Madison University

Steps to **designing** a measure



Where to **find** existing measures?



- **Measures are carefully curated**
- **Psychometric information is summarized**
- **Experts review the quality of measures**

- **Collection of measures**
- **Psychometric information is summarized**

- **Collection of measures**

Tier 1 example

Repository: Mental Measurements Yearbooks

REPOSITORIES OF PRE-EXISTING MEASURES				
Tier 1 - Provides psychometric information and their own rating of the quality of the measure **4 Repositories**				
Source	**CAS Standards**	**Description of Resource**	**Information About the Characteristics of Measures**	**Examples of Measures**
Mental Measurements Yearbooks Hard Copy Books https://buros.org/test-reviews-information Test Reviews online https://marketplace.unl.edu/buros/	- Knowledge acquisition, construction, integration, & application* - Cognitive complexity* - Intrapersonal development* - Interpersonal competence* - Humanitarianism & Civic Engagement* - Practical competence*	The Mental Measurements Yearbooks (MMY) are published by the Buros Center for Testing. The MMY includes commercially available test reviews that evaluate information to encourage informed selection of measures. Since all of the measures reviewed in the MMY are commercial in nature, no example items or full measures are provided. The MMY is currently in its 21st edition, and a full list of all measures reviewed in the MMY series is available here. .	For each measure, the MMY entries include its purpose, target population, publication date, administration type (individual or group), price, time for completion, and reviewer, author and publisher information. A description of the measure and its development is provided along with technical information, commentary from the reviewer, summary of the measure's quality, and recommendation for use. Technical details include normative data, reliability information (test-retest, internal consistency, and inter-rater), and validity information (internal, external, convergent, and divergent). Two professional reviewers evaluate each measure and both of their reports are included for each entry.	1. College Survival and Success Scale 2. Multi-Dimensional Intelligence Test 3. Cornell Critical Thinking Tests 4. Test of Everyday Reasoning 5. Stress Assessment Questionnaire These links will provide access to the reviews if your academic library has a database subscription. If not, simply enter the name of the measure or topic into the Test Reviews Online link to the left.

Tier 1 example

Repository: Mental Measurements Yearbooks

Detailed Record

HTML Full Text

Related Information

Find Similar Results
using SmartText Searching.

Result List Refine Search 2 of 9

California Critical Thinking Skills Test [Revised]

Acronym:	CCTST.
Authors:	Facione, Peter A.; Facione, Noreen C.; Blohm, Stephen W.; Giancarlo, Carol Ann F.
Publication Date:	1990-2007.
Publisher Information:	Insight Assessment—The California Academic Press LLC, 217 La Cruz Avenue, Millbrae, CA, 94030, info@insightassessment.com, www.insightassessment.com .
Purpose:	'Specifically designed to measure the skills dimension of **critical thinking**.'
Test Category:	Intelligence and General Aptitude.
Population:	College students and **adults**.
Administration:	Group.
Time:	(45) minutes or unlimited.
Price Data:	Available from publisher.
Comments:	This test can be administered online or via paper and pencil
Sublistings:	a)FORM A. Scores, 4: Analysis, Inference, Evaluation, Total. b) FORM B. Scores, 4: Analysis, Inference, Evaluation, Total. c) FORM C. Scores, 6: Analysis, Inference, Evaluation, Deductive Reasoning, Inductive Reasoning, Total..
Cross References:	For reviews by Matthew E. Lambert and William E. Martin, Jr., see 18:21; for reviews by Robert F. McMorris and William B. Michael of an earlier edition, see 12:58.
Reviewers:	Lambert, Matthew E.; Martin, William E..
Special Editions:	All forms available in English; Form A is available in Chinese (Beijing and Taiwan), Hebrew, Korean, Spanish (Mexico), and Thai; Form B is available in Portuguese; Form 2000 is available in French (Canadian), Italian, Korean, and Spanish (Spain)..
Yearbook Volume:	18.
Yearbook Reference:	R. A. Spies, J. F. Carlson, & K. F. Geisinger (Eds.), The eighteenth mental measurements yearbook. 2010.
Published Test Description:	The California **Critical Thinking** Skills Test [Revised]. Purpose: 'Specifically designed to measure the skills dimension of **critical thinking**.' Population: College students and **adults**. Publication Dates: 1990-2007. Acronym: CCTST. Administration: Group. Price Data, 2008: $60 per manual (2007, 23 pages) and example test booklet; $250 per one-time client online testing set-up fee; $20 each first 50 online testing uses; $15 each 51-500 online testing uses; $10 each online testing use over 500; $20 each first 50 paper-and-pencil test booklets, answer forms, and scoring sheets; $15 each 51-500 paper-and-pencil test booklets, answer forms, and scoring sheets; $10 each paper-and-pencil test booklets, answer forms, and scoring sheets; discounts available for nonprofit organizations. Foreign Language Editions: All forms available in English; Form A is available in Chinese (Beijing and Taiwan), Hebrew, Korean, Spanish (Mexico), and Thai; Form B is available in Portuguese; Form 2000 is available in French (Canadian), Italian, Korean, and

Tools

Google Drive
Add to folder
Print
E-mail
Save
Cite
Export
Create Note
Permalink
Listen
Translate

Tier 1 example

Repository: Mental Measurements Yearbooks

California **Critical Thinking** Skills Test [Revised]

Review of The California **Critical Thinking** Skills Test [Revised] by MATTHEW E. LAMBERT, Clinical Assistant Professor of Neuropsychiatry, Texas Tech University Health Sciences Center, Department of Neuropsychiatry, Lubbock, TX:

DESCRIPTION. The California **Critical Thinking** Skills Test [Revised] (CCTST) is a 34-item multiple-choice test, available in three forms, designed to assess the **critical thinking** skills deemed essential to college education. It is available in three forms (A, B, and 2000) that can be administered via paper-and-pencil or an internet connection to the publisher's test administration site. The paper-and-pencil format requires a computer-scanned answer form on which test takers "bubble-in" their answers. The individual items address various skills of evaluating information, drawing inferences, and justifying objections to inferences. The updated Form 2000 focuses on evaluating **critical thinking** skills associated with ideas related to the new century. Item content is nonspecific to any one college discipline with the focus more on underlying skills necessary for a college education.

Standard administration time is 45 minutes, although it can be increased if local norms are to be developed. Regardless of administration procedures, all scoring is completed by the test publisher with results being supplied via electronic media either through the online test site or transmitted to the administrator. In addition, various forms are available in different languages that include English, Chinese, Hebrew, Korean, Spanish, Thai, French, Italian, and Portuguese. CCTST scores include: Analysis, Inference, Evaluation, Deductive Reasoning, Inductive Reasoning, and Total, each of which reflects various aspects of the **critical thinking** concept.

The 23-page basic manual discusses the development, administration and scoring, and psychometric underpinnings of the instrument. A supplemental manual is also provided, which addresses paper-and-pencil versus online administrations.

DEVELOPMENT. The items that comprise the CCTST Forms A and B were drawn from a pool of about 200 items created as part of a longterm study to assess **critical thinking** and designed to be discipline neutral. References to sex or social class were avoided with equal numbers of male and female referents in the items. The base item pool had been previously analyzed and determined to adequately discriminate **critical** thinkers and had high item-total correlations. Item selection for the final forms was designed to reflect five **critical thinking** skills previously identified by experts as the basis for problem solving and decision making: Interpretation, Analysis, Evaluation, Explanation, and Inference. Unfortunately, no information is provided as to how individual items were selected from the overall item pool.

Form 2000 consists of 22 items from Form A and 12 new items that require application of **critical thinking** skills to visual information (i.e., diagrams and charts) versus Forms A and B, which present only textual information. No information is presented about how the additional Form 2000 items were constructed or selected.

Regardless of the test form, items are grouped in order of **critical thinking** complexity although the manual does little to discuss a hierarchical concept of **critical thinking** skills. Distractors were written to directly address errors associated with the specific **critical thinking** skill to be assessed. Three or four distractors are associated with each item. The rationale for using the different numbers of distractors is not discussed.

TECHNICAL. Internal consistency estimates for the CCTST were derived from the original validation studies and produced Kuder Richardson-20 values ranging from .68 to .70 for Form A when it was used as a pretest or posttest, respectively, and .71 for Form B for a subgroup of individuals who completed both forms in an equivalency study. For a separate group of graduate nursing students who completed Form B, the Kuder Richardson-20 estimate was .75. Alternate form reliability between Form A and Form B was .78 for students who took both forms. No efforts to address test-retest reliability were presented in the test manuals. Thus, the CCTST has demonstrated adequate internal consistency and alternate form reliability.

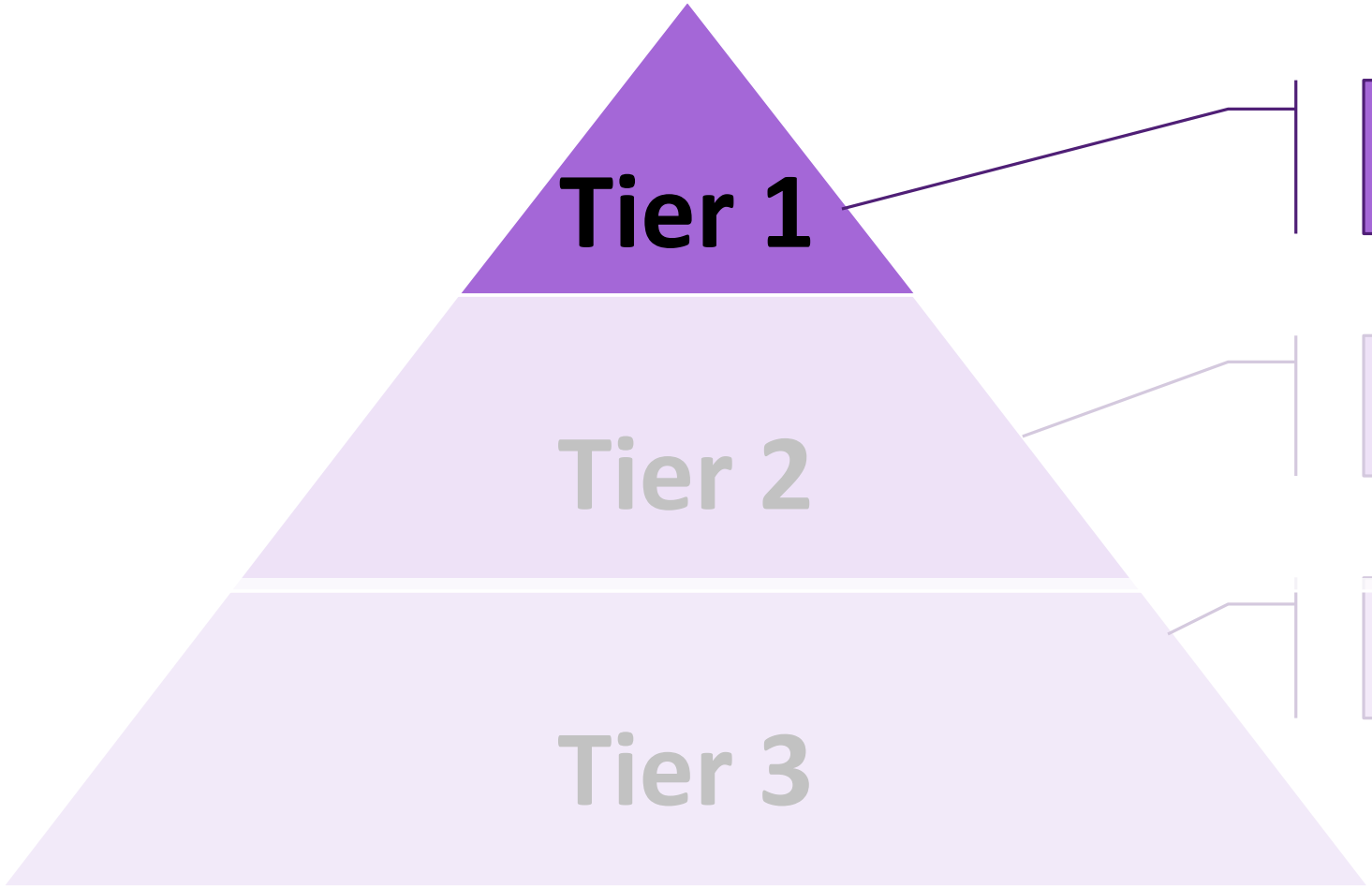
Comparison of scores from CCTST items that make up Form A and Form 2000 revealed correlations of .91 and .87, drawn from two divergent samples. Again, Kuder Richardson-20 estimates of internal consistency for Form 2000 were .80 for a sample drawn from a large public university and .78 for a smaller sample of health science school students. As such, the Form 2000 appears to have slightly greater internal consistency than Forms A and B.

Content validity is discussed in terms of the manner in which items were selected from a pool previously analyzed for their ability to discriminate well between individuals in terms of **critical thinking** skills and by high interitem correlations. The discussion surrounding item selection related to content validity is limited and could benefit from significant expansion.

Construct validity is supported by CCTST pretest to posttest score improvement for students taking a required college level **critical thinking** course. A .74 score increase was noted across administrations, which was significant at a .008 level. Yet, a matched-pairs analysis demonstrated an average 1.45 point score gain and the average student in the paired sample moved from the 55th to the 70th percentile as compared to pretest scores. Control groups for both cross-sectional and matched-pair groups did not demonstrate significant gains. Other experiments controlling for test experience did not demonstrate differences between test-experienced and test-naïve students following a **critical thinking** course.

Similarly, criterion validity was assessed by comparing CCTST scores to various measures of academic performance and demographic variables. Dependent upon the measures, significant correlations

Where to **find** existing measures?



- **Measures are carefully curated**
- **Psychometric information is summarized**
- **Experts review the quality of measures**

- Collection of measures
- Psychometric information is summarized

- Collection of measures

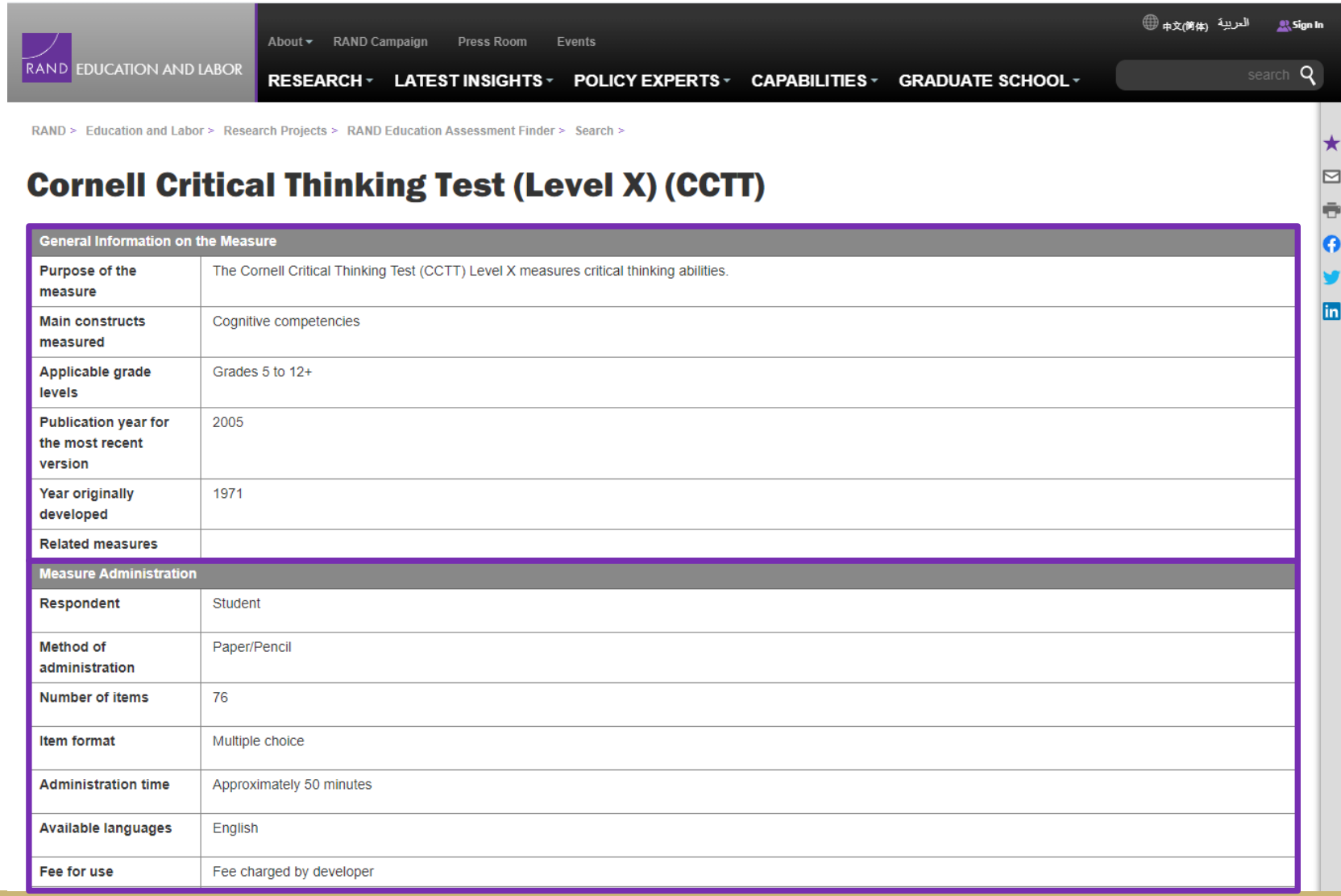
Tier 2 example

Repository: RAND Education Assessment Finder

Tier 2 – Provides psychometric information, but not their own rating of the quality of the measure 18 Repositories				
Source	**CAS Standards**	**Description of Resource**	**Information About the Characteristics of Existing Measures**	**Examples of Measures**
RAND Education Assessment Finder https://ww…g/educati… labor/projects/assessments.html	- Knowledge acquisition, construction, integration, & application - **Cognitive complexity*** - **Intrapersonal development*** - **Interpersonal competence*** - Humanitarianism & Civic Engagement - Practical competence	The RAND Education Assessment Finder was created by the RAND Corporation to benefit three specific groups: practitioners, researchers, and policymakers. The main inclusion criteria of this repository's educational assessments required that: (1) measures be related to interpersonal, intrapersonal, or higher-order cognitive competencies; (2) measures be appropriate to use in educational settings and not be specific to clinical settings or curriculum-embedded assessments; and (3) be appropriate to use with students in the United States. For more information on the development of this repository, you can visit the following article: Schweig, J., Baker, G., Hamilton, L. S., & Stecher, B. M. (2018). *Building a Repository of Assessments of Interpersonal, Intrapersonal, and Higher-Order Cognitive Competencies*	The repository search can be filtered by keywords, competency, grade level (pre-K through postsecondary), respondent, method of administration, Administration time, item format (selected response, free response, or performance task), and fee for use (commercial or non-commercial). There is also a "compare" feature, where you could select more than one measure for a close-up comparison. The information presented for each measure include: • General information: purpose, construct, grade levels, publication year, year developed, related measures. • Measure administration: respondent, method, number of items, item format, administration time, available languages, fee, credentials needed for administration. • Scoring: overall score reporting, subscore reporting, scoring procedure, interpretive information. • Evidence of technical quality: population, reliability, validity. • Locating the measure: link(s) to obtain a copy of the measure.	1. Cornell Critical Thinking Test 2. Test of Everyday Reasoning 3. Torrance Test of Creative Thinking—Verbal 4. Multidimensional Self Concept Scale 5. Personal Skills Map—Long Version

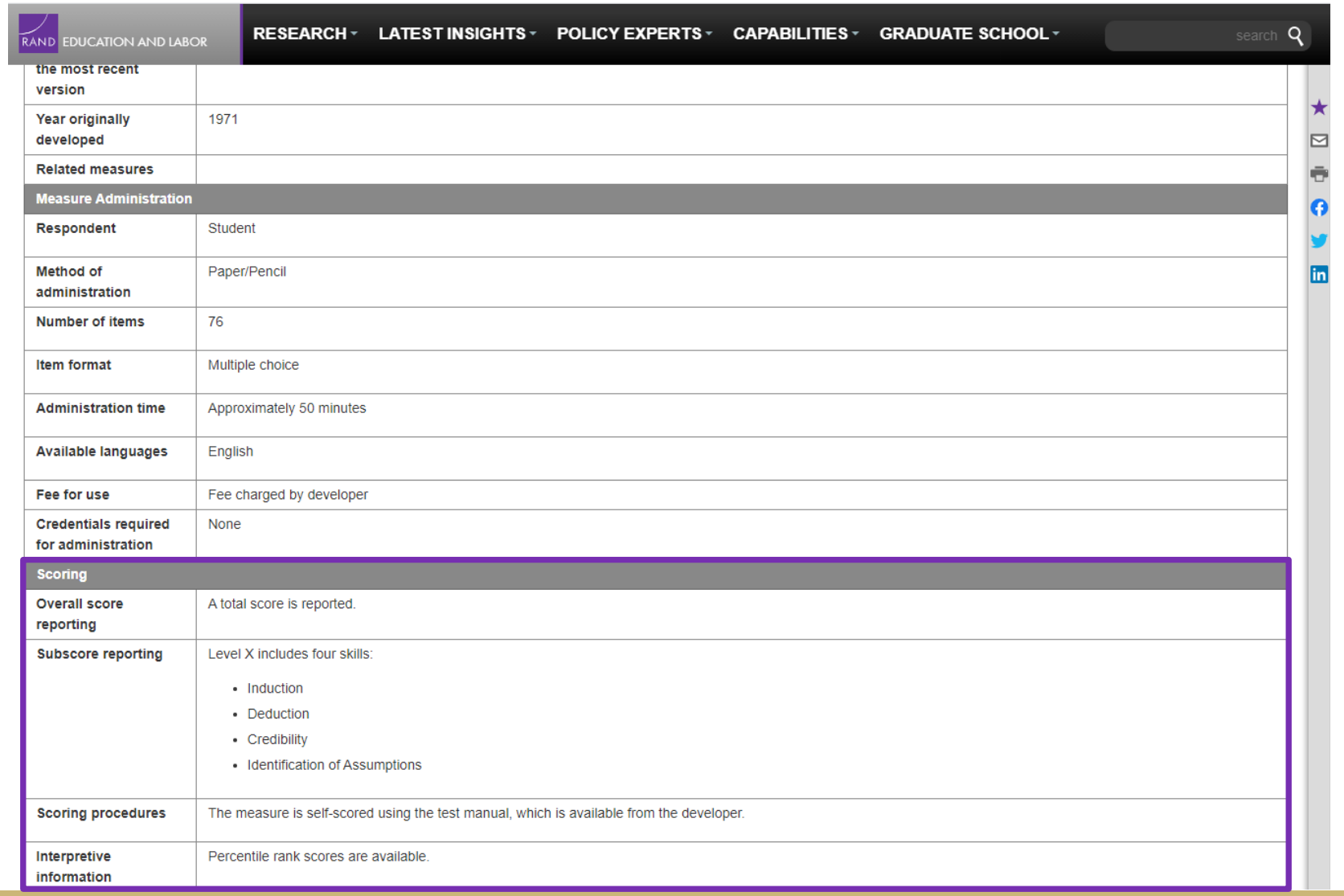
Tier 2 example

Repository: RAND Education Assessment Finder



Tier 2 example

Repository: RAND Education Assessment Finder

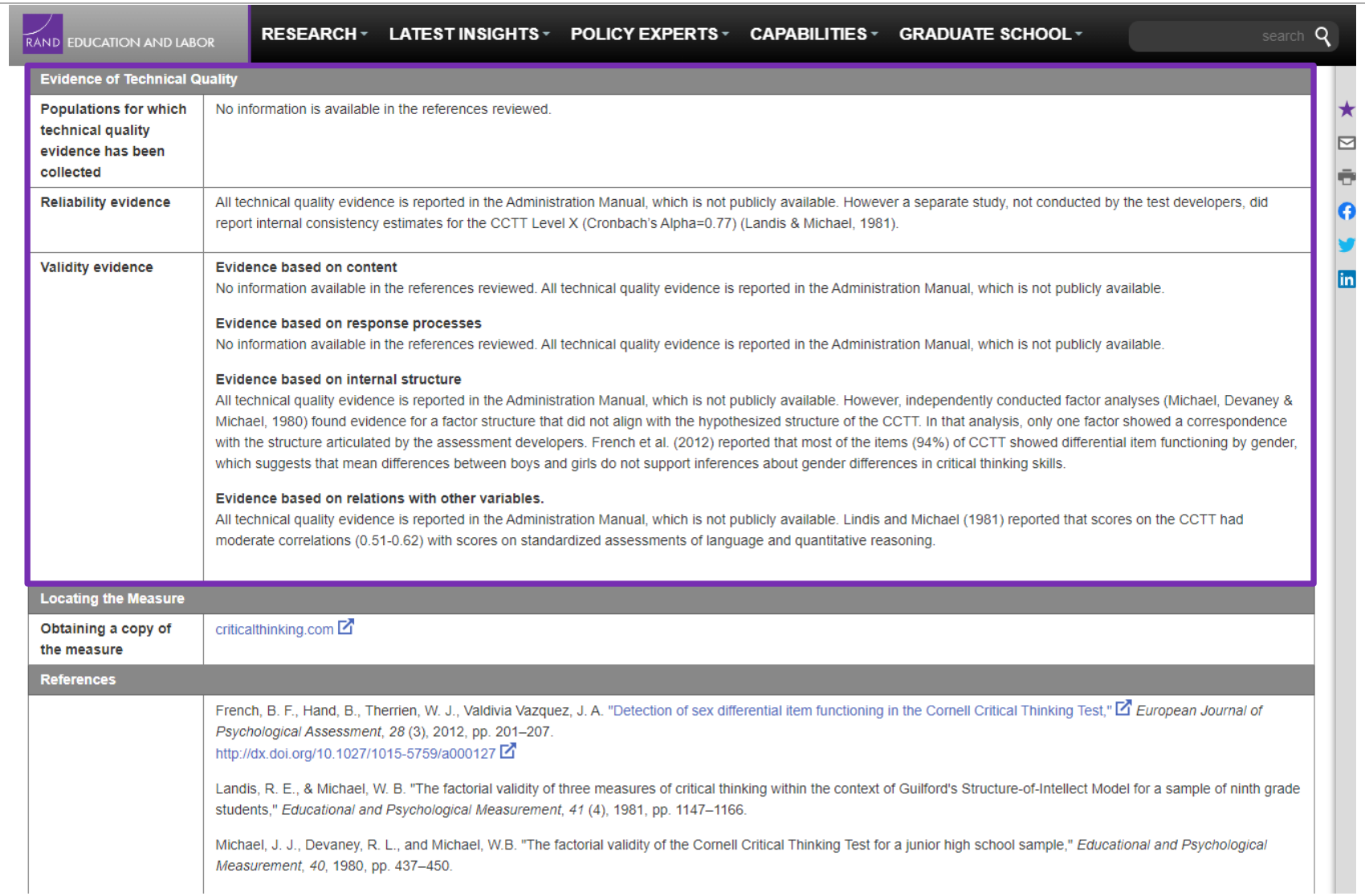


RAND EDUCATION AND LABOR | RESEARCH | LATEST INSIGHTS | POLICY EXPERTS | CAPABILITIES | GRADUATE SCHOOL | search

the most recent version	
Year originally developed	1971
Related measures	
Measure Administration	
Respondent	Student
Method of administration	Paper/Pencil
Number of items	76
Item format	Multiple choice
Administration time	Approximately 50 minutes
Available languages	English
Fee for use	Fee charged by developer
Credentials required for administration	None
Scoring	
Overall score reporting	A total score is reported.
Subscore reporting	Level X includes four skills: • Induction • Deduction • Credibility • Identification of Assumptions
Scoring procedures	The measure is self-scored using the test manual, which is available from the developer.
Interpretive information	Percentile rank scores are available.

Tier 2 example

Repository: RAND Education Assessment Finder

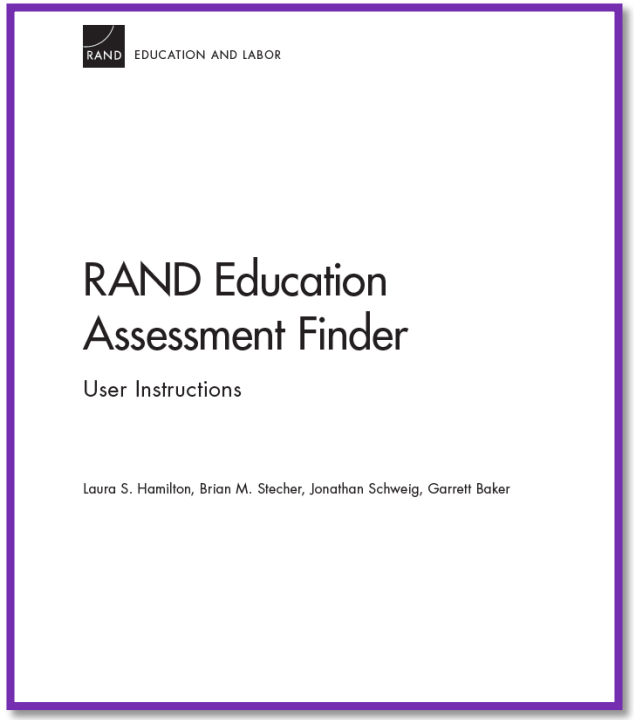


RAND EDUCATION AND LABOR | RESEARCH | LATEST INSIGHTS | POLICY EXPERTS | CAPABILITIES | GRADUATE SCHOOL | search

Evidence of Technical Quality	
Populations for which technical quality evidence has been collected	No information is available in the references reviewed.
Reliability evidence	All technical quality evidence is reported in the Administration Manual, which is not publicly available. However a separate study, not conducted by the test developers, did report internal consistency estimates for the CCTT Level X (Cronbach's Alpha=0.77) (Landis & Michael, 1981).
Validity evidence	**Evidence based on content** No information available in the references reviewed. All technical quality evidence is reported in the Administration Manual, which is not publicly available. **Evidence based on response processes** No information available in the references reviewed. All technical quality evidence is reported in the Administration Manual, which is not publicly available. **Evidence based on internal structure** All technical quality evidence is reported in the Administration Manual, which is not publicly available. However, independently conducted factor analyses (Michael, Devaney & Michael, 1980) found evidence for a factor structure that did not align with the hypothesized structure of the CCTT. In that analysis, only one factor showed a correspondence with the structure articulated by the assessment developers. French et al. (2012) reported that most of the items (94%) of CCTT showed differential item functioning by gender, which suggests that mean differences between boys and girls do not support inferences about gender differences in critical thinking skills. **Evidence based on relations with other variables.** All technical quality evidence is reported in the Administration Manual, which is not publicly available. Lindis and Michael (1981) reported that scores on the CCTT had moderate correlations (0.51-0.62) with scores on standardized assessments of language and quantitative reasoning.
Locating the Measure	
Obtaining a copy of the measure	criticalthinking.com
References	
	French, B. F., Hand, B., Therrien, W. J., Valdivia Vazquez, J. A. "Detection of sex differential item functioning in the Cornell Critical Thinking Test," *European Journal of Psychological Assessment*, *28* (3), 2012, pp. 201–207. http://dx.doi.org/10.1027/1015-5759/a000127 Landis, R. E., & Michael, W. B. "The factorial validity of three measures of critical thinking within the context of Guilford's Structure-of-Intellect Model for a sample of ninth grade students," *Educational and Psychological Measurement*, *41* (4), 1981, pp. 1147–1166. Michael, J. J., Devaney, R. L., and Michael, W.B. "The factorial validity of the Cornell Critical Thinking Test for a junior high school sample," *Educational and Psychological Measurement*, *40*, 1980, pp. 437–450.

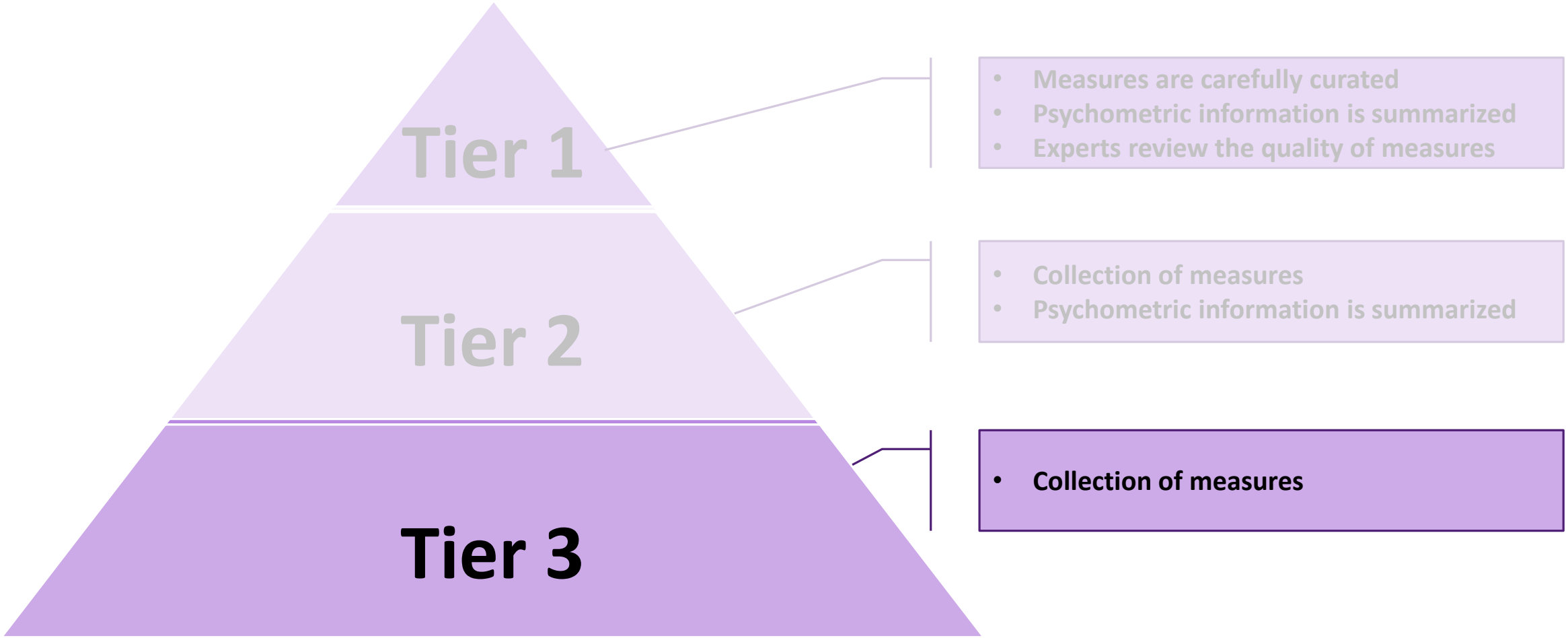
Tier 2 example

Repository: RAND Education Assessment Finder



"Because interpreting validity evidence is complex and generally requires measurement expertise, users are encouraged to seek input from measurement experts to evaluate the adequacy and relevance of the available evidence for a particular assessment purpose. For more information on validity evidence, see the "Choosing and Using" guidance."

Where to **find** existing measures?



Tier 3 example
Repository: PsycTests

Tier 3 – Provides no psychometric information or their own rating of the quality of the measure 10 Repositories				
Source	CAS Standards	Description of Resource	Information About the Characteristics of Existing Measures	Examples of Measures
PsycTests https://www.apa.org/pubs/databases/psyctests	**- Knowledge acquisition, construction, integration, & application*** **- Cognitive complexity*** **- Intrapersonal development*** - Interpersonal competence - Humanitarianism & Civic Engagement **- Practical competence***	PsycTests is a database run by APA that contains an extensive collection of resources associated with psychological measures. PsycTests currently houses 61,460 records and is updated on a monthly basis. The included resources date as far back as 1896, but 79% are from 2000 and later. Measures are collected from a variety of sources: • Directly from authors • Peer reviewed journals • Books and handbooks • Dissertations • Websites • Archives of the History of American Psychological Test Collection • Test publishers (descriptive information only) Must log in with institutional access, then select PsycTests in the normal PsycNet database.	PsycTests provides a descriptive summary of each measures, its development and its administration. The majority of the measures are available for download. Each measure has a "Master Test Profile" that includes a summary of the measure, purpose, construct, and contact information of authors. This resource does not provide psychometric reviews of measures; it is a collection of a variety of measures. No information regarding reliability or validity is provided. If validity, reliability, and factor analysis information are included in an article, the record will indicate the presence of such information. Examples of subject areas that measures cover include: • Educational Measures • Aggression, coping, or functional status questionnaires • Aptitude and achievement measures • Resilience, anger response, or substance abuse inventories • Intelligence Tests • Neuropsychological assessments	1. Ethical Reasoning Inventory 2. College Student Empowerment Scales 3. Measures of Obstacles to Succeeding Academically in College 4. Perceptions of the LGBTQ College Campus Climate Scale 5. Student Adaptation to College Questionnaire – Modified

Tier 3 example

Repository: PsycTests

Detailed Record

Get JMU Access

Related Information

Find Similar Results
using SmartText Searching.

‹ Result List Refine Search ‹ 1 of 2 ›

Psychological Critical Thinking Exam (PCTE)

Test Type:	Original
Other Version:	9999-43670-000, Psychological **Critical Thinking** Exam--Revised, Revision
Acronym:	PCTE
Source Used:	The majority of metadata for this record was created from PsycINFO Record: 1999-03851-011
Purpose:	The purpose of the Psychological **Critical Thinking** Exam is to assess psychology students' ability to evaluate claims in the field of psychological science.
Description:	The Psychological **Critical Thinking** Exam (PCTE; Lawson, 1999) is a 14-item measure of the ability of psychological students to think critically, or evaluate claims in a way that explicitly incorporates basic principles of psychological science. This measure is based on a list of 9 questions that psychology majors can ask themselves while evaluating claims (e.g., "Is there a control or comparison group against which to assess the performance of the experimental group?"). These were constructed using basic principles of psychological science discussed in Stanovich (1998) and undergraduate research textbooks and on other authors' **critical thinking** guidelines (Smith, 1995; Tavris & Wade, 1997). To develop the PCTE, psychology faculty wrote various types of claims, each of which related to one or more of the questions. The authors selected 14 claims that, as a group, relate to all of the questions. An example is as follows: "A researcher tested a new drug designed to decrease depression. She gave it to 100 clinically depressed patients and discovered that their average level of depression, as measured by a standardized depression inventory, declined after 4 months of taking the drug. She concluded that the drug reduces depression. (Question 3)." Students are asked to identify, in writing, the problems (if any) with each claim. (PsycTESTS Database Record (c) 2019 APA, all rights reserved)
Test Year:	1999
Author:	Lawson, T. J.
Email:	Lawson, T. J.: tim_lawson@mail.msj.edu
Affiliation:	Lawson, T. J.. College of Mount St. Joseph, Cincinnati, Ohio, United States
Correspondence:	Lawson, T. J., College of Mount St. Joseph, Psychology Program Cincinnati, Ohio, United States, 45233, tim_lawson@mail.msj.edu
Instrument Type:	Vignette/Scenario
Format:	For each of the 14 scenarios/claims, respondents are asked to identify the problems (if any).
Language Present:	English
Language Available:	English
Construct:	Psychological **Critical Thinking**

Tools

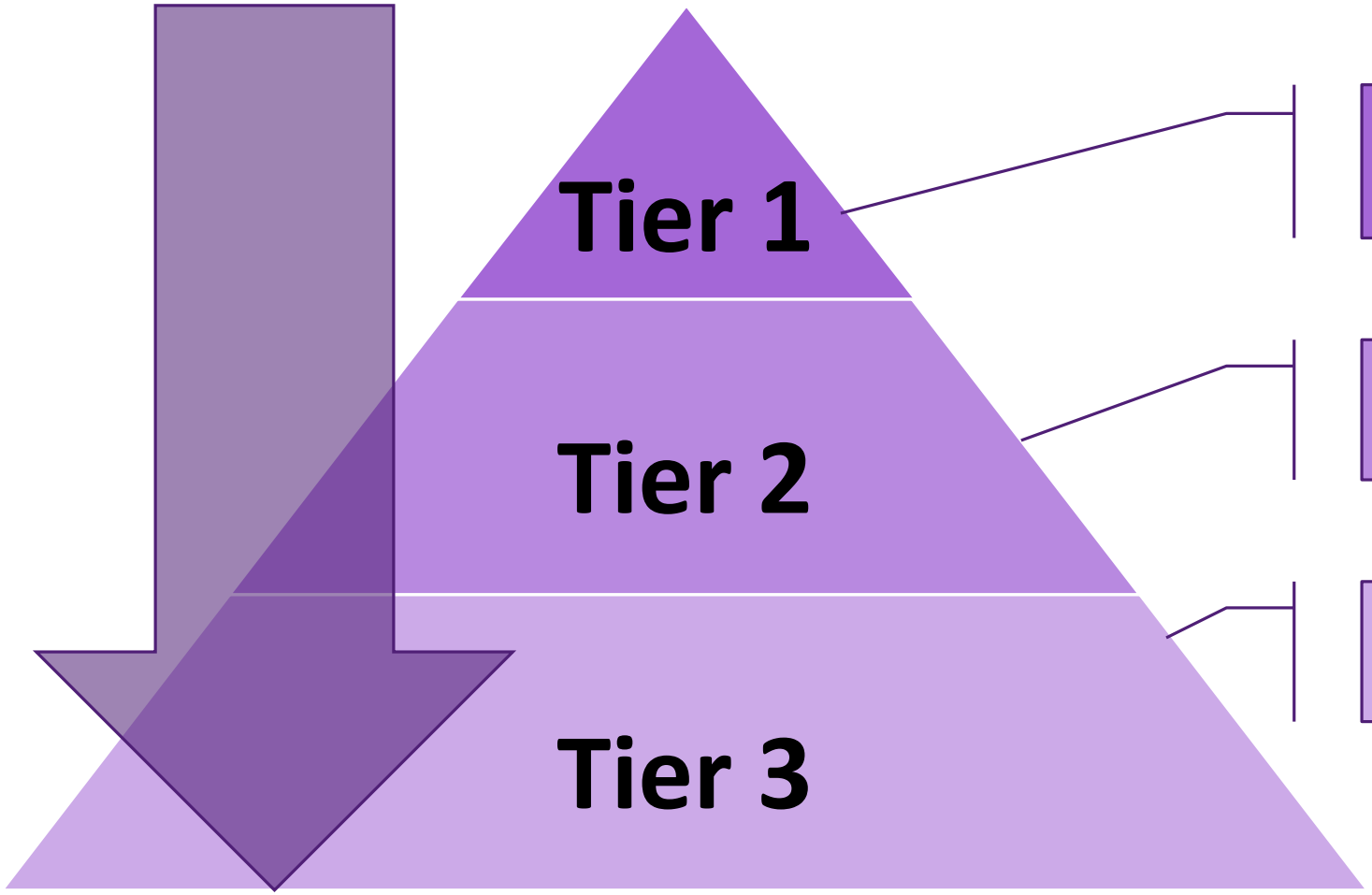
Google Drive
Add to folder
Print
E-mail
Save
Cite
Export
Create Note
Permalink

Tier 3 example

Repository: PsycTests

Construct:	Psychological **Critical Thinking**
Commercial Availability:	No
Permissions:	Contact Publisher and Corresponding Author
Fee:	No
Test Items Available:	No
Classification:	5400 Cognitive Processes, Memory, and Decision Making 5900 Education, Teaching, and Student Characteristics
Reliability:	No reliability indicated.
Validity:	No validity indicated.
Factor Analysis:	No factor analysis indicated.
Test Location:	Text, Page 207, 1999-03851-011 Appendix, Page 209, 1999-03851-011
Number of Test Items:	This measure consists of 14 items.
Administration Method:	Paper
Age Group:	Adulthood (18 yrs & older); Young Adulthood (18-29 yrs)
Population Group:	Human; Male; Female
Other Population Details:	Sample: Senior Psychology Majors
Keywords:	Claim Evaluation; Problem Identification; Psychological **Critical Thinking** Exam; Psychological Science Principles; Psychology Majors; Psychology Students; Test Development; Undergraduate Students
Index Terms:	Academic Specialization; Behavioral Sciences; Cognitive Ability; College Students; **Critical Thinking**; Educational Measurement; Evaluation; Experimentation; Problem Solving; Psychology; Test Construction; Undergraduate Education
Source Citation:	**Test Development** Lawson, Timothy J. (1999). Assessing psychological **critical thinking** as a learning outcome for psychology majors. Teaching of Psychology, 26(3), 207-209. doi: 10.1207/S15328023TOP260311 **PsycINFO Record**: 1999-03851-011

Summary of **tier utility**

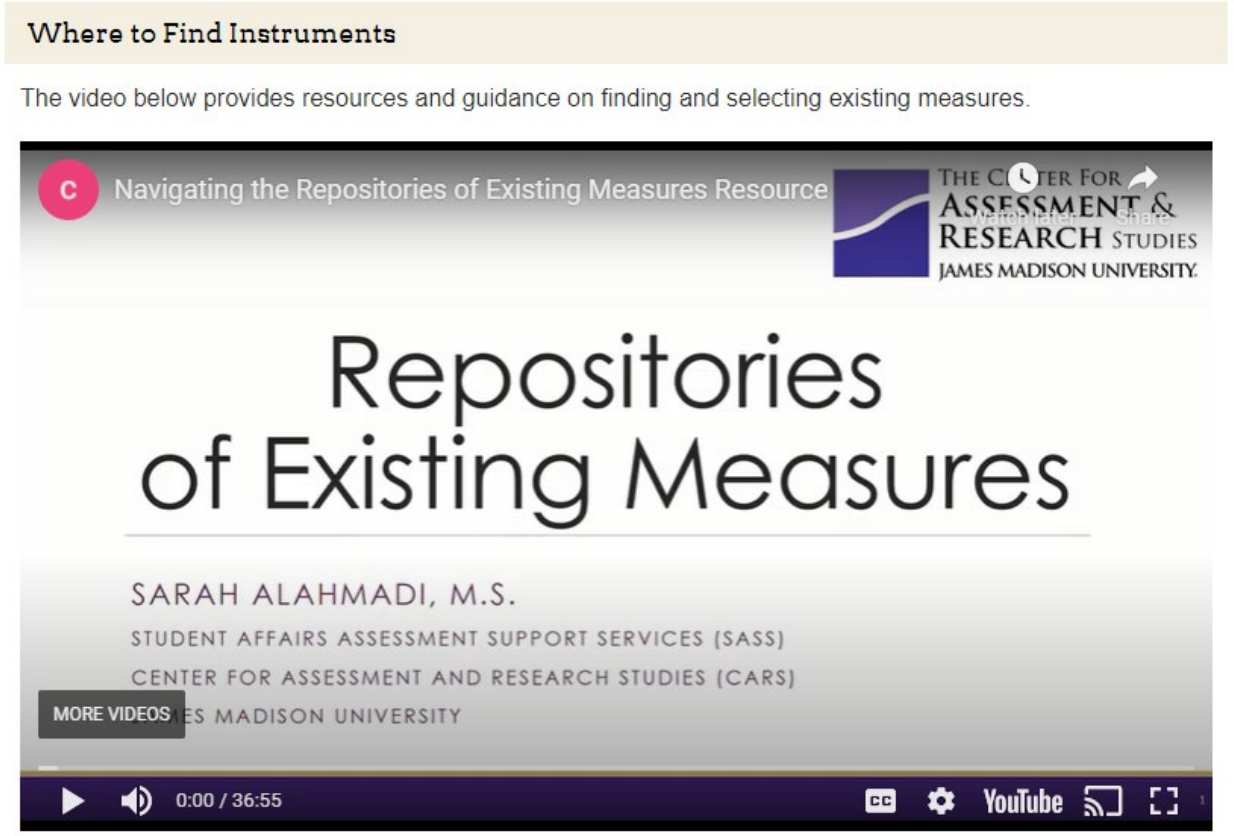


- **Measures are carefully curated**
- **Psychometric information is summarized**
- **Experts review the quality of measures**

- **Collection of measures**
- **Psychometric information is summarized**

- **Collection of measures**

How to **use** the resource **step-by-step**?



Sarah Alahmadi, MS
alahmasi@jmu.edu

Student Affairs Assessment Support Services (SASS)
www.jmu.edu/assessment/sass
Center for Assessment & Research Studies (CARS)
James Madison University

www.jmu.edu/assessment/sass/AC-step-three.shtml#where_to_find

JMU The Center for Assessment and Research Studies

Using a Questioning Approach to Increase Value for Assessment

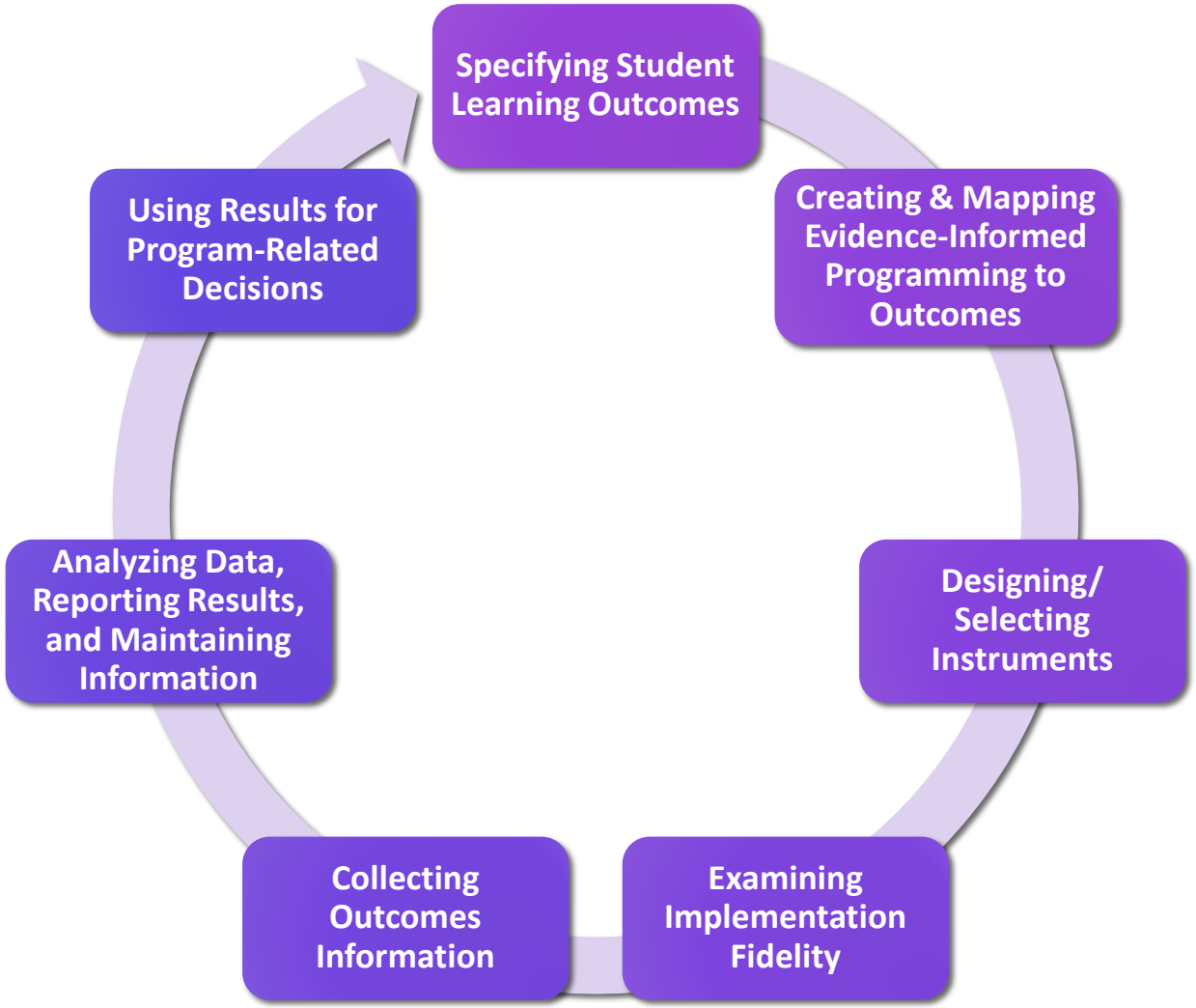
CHRIS PATTERSON, MA

HE/HIM/HIS

“Assessment is too hard”

“Assessment doesn’t help me”

The Outcomes Assessment Cycle



Influencing Meaning-Making

- Meaning-making influences behaviors (Walton & Wilson, 2018)
- We can influence the way one interprets and makes meaning of an experience through asking questions (Miller & Rollnick, 2009; Walton & Wilson, 2018)
- *We can influence assessment-related behaviors by asking relevant questions about how one applies assessment to their work!*

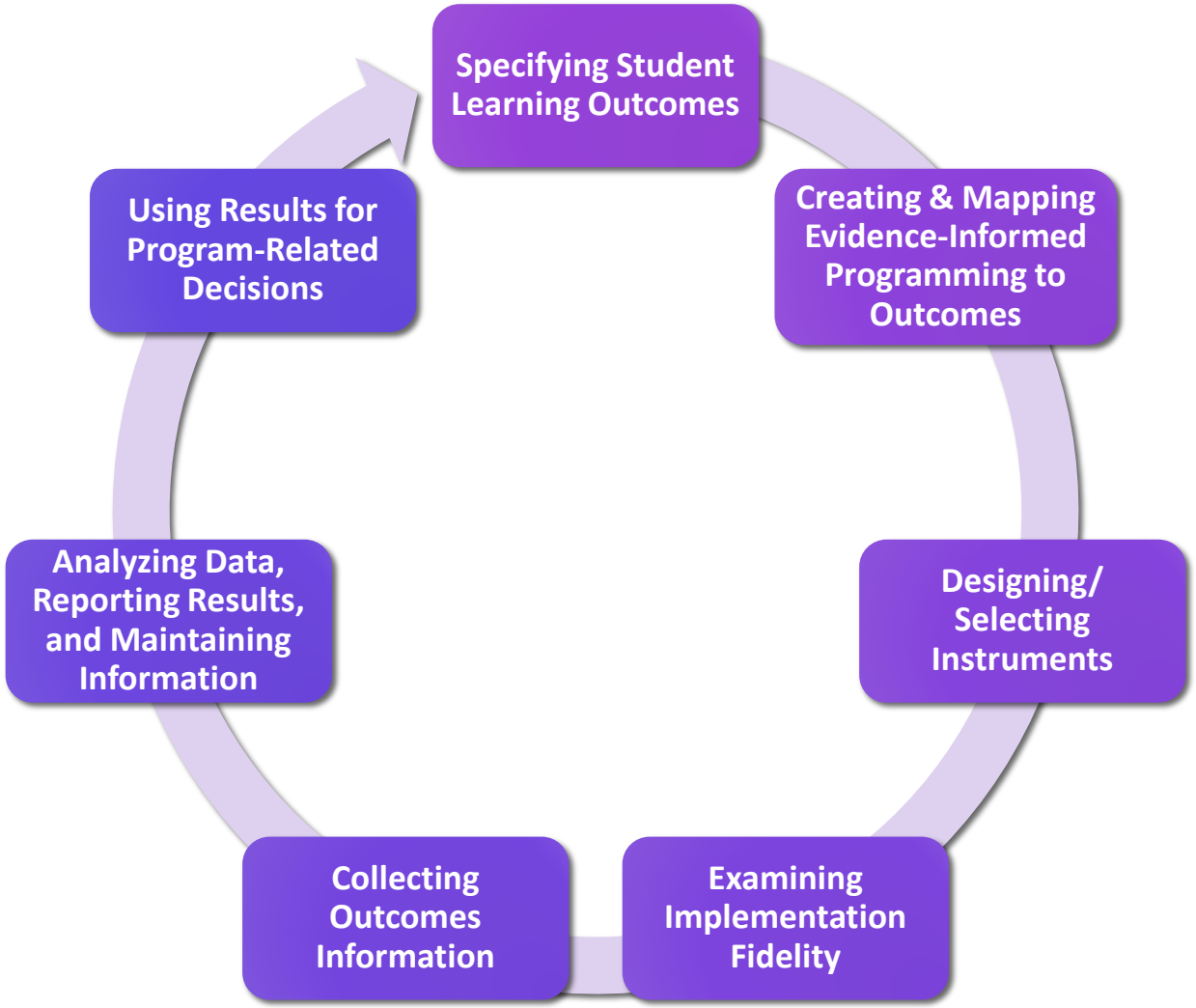
Activity on Changing Behaviors

STEP 1: Write down at least one question that can be answered by engaging in outcomes assessment. Explain why that question (or those questions) are relevant to work as an educator.

STEP 2: Read through the questions below and compare the questions you generated in Step 1 to the questions stated here.

- As an educator, what do you believe your students should know, think, or be able to do?
- What programming would you, the curriculum designer, create to foster the desired learning and development?
- As an educator, how would you measure the student learning and development outcomes?
- What evidence would you, the curriculum designer, gather to describe the programming the students actually experienced?
- As an educator, how and when would you collect outcomes data to best understand student learning and development?
- As an evidence-informed educator, how would you analyze student learning and development data and interpret the results?
- As a designer of learning and development opportunities, how would you use the assessment results to improve your programming?

The Outcomes Assessment Cycle



Questions-Based Approach to Assessment

Assessment Cycle Step	Question
Creating Student Learning Outcomes	As an educator, what do you believe your students should know, think, or be able to do?
Creating and Mapping Programming to Outcomes	What programming would you, the curriculum designer, create to foster the desired learning and development?
Select or Design Instrumentation	As an educator, how would you measure the student learning and development outcomes?
Collect Implementation Fidelity	What evidence would you, the curriculum designer, gather to describe the programming the students actually experienced?
Collect Outcomes Data	As an educator, how and when would you collect outcomes data to best understand student learning and development?
Analyze Data and Interpret Results	As an evidence-informed educator, how would you analyze student learning and development data, and interpret the results?
Make Changes to Programming	As a designer of learning and development opportunities, how would you use the assessment results to improve your programming?

Activity on Changing Behaviors

STEP 3: Explain how your questions from Step 1 relate to the questions above. How are they similar? How are they different?

STEP 4: Think of programming you designed, are currently designing, or have implemented. Draft answers to the 7 questions stated in Step 2. It is typical for this step to take time and effort. It is difficult, and worthwhile. You are articulating a plan to understand your impact on student learning and development. You will update your initial responses as you engage in the outcomes assessment process. At this point, what is most important is to begin processing and answering these questions.

Activity on Changing Behaviors

STEP 5: Answer the following: Assessment is helpful for your work as an educator, isn't it?

STEP 6: Answer the following questions about engaging in specific assessment-related actions in the future. Respond by simply writing "yes" or "no" for each question.

- Do you predict you will articulate malleable and feasible student learning or development outcomes?
- Do you predict you will use evidence to create effective programming?
- Do you predict you will establish a way to measure your student learning or development outcomes?
- Do you predict you will collect implementation fidelity data?
- Do you predict you will collect outcomes data?
- Do you predict you will analyze and interpret implementation fidelity and outcomes data?
- Do you predict you will use results to make changes to programming?

Thank you!