



Contents

Introduction.....2

Planning Logistics.....4

Travel Grants.....5

Lessons Learned.....6

Materials.....7

Appendix A: 4-VA Collaborative Research Grant Proposal.....8

Appendix B: Correspondence leading up to the event.....11

Appendix C: Assessment.....15

Appendix D: Assessment Report.....19

Presentation: Telling Our Stories: Using Assessment Data for Learning and Improvement.....23

Handout 1: And the Award Goes To.....104

Handout 2: Threats to Internal Validity.....106

Handout 3: Let's Get Real.....108

Handout 4: Breakout Activity on Practical Challenges.....111

Facilitator's Guide for Breakout Activity on Practical Challenges.....113

Presentation: Elements of Evidence-Based Storytelling.....115

Handout 5: Evidence-Based Storytelling Activity.....141

Facilitator's Guide for Evidence-Based Storytelling Activity.....152



Introduction

On Friday, April 19th, 2019 the Virginia Assessment Group hosted a spring drive-in telepresence conference, entitled *"Telling our stories: Using assessment data for learning and improvement."* The one-day conference was funded by a 4-VA collaborative research grant, and involved a planning team representing four public universities, a community college, and the State Council of Higher Education in Virginia. The event was a hybrid conference with participants meeting across six locations: George Mason University, James Madison University, Old Dominion University, University of Virginia, Virginia Commonwealth University (Cisco Headquarters), and Virginia Polytechnic Institute. Participants at each location then met virtually with participants from the other locations, via 4-VA telepresence technology. The event addressed the mission of 4-VA by leveraging the expertise of professionals across the nation to address the need for high quality professional development with low travel costs for faculty and assessment professionals. The 168 conference registrants represented 50 organizations: 31 universities, 15 community colleges, and 4 professional organizations.

The day-long conference began with a welcome from the Virginia Assessment Group president, Ryan Otto (Roanoke College) at the Virginia Tech location, and review of agenda by Kelsey Kirland from Old Dominion University. The morning workshop was presented by James Madison University Assessment and Measurement doctoral students, Andrea Pope and Caroline Prendergast, Psychological Sciences master's student, Morgan Crewe, and faculty member, Jeanne Horst. The morning workshop, entitled *"Can we back up that claim? Making important data collection design decisions"* addressed the appropriate inferences that can be drawn from assessment data collection designs. Participants grappled with how to make appropriate inferences from the data collection designs that are possible given common constraints. In the afternoon, Jodi Fisler (SCHEV), and Gianina Baker (NILOA) presented a workshop entitled *"Evidence-based storytelling."* Participants then viewed a video produced by Jillian Kinzie (NILOA), illustrating examples and rationale for presenting assessment findings that tell the story of student learning. Following the video, participants engaged in an activity in which they tailored a data report to a specified stakeholder audience. Gianina Baker closed the afternoon, providing reflections and suggestions for effective evidence-based reporting.

Conference materials (including worksheets and PowerPoint slides) may be found at (copy/paste link into browser):

<https://drive.google.com/open?id=1JwLaJX9gWAvu3iB6ld8WeAXTibHdsuem>

Schedule

Topic: Telling Our Stories: Using Assessment Data for Learning and Improvement **Date:**

Friday, April 19th, 2019

Time: 8:30 a.m. – 4:30 p.m.

Location: Your closest 4-VA campus

- George Mason University, Fairfax
- James Madison University, Harrisonburg
- Old Dominion University, Norfolk
- University of Virginia, Charlottesville
- Virginia Commonwealth University (held at CISCO campus), Richmond
- Virginia Tech, Blacksburg

Schedule:

8:30 a.m. Check in

9:00 a.m. Welcome from Virginia Assessment Group President

9:30 a.m. Workshop 1 “Answering the Right Questions”

12:00 p.m. Lunch

1:00 p.m. Workshop 2 “Evidence-Based Storytelling” by NILOA

4:00 p.m. Wrap up

Audience: We invite higher education professionals who have significant responsibility for learning outcomes assessment. Travel grants for our colleagues at community colleges will be available.

Cost per person: FREE

The Spring Drive-in is funded by a 4-VA Collaborative Research Grant and brought to you by the Virginia Assessment Group.

Please cite as: Prendergast, C., Pope, A., & Horst, S. J. (2020, February). *Evidence-based storytelling toolkit: Example from the Virginia Assessment Group and James Madison University*. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Acknowledgments: Morgan Crewe and the Virginia Assessment Group planning team

Planning Logistics

The logistics were managed by a team of eight Virginia Assessment Group board members that included representation from Virginia private colleges, four-year public universities, two-year colleges, and the State Council for Higher Education in Virginia (SCHEV). Beginning in December 2018, the team held regular virtual meetings – weekly at some points – to plan the April 2019 event. Personnel in the 4-VA grant office provided logistical support, in terms of scheduling, technology expertise, and dispersion of finances across the six locations. The following provides an overview of the logistics that future planners may wish to consider.

Hosts. Because the event was conducted across six sites, it was important to assign a host for each location. The host was employed at one of the six locations who would then collaborate with a planning team member and manage details at that location.

Room scheduling. The 4-VA office reserved rooms at each of the six locations. On-site hosts previewed the rooms to ensure the seating capacity and layout, availability of rest rooms, suitability for lunch, and available technology. Some hosts reserved an additional room for eating lunch.

WIFI. The hosts at each location reserved WIFI for participants.

Parking. When necessary, hosts secured parking passes for participants visiting from external locations. Upon registration, directions and maps of available parking were provided for each location.

Food. As a result of grant funding, we were able to provide breakfast and lunch to participants. The 4-VA personnel transferred grant monies to offices at each location. Location hosts then selected and ordered the food.

Registration. Registration was managed centrally through the Virginia Assessment Group conference registration software. Participants were able to select the site they planned to attend, and then received information tailored to that specific location. As part of the registration, we asked about dietary needs and any necessary accommodations.

Dispersion of grant money. The grant money was dispersed by the 4-VA office to each of the respective six locations. Some universities required that all monies be funneled through a central sponsored programs office, necessitating host training to receive the money. However, other universities managed the money through a 4-VA office and dispersed the money through directly through a budget line code.

Technology coordination. Several weeks prior to the event, team members, presenters, and on-site hosts met with 4-VA personnel for a test run-through. This was a key activity for developing familiarity with the technology. Logistics, such as silencing microphones, sharing screens, transitioning across locations, and other details were covered. The test run gave presenters a chance to become familiar with the technology, putting them at ease for the actual day. Having a knowledgeable 4-VA staff member lead the test run was critical.

Advertisement. Advertisement for the event was primarily conducted by email and web presence. Two months before the event, a Save the Date was disseminated to a list of Virginia assessment professionals (members and non-members of the Virginia Assessment Group) and on the Virginia Assessment Group webpage.

Travel Grants

Available travel money for assessment professionals is not always equitable across institutions. For the year that the Drive-in was offered, the community colleges were under a restricted travel budget. Despite the fact that Drive-in participants could select their closest location, travel expenses were still incurred. Therefore, a portion of the grant included money for community college participants.

The 4-VA grant permitted a \$50 travel stipend for fifteen community college participants. The money was dispersed to the Virginia Community College System and a travel grant was awarded to the first person to register from each of the community colleges.

The following email was sent to recipients of the grants:

Hi _____,

On behalf of the Virginia Assessment Group, we are excited to have you register for the Spring Drive-in, Telling Our Stories: Using Assessment Data for Learning Improvement on April 19, 2019. This event is made possible by a 4-VA Collaborative Research Grant.

As a part of the grant, we are able to offer a \$50 travel grant to one individual from 15 of the state community colleges. Because of your timely registration, you qualify for one of these travel grants, which will be disbursed by the Virginia Community College System during your reimbursement process. Awarding of the travel grant is dependent on your attendance at one of the 4-VA Campuses. If you do not attend, the grant will be awarded to another community college member.

If you have questions about the reimbursement process please contact Justin Horton (CCed), a V.A.G board member at Thomas Nelson Community College.

We look forward to seeing you on April 19th!

Lessons Learned

Plan carefully when using new forms of technology. This workshop was unique in its ability to simultaneously connect participants at six sites across Virginia. However, the benefits of the telepresence technology were accompanied by some unexpected complications. The facilitators at each of the six sites participated in test calls in the week prior to the event to ensure functioning of the telepresence technology and to acquaint themselves with the quirks of each location. However, a number of technological issues nonetheless occurred during the day of the event. Most issues were easily remedied, but others required quick thinking and adaptation from each of the facilitators. We recommend that facilitators of telepresence-based workshops become extremely familiar with relevant technology prior to the planned event. We also recommend developing contingency plans for common issues (e.g., spotty internet connection, audiovisual glitches, etc.). Ensuring that facilitators at all sites have each other's contact information handy is also important so that facilitators can contact each other if a problem arises at one of the sites.

Prepare for travel issues. The single-day format of the workshop meant that each participant needed to drive to and from the site in order to participate. While this decreased the cost of participation, it also meant that many participants needed to spend a significant amount of their days traveling. Unfortunately, the event coincided with a severe storm in part of the state. School closures and flooding forced some participants to leave after lunch. Of course, even the best workshop facilitators are unable to control the weather. However, this taught us an important lesson about ensuring all participants have access to the materials, even if an unexpected event causes some participants to leave early.

Always have a backup plan. Multi-site workshops, like this one, rely heavily on technology. While technology enables new forms of collaboration and professional development, it also opens up new opportunities for plans to go awry. It is therefore important to have a backup plan for each of the activities and presentations. For example, we should have made sure that a facilitator at each site was ready to provide each presentation or facilitate each activity if telecommunications went down. Even if the workshop only lasts a day, participants have devoted precious time to attending. It is therefore the responsibility of the facilitators to ensure that time is well-spent.

When trying something new, be sure to welcome feedback from participants. The novel approach to this workshop was exciting, but we also know that our first attempt left room for improvement. We asked participants to reflect on their experiences in the workshop and specifically requested feedback on the telepresence-based approach to multisite professional development. The feedback we received will be extremely useful if we attempt to conduct a similar workshop in the future.

Materials

- AM Materials
 - 1. Presentation –Morning Workshop
 - 2. Handout – And the Award Goes To
 - 3. Handout – Threats to Internal Validity
 - 4. Handout – Let’s Get Real
 - 5. Handout – Breakout Activity on Practical Challenge
 - 5a. Facilitation guide for breakout activity
- PM materials
 - 6. Presentation – Kinzie Communicating Value
 - 7. Presentation – NIOLOA Elements of Evidence-Based Storytelling
 - 8. Handout – Evidence-Based Storytelling Activity
 - 8a. Facilitation guide for evidence-based storytelling activity

Appendix A: 4-VA Collaborative Research Grant Proposal

October 2018

1. Project Summary (1-2 pages)

a. Title: “Meeting Assessment Professional Development Needs across Virginia: A Virtual Drive-In”

b. Director (PI): S. Jeanne Horst, *Center for Assessment & Research Studies*

c. Partners/Collaborators: Collaborators are assessment leaders from across the state of Virginia who serve as Virginia Assessment Group board members, State Council of Higher Education for Virginia (SCHEV) leaders, or graduate students in Assessment and Measurement at James Madison University (JMU). The collaborators are well-positioned to identify professional development needs throughout the state, and include (in alphabetical order): Morgan Crewe, *JMU*; Kristy L. Crickenberger, *Washington and Lee University*; Jodi Fidler, *SCHEV*; Stephanie Foster, *George Mason University*, *Virginia Assessment Group past-president*; Justin Horton, *Thomas Nelson Community College*; Kelsey Kirland, *Old Dominion University*; Jason Lyons, *Christopher Newport University*; Ryan Otto, *Roanoke College*, *Virginia Assessment Group president*; Andrea Pope, *JMU*; Caroline Prendergast, *JMU*; and, Linda Townsend, *Longwood University*.

Support from: Keston H. Fulcher, *Center for Assessment and Research Studies*

d. Type of Grant: This proposal includes a request for funding in the form of a collaborative research grant. This proposal meets the mission of 4-VA by leveraging the expertise of professionals across Virginia to address the need for high quality professional development with low travel costs for faculty and assessment professionals. This event will be held locally to maximize university and community college resources. We will leverage faculty and administrator expertise by hosting the conference at 5-6 sites across Virginia via telepresence. We will increase the capacity of assessment professionals by hosting workshops on shared areas identified for professional development. Partners are members of the [Virginia Assessment Group](#) board and the State Council of Higher Education for Virginia. The Virginia Assessment Group is “dedicated to the goals of promoting the continued high quality of higher education in the Commonwealth of Virginia through assessment practices. Virginia Assessment Group serves as a forum for the expression of ideas about assessment practices, and as a network for communication and collaboration among public and private institutions and state agencies. Virginia Assessment Group also functions to promote the professional development of its members.” With this funding, we believe we will be able to make a valuable and impactful contribution to the high quality education offered by VA colleges and universities.

2. Narrative (1-3 pages)

a. Project activity summary. The current proposal describes a request for funding to support a Virtual “Drive-In” day-long conference during spring or summer 2019. The proposed one-day event would target the needs of assessment professionals across the state of Virginia. The day-long (9:00 a.m. to 3:00 p.m.) conference would include short educational sessions on topics relevant to assessment professionals. Specific topics for conference workshops will be identified by the grant partners, based upon feedback from Virginia Assessment Group members in November 2018. The proposed conference would be hosted at five or six 4-VA locations. Each location would provide scheduled presentations throughout the day-long conference, and opportunity for collaboration across locations would be provided.

Need for Professional Development. As the practice of higher education assessment flourishes, there is increasing demand to meet professional development needs (Ariovich, Bral, Gregg, Gulliford, & Morrow, 2018). Entry into the profession is from diverse backgrounds, resulting in a variety of skill sets for assessment processes, data collection, analysis, interpretation, reporting, and use of results. Consequently, there are ongoing needs for professional development in assessment.

A recent analysis of data from two large surveys of assessment professionals identified specific professional development needs (Ariovich et al., 2018). Preferred methods for meeting professional development needs were conferences and webinars. The current proposal aims to combine the favored methods into a hybrid conference format by leveraging 4-VA telepresence rooms and Webex technology. Given the current fiscal constraints that have resulted in limited or eliminated travel funding for our community college colleagues, a small travel grant (\$50) is proposed for fifteen community college representatives who wish to participate in the on-site event. Graduate student involvement will also be encouraged and welcomed, in order to encourage and develop the future generation of assessment professionals.

b. Research questions. The following research questions will be addressed: 1) To what extent does a one-day virtual professional development event influence the perceived knowledge, skills, and confidence of assessment professionals and faculty in Virginia? 2) What benefits does a blended conference offer? 3) What do assessment professionals and faculty in Virginia colleges and universities identify as needs for professional development?

c. Research plan (framework/methods/phases/specific approach). An evaluation will be administered at two time-points: 1) at the close of the seminar, we will gather open-ended and quantitative information on each of the three research questions; and 2) at six months, we will gather follow-up information regarding any changes implemented as a result of the conference. Coding and analysis of all data will be completed prior to January 31, 2020. Recommendations for future conferences and events will be made, based upon the findings.

d. Assessment plan. Post-seminar evaluations will be disseminated via Qualtrics survey administration software. Participants will be asked to reflect on the knowledge and skills gained from the conference, via several open-ended questions. Likert-type questions about event organization and content will also be included.

e. Outcomes and deliverables. Outcomes for this conference are three-fold: 1) To increase the number of local professional development opportunities; 2) To decrease travel expenditures by hosting a virtual conference at local sites; and 3) To build capacity of assessment professionals through workshops on improving student learning.

3. Timeline

Timeline			
November-December 2018	January-March 2019	April-June 2019	July 2019-January 2020
Determine topics from VAG members; invite speakers (confirmation by 12/31/18); select date for event; discuss logistics	Finalize schedule and logistics; reserve rooms at the 4-VA locations; disseminate information	One day conference held virtually at five or six 4-VA locations; conduct post-conference assessment (Qualtrics survey)	Conduct 6-month follow-up assessment; aggregate findings & make recommendations for future conferences, based upon findings.

4. Budget

Budget			
Item	Amount	Cost	Total
Breakfast	150	\$ 12.00	\$ 1,800.00
Lunch	150	\$ 16.25	\$ 2,437.50
Travel Grants	15	\$ 50.00	\$ 750.00

\$ 4,987.50

In-kind support: Planning sessions (collaborators), data collection and analysis, and reporting. We would like to ask for use of 4-VA technology classrooms (4-VA), technology support, and help in contacting Virginia Tech and University of Virginia. We currently have partner representatives from the remaining four 4-VA schools.

Complementary funds: The current budget caps participation at 150, which would fund food costs for people at all locations. Complementary funds from the other 4-VA locations would allow more people to attend the conference.

5. Proposals that involve JMU departmental resources or release time will require approval from the appropriate academic unit head: No release time is required for the full-time JMU faculty member.

Reference

Ariovich, L., Bral, C., Gregg, P., Gulliford, M., & Morrow, J. A. (2018, May). *The assessment profession in higher education: Addressing the varied professional development needs of practitioners*. Lexington, KY: Association for the Assessment of Learning in Higher Education and Watermark.

Appendix B: Correspondence leading up to the event

Save the date was emailed 2/11/19, which was two months prior to the event

Spring Drive-in

Topic: Telling Our Stories: Using Assessment Data for Learning and Improvement

Date: Friday, April 19th, 2019

Time: 8:30am - 4:30pm

Location: Your closest 4-VA Campus*

Join us at one of these drive-in sites:

George Mason University, Fairfax

James Madison University, Harrisonburg

Old Dominion University, Norfolk

University of Virginia, Charlottesville

Virginia Commonwealth University (to be held at CISCO campus),
Richmond

Virginia Tech, Blacksburg

Audience: We invite higher education professionals who have significant responsibility for learning outcomes assessment. Travel grants for our colleagues at community colleges will be available.

Cost per person: FREE

Registration opens in late February

The Spring Drive-in is funded by a 4VA Collaborative Research Grant and brought to you by the Virginia Assessment Group

Registration opened on 3/3/19, one and one-half months prior to the event.

Spring Drive-in

The Virginia Assessment Group will be hosting a Spring Drive-in. This is a one-day training and networking event that will be held at six sites around the Commonwealth, connected by telepresence technology. You will have the opportunity to build your professional skills and network with colleagues.

Topic: Telling Our Stories: Using Assessment Data for Learning and Improvement

Date: Friday, April 19th, 2019

Time: 8:30am - 4:30pm

Location: Your closest 4-VA Campus*

Join us at one of these drive-in sites:

George Mason University, Fairfax

James Madison University, Harrisonburg

Old Dominion University, Norfolk

University of Virginia, Charlottesville

Virginia Commonwealth University (to be held at CISCO campus),

Richmond

Virginia Tech, Blacksburg

Proposed Schedule:

8:30am - Check in

9:00am - Welcome from Virginia Assessment Group President

9:30am - Workshop 1 "Answering the Right Questions...Can we back that claim up?"

12:00pm - Lunch

1:00pm - Workshop 2 "Evidence-based Storytelling" by NILOA

4:00pm - Wrap up

Audience: We invite higher education professionals who have significant responsibility for learning outcomes assessment. Travel grants for our colleagues at community colleges will be available.

Cost per person: FREE

The Spring Drive-in is funded by a 4-VA Collaborative Research Grant and brought to you by the Virginia Assessment Group

When registering, participants received the following email.

VAG Spring Drive-In 2019

2019 Spring Drive-in JMU Registration

The Spring Drive-In is a one-day training and networking event to be held at six sites around the Commonwealth, connected by telepresence technology. This event is funded by a 4-VA Collaborative Research Grant and brought to you by the Virginia Assessment Group. This registration is for participation at the James Madison University site.

Friday, April 19, 2019 from 8:30 AM to 4:30 PM EDT

James Madison University

Lakeview Hall, Room 1165

298 Port Republic Rd

Harrisonburg, VA 22807

Thank you again for registering for the Spring Drive-in event at James Madison University. This email is confirmation of your successful registration. If any of the information displayed below is incorrect, please contact us at vaassessgroup@gmail.com.

The Spring Drive-in is funded by a 4-VA Collaborative Research Grant and brought to you by the Virginia Assessment Group. There will be no fees for this event.

Proposed Schedule:

8:30am – Check in

9:00am – Welcome from Virginia Assessment Group President

9:30am – Workshop 1

12:00pm – Lunch

1:00pm – Workshop 2

4:00pm – Wrap up

The Virginia Assessment Group website will host the most updated information about the event:
<http://virginiaassessment.org/conference/spring-drive-in/>

Closer to the event, you will receive an email with information about parking, building and room information, getting around campus, and more. If you have any questions leading up to the event, please contact us at vaassessgroup@gmail.com.

We look forward to seeing you in April!

If you would like to live tweet during this event, please use the hashtag #2019DriveinVA.

Participants were asked to complete an assessment before and after the event.

A copy of the pre-event assessment email is pasted, below.



Dear ,

We are looking forward to seeing you this Friday (April 19th) for the Virginia Assessment Group's Drive-In workshop, *Using Data to Tell Our Stories*. In preparation for the event, we would like to learn a little more about you. Please complete the brief survey below prior to Friday's workshop. This should take no more than two minutes of your time.

Follow this link to the Survey:

[Take the Survey.](#)

Or copy and paste the URL below into your internet browser:

http://cnu.az1.qualtrics.com/jfe/preview/SV_0qzn9aqtrsZ7CjX?Q_CHL=preview

Thank you in advance for your time. We look forward to seeing you soon!

Best regards,

The Virginia Assessment Group Drive-In Planning Team

Appendix C: Assessment

The following is the post-event assessment, which includes pre-event items, plus several post-event items.



Did you attend the *morning* session of the Spring Drive-in Event?

☐ Yes

☐ No

Did you attend the *afternoon* session of the Spring Drive-in Event?

☐ Yes

☐ No

What is your primary role?

☐ Full-time faculty

☐ Part-time faculty

☐ Graduate student

☐ Assessment professional

☐ Academic support professional (includes Libraries, Writing Center, etc.)

☐ Staff

☐ Other

How confident are you in your ability to manage the collection of assessment data?

☐ Very Confident

☐ Confident

☐ Somewhat Confident

☐ Slightly Confident

☐ Not Confident

How confident are you in your ability to communicate the results and value of assessment?

☐ Very Confident

☐ Confident

☐ Somewhat Confident

☐ Slightly Confident

☐ Not Confident

To what extent did the Spring Drive-in Event help you learn or think differently about each of the following?

	A great deal	Quite a bit	Somewhat	Very little	Not at all	Not applicable
Your approach(es) to data collection	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Assessment or research design	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Reporting results	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Using results	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Think about the assessment data collection challenges you commonly face. What is one method/strategy you learned during the drive-in that will help you address these challenges?

This event blended a physical location ("drive-in") with telepresence, rather than restricting the meeting to one location. What did you find useful about this approach?

Were there aspects of the blended telepresence format that you did not find effective?

☐ Yes (please explain)

☐ No

What additional professional development topics would you like to see at future events?

Do you have any additional comments or feedback?

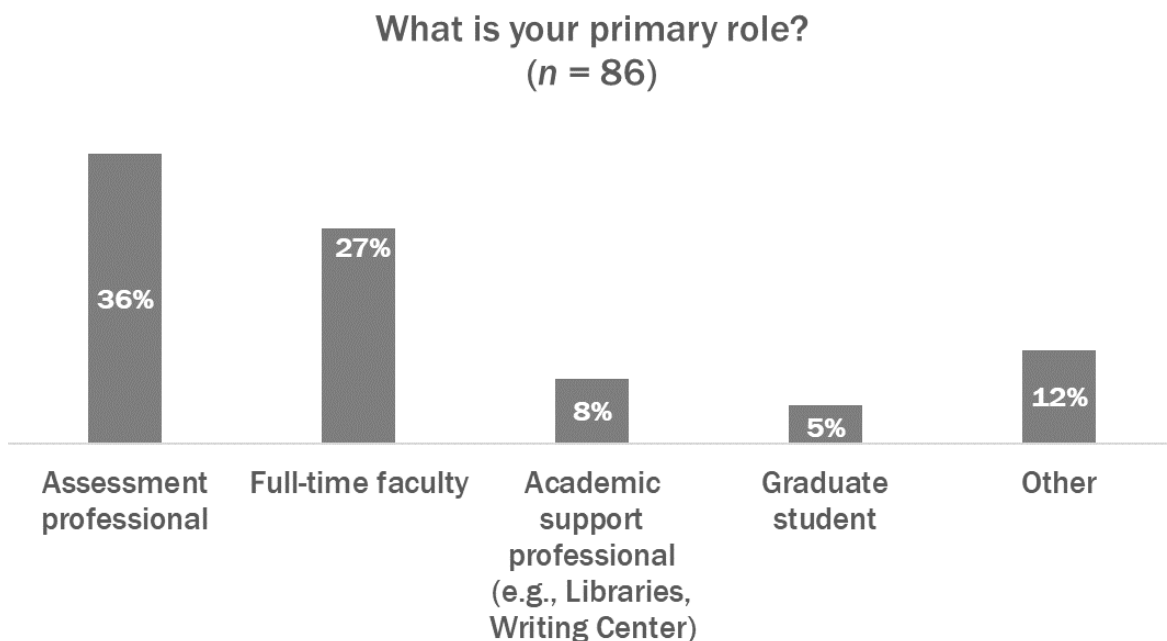
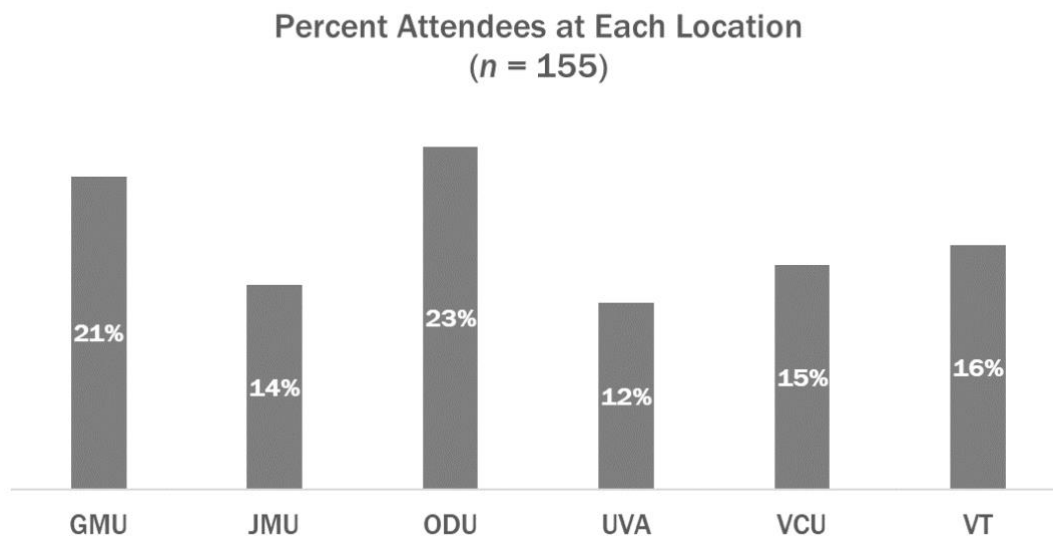


We thank you for your time spent taking this survey.
Your response has been recorded.

Appendix D: Assessment Report

Who attended the Drive-In?

Just under half of the participants attended at Old Dominion University (ODU) and George Mason University (GMU), combined. The remaining participants were distributed across the other four locations. The majority of survey respondents ($n = 86$) were Assessment Professionals or Full-Time Faculty members. Others included student affairs professionals, accreditation professionals, college deans, and other administrators.



Did respondents' self-reported confidence increase after attending the Drive-In?

Two survey questions were included to address the planning question "To what extent does a one-day virtual professional development event influence the perceived knowledge, skills, and confidence of assessment professionals and faculty in Virginia?" Respondents rated their confidence in two assessment tasks (i.e., 1) managing the collection of assessment data, and 2) communicating the results and value of assessment) before and after the Drive-in. List wise deletion was used to account for missing data, thus the sample size considered below includes data from only 39 respondents (only 23% response rate) with complete data at both time points. Respondents used the following rating scale: 1 = "Not Confident", 2 = "Slightly Confident", 3 = "Somewhat Confident", 4 = "Confident", 5 = "Very Confident". Higher scores represent higher level of confidence.

Results from these items are reported in the table below. Given the issues with attrition and small sample size, we cannot rule out plausible threats to validity, and causal interpretations should be made cautiously. See Tables 1 and 2 for a summary of parametric and non-parametric analyses (respectively).

Table 1

Pre-Post Self-Reported Confidence Gains (N = 39)

	Mean (SD)		Mean Difference	Cohen's <i>d</i>	Significance
	Pre-Drive In	Post-Drive In			
Confidence in ability to manage the collection of assessment data	3.54(.97)	3.85(.78)	0.308	0.47	$t(38) = 2.93, p = .006$
Confidence in ability to communicate the results and value of assessment	3.51(.97)	3.97(.74)	0.462	0.48	$t(38) = 3.06, p = .004$

Table 2

Non-parametric Analysis of Difference from Pre-to-Post-Test (Wilcoxon Signed Rank)

	Mode		Neg. Differences	Pos. Differences	Ties	Significance
	Pre-Drive In	Post-Drive In				
Confidence in ability to manage the collection of assessment data	4	4	2	12	25	$p = 0.13$
Confidence in ability to communicate the results and value of assessment	3	4	4	17	18	$p = .007$

To what extent did the Spring Drive-in Event help you learn or think differently about each of the following? Following the Drive-in respondents were asked to what extent the Spring Drive-in helped them to think differently in four different areas:

- 1) own approaches to data collection,
- 2) assessment or research design,
- 3) reporting results, and
- 4) using results.

Respondents used a 1-6 point scale where 1 = “Not applicable”, 2 = “Not at all”, 3 = “Very little”, 4 = “Somewhat”, 5 = “Quite a bit”, 6 = “A great deal”. Thus, the higher the score, the more impact the Drive-in had on respondents’ thinking processes. The mean levels of each area are reported below. Values of 1 (“Not Applicable”) were not counted in the averages. Average scores were between 4 - “Somewhat” and 5 - “Quite a bit”. No one selected option 2 – “Not at all.” Tables 3 and 4 present the findings.

Table 3
Post-Drive-In Responses (n = 44)

Item	Mean	SD	95% CI	
			Lower	Upper
Own approaches to data collection	4.66	.78	4.42	4.89
Assessment or research design	4.66	.81	4.41	4.90
Reporting results	4.80	.76	4.56	5.03
Using results	4.72	.76	4.49	4.96

Table 4
Frequency (Percent) for Each Response Option

Item	N/A	Not at all	Very little	Somewhat	Quite a bit	A great deal
Own approaches to data collection	6	0	0	2	17	19
	(13.6%)	(0%)	(0%)	(4.5%)	(38.6%)	(43.2%)
Assessment or research design	5	0	0	4	12	23
	(11.4%)	(0%)	(0%)	(9.1%)	(27.3%)	(52.3%)
Reporting results	7	0	0	2	12	23
	(15.9%)	(0%)	(0%)	(4.5%)	(27.3%)	(52.3%)
Using results	7	0	0	1	17	18
	(16.3%)	(0%)	(0%)	(2.3%)	(39.5%)	(41.9%)

Open-Ended Feedback

What are your common assessment data collection challenges?

Prior to the Drive-In, respondents were asked to provide their thoughts on the most common challenges in assessment data collection. From the 65 comments provided, several themes emerged. About a third of comments ($n = 19$) were related to *getting support from faculty members when collecting data*.

Comments related to faculty members ranged from simply gathering the data to working with faculty to create meaningful assessment tools. *Lack of experience or lack of confidence in collecting data* emerged from 11 responses. Some respondents were new in their roles, others expressed *lack of familiarity with*

best practices. Another 10 responses were related to *lack of data or not obtaining quality data.* Other themes had to do with *lack of time and resources, reporting assessment data, storing data efficiently, and creating or using quality tools to assess.* These themes could potentially provide information for future professional development events.

Think about the assessment data collection challenges you commonly face. What is one method/strategy you learned during the drive-in that will help you address these challenges?

Following the Drive-In event, respondents were asked what methods or techniques they learned at the Drive-in that could help them in their data collection practices. Based on the feedback, they reported learning about threats to validity and researcher designs. Five respondents specifically mentioned the use of pre and post methods. Storytelling was another theme mentioned several times, along with small sample size designs, and considering alternative explanations to assessment results.

This event blended a physical location ("drive-in") with telepresence, rather than restricting the meeting to one location. What did you find useful about this approach?

Slightly more than one-half of respondents noted the travel convenience and how easy was to find a location that fit them, which encouraged them to attend the Drive-in. Another advantage of this approach was that respondents felt a part of a bigger community, because they were able to communicate with other respondents from around the state. Respondents reported that having a site moderator and having small group discussion was extremely helpful before joining the bigger discussion.

Were there aspects of the blended telepresence format that you did not find effective?

This question was addressed in both selected-response (Yes/No) and open-ended response format. Fifty percent ($n = 21$) selected "Yes" and 50% ($n = 21$) selected "No." Those who selected "Yes," were offered the option to explain. The majority of those selecting "Yes" mentioned technical issues, mainly the audio and microphones. Feedback suggested that presenters should be at a physical location.

What additional professional development topics would you like to see at future events?

Respondents provided several topics for future developmental events. A few respondents wanted to dig deeper into curriculum maps and examples from other institutions. Others wanted to see workshops that would address practical challenges with assessment practice. Lastly, tools and methods to analyze different formats of assessment data emerged as a potential theme. There were other suggestions that involve accreditation, survey response rates, implementation fidelity, storytelling, research design limitations, and more extensive workshops on data collection.



TELLING OUR STORIES:
**Using Assessment Data for Learning and
Improvement**

**SPRING DRIVE-IN
2019**

Brought to you by

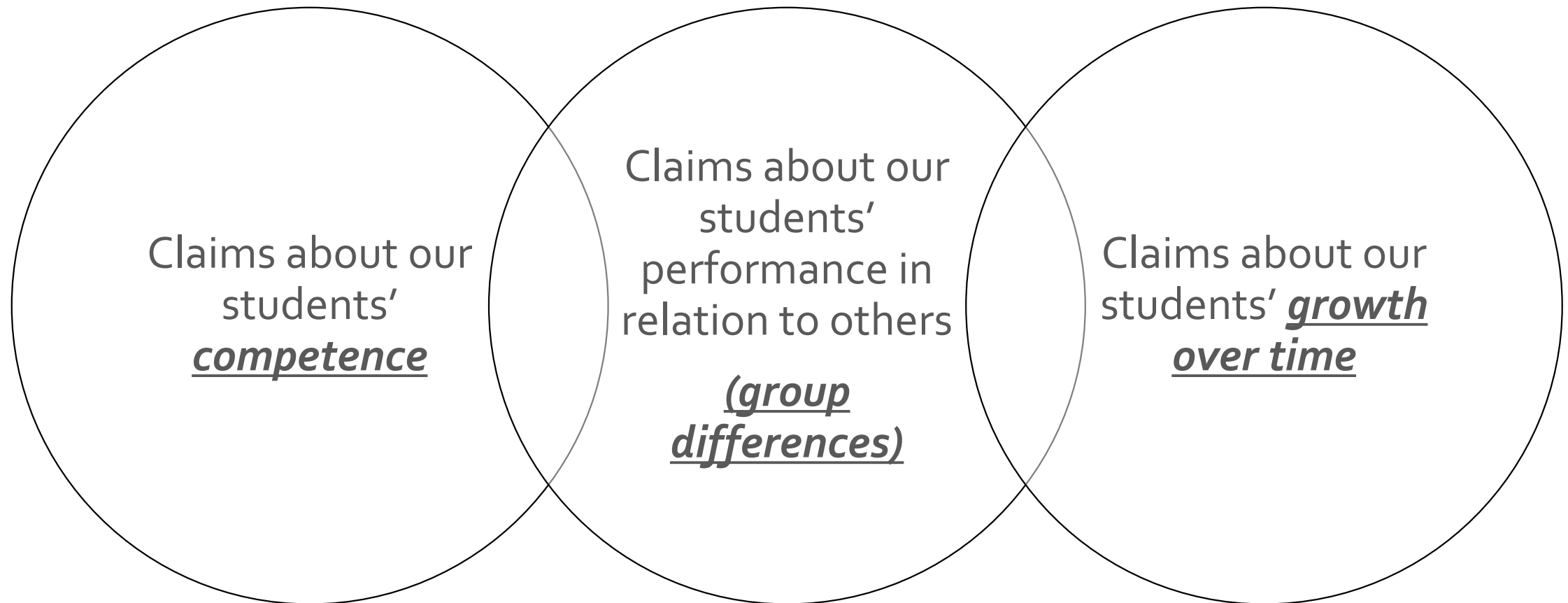


Can we Back Up that Claim?

Making Important Data Collection Design
Decisions

Andrea **Pope**, Caroline **Prendergast**, Morgan **Crewe**, & Jeanne **Horst** | James Madison University

Why do we do assessment?



A Primary Reason



And the Award Goes to...

Activity – Morning Handout #1

Morning Handout #1: And the Award Goes To...

The mission of the Center for Curricular and Co-Curricular Excellence (CCCE) is to recognize and support exemplary academic and student affairs programs that are contributing substantially to students' learning and development.

For the last five years, the CCCE has awarded one exemplary program a \$10,000 grant to support program expansion. As director of the Center, you are responsible for reviewing this year's nominations for the Program Excellence Award. The award criteria are as follows:

Demonstrating Causality

- Making statements about causality is *hard*. It requires eliminating all other reasonable explanations for student growth/achievement
- *Fortunately*, we can eliminate many of the most common alternative explanations for student achievement using strong data collection designs
- *Unfortunately*, these strong data collection designs are often infeasible for educators—whether due to time, money, logistics, or ethics



A solid teal square is positioned on the left side of the slide, partially overlapping a dashed teal rectangular frame.

Making Causal Claims

The Ideal Data Collection Design

A solid gray square is positioned on the right side of the slide, partially overlapping the dashed teal rectangular frame.

Data Collection Design Basics

When will the data be collected?

- **Are the data collected at multiple time points?**

Who will the data be collected from?

- **Is there a comparison group?**

How is participation in the program determined?

- **Are students randomly assigned to groups?**

Data Collection Design Basics

When will the data be collected?

- **Are the data collected at multiple time points?**

Who will the data be collected from?

- **Is there a comparison group?**

How is participation in the program determined?

- **Are students randomly assigned to groups?**

Single vs. Multiple Time Points

Alternative Spring Break Program (ASB)

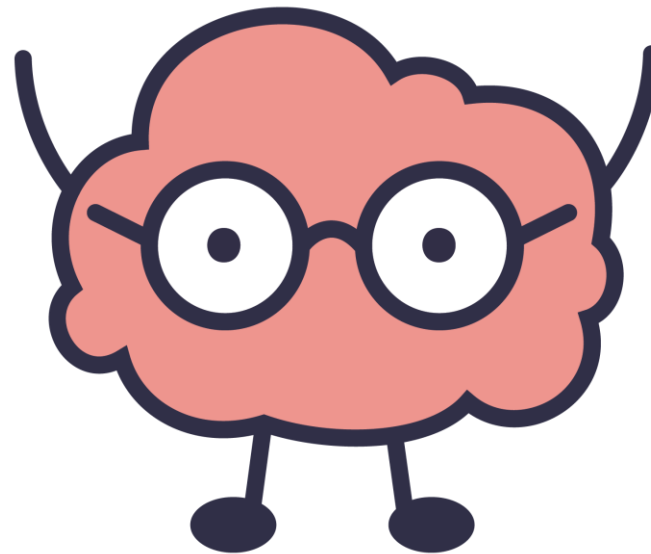
Learning Outcome: Civic-Mindedness

Program Description: Each year, CSL offers a variety of week-long service breaks to locations in the United States and abroad. These alternative breaks provide opportunities for participants to engage directly with community members through hands-on projects and activities. There are three important elements of ASB: education, reflection, and re-orientation.

Summary of Assessment Results: After completing ASB, students were required to complete a measure of civic-mindedness. Civic-mindedness is defined as displaying a concern for the public good or humanity as a whole. According to the survey, 87% of ASB participants demonstrated high levels of civic-mindedness, 9% of ASB participants demonstrated moderate levels of civic-mindedness, and only 4% of ASB participants demonstrated low levels of civic-mindedness.

Alternative Explanations

The “Growth” Illusion



Were students proficient *before* the program
or intervention?

Benefits of Multiple Time Points

Allows practitioners to make claims about growth/change over time

If students are high at pretest, may be an indication that programming isn't necessary or should be targeted at a higher level

Data Collection Design Basics

When will the data be collected?

- **Are the data collected at multiple time points?**

Who will the data be collected from?

- **Is there a comparison group?**

How is participation in the program determined?

- **Are students randomly assigned to groups?**

Comparison Group vs. No Comparison Group

International Business, B. A.

Learning Outcome: Written Communication

Program Description: All students are required to take a 3-credit writing seminar the semester after enrolling in the major. Additionally, all students must complete two writing-intensive major courses. As part of these courses, students complete a semester-long writing project. Multiple drafts of the project are submitted throughout the semester and students receive detailed, formative feedback between each draft.

Summary of Assessment Results: Assessment results from the 2017-2018 academic year indicate that students entered the major with poor writing skills (1.5 out of 5 on writing rubric). After two years of coursework, including the writing seminar and writing-intensive courses, students demonstrated great improvement (scored 3.5 out of 5 on writing rubric).

Alternative Explanations

Maturation Effect



Were changes due to normal developmental processes over time?

History Effect



Were changes due to some event that occurred by chance at the same time as the intervention?

Benefits of Using a Control Group

Allows practitioners to see how program participants perform compared to students who did not participate in the program

Can help eliminate other potential explanations for growth including maturation effect and history effect

May reveal situations in which programming isn't necessary (or should be targeted at a higher level) because students appear to acquire the intended knowledge/skills without it

Data Collection Design Basics

When will the data be collected?

- **Are the data collected at multiple time points?**

Who will the data be collected from?

- **Is there a comparison group?**

How is participation in the program determined?

- **Are students randomly assigned to groups?**

Random vs. Non-Random Assignment

Student Academic Success Program (SAS)

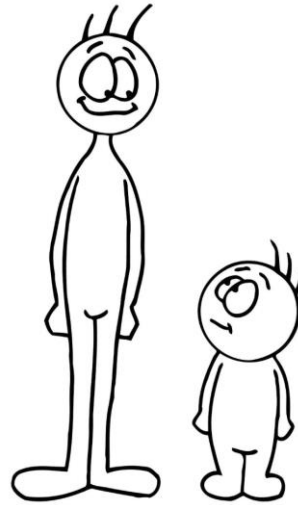
Learning Outcome: Academic Self-Efficacy

Program Description: SAS is an voluntary eight-week course designed to help students on academic probation return to good academic standing (i.e., achieve a cumulative grade point average of 2.0 or higher). This is accomplished by targeting knowledge, attitudes, and skills empirically related to academic success, such as knowledge of academic support resources, organizational skills, and academic self-efficacy (belief in one's ability to achieve in an academic setting).

Summary of Assessment Results: At the end of the fall semester, a measure of academic self-efficacy was sent to all students on academic probation. Results were presented for two groups: students who chose to complete the program (SAS participants), and students who did not. Notably, SAS participants scored two standard deviations higher than their peers who did not complete the course. In other words, SAS participants had much greater confidence in their academic abilities than students who did not complete the course.

Alternative Explanations

Self-Selection Bias



Were differences between groups on the outcome preexisting?

Benefits of Using Random Assignment

Eliminates pre-existing differences at pretest as a potential explanation for differences at posttest (i.e., addresses selection bias)

With a large enough sample, there should be no significant difference at pretest between the treatment group (program participants) and the comparison group

Can make (tentative) claims about program effectiveness for all students in the population, not just the specific students who were assessed.

Summary

Elements of an ideal data collection design:

- Pretest-Posttest (multiple time points)
- Comparison Group
- Random Assignment

Without these elements, the following threats may become plausible:

- The Illusion of Growth
- Maturation Effect
- History Effect
- Self-Selection Bias

Additional Threats to Validity

Morning Handout #2: Threats to Internal Validity

Maturation Effect: The observed effect is due to normal developmental processes or changes over time, not the program.

- Program A (implemented during the first semester of college) claims to have increased students' sense of independence. However, studies show students naturally gain more independence during their first semester of college even without an intervention.

History Effect: The observed effect is not due to the program, but to some other unaccounted for event.

- Program B claims to have reduced instances of sexual assault on campus. However, sexual assault prevention is a university-wide initiative and upon further investigation, the facilitators of Program B realize their participants also received programming related to sexual assault prevention in their residence halls. Could the reduction in instances of sexual assault be due to this residence life program instead?

Designs for Desired Inferences

Desired Inference	Data Collection Design
Inferences about Students	
Posttest Only Design	Student(s) demonstrated a certain level of competency
Pretest-Posttest Design	Student(s) increased their knowledge/skills from the beginning of the program to the end of the program
Control Group Design (No Pretest)	Student(s) in the program demonstrated greater competency than students who were not in the program
Pretest-Posttest w/ Comparison Group Design	Student(s) in the program grew more than students who were not in the program
Inferences about the Program	
Pretest-Posttest w/ Comparison Group & Random Assignment Design	(Tentative) The program increased students' knowledge and skills

A solid teal square is positioned on the left side of the slide.

Let's Get Real

Moving from “Ideal” to “Real”

A dashed teal rectangular border frames the central text area on the right side of the slide.A solid grey vertical bar is located on the far right edge of the slide.

“The cold reality is that decisions will be made, policies developed, and practices implemented regardless of the availability of assessment results. So the question becomes, When it comes to the usefulness of a study for policy and practice, is a study with substantial limitations better than no study at all?”
– Upcraft & Schuh, 2002

Upcraft, M. L., & Schuh, J. H. (2002). Assessment vs. research: Why we should care about the difference. *About Campus*, 7(1), 16-20.

When is a pretest...

- Not practical?
 - When an intervention does not have a clear beginning or end
 - When the participant group isn't clear
 - Brief interventions where testing effects could be problematic or insufficient time for testing
- Not ethical?
 - When the test could act as a barrier to entry to a useful/necessary intervention

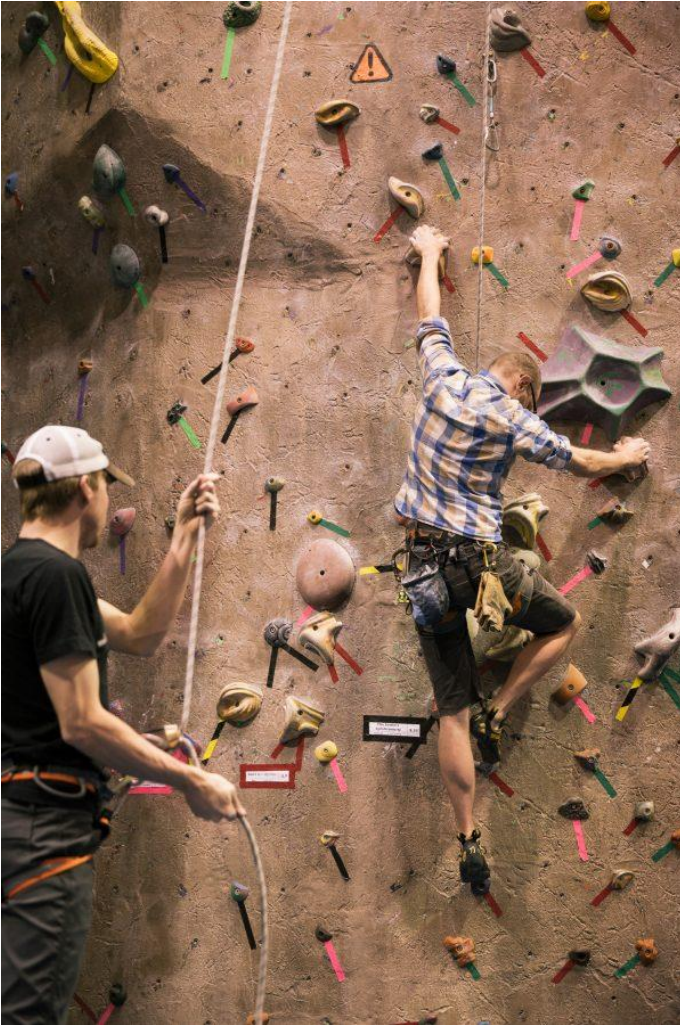
When is a comparison group...

- Not practical?
 - When a comparable group is unavailable or difficult to access
 - When a motivated group is not possible to access
- Not ethical?
 - When delaying or withholding treatment would cause harm

When is random assignment...

- Not practical?
 - When we can't tell students what to do (choice of major, choice of classes, etc.)
- Not ethical?
 - When delaying or withholding treatment would cause harm
 - When random assignment limits autonomy or restricts justice

Scenario 1



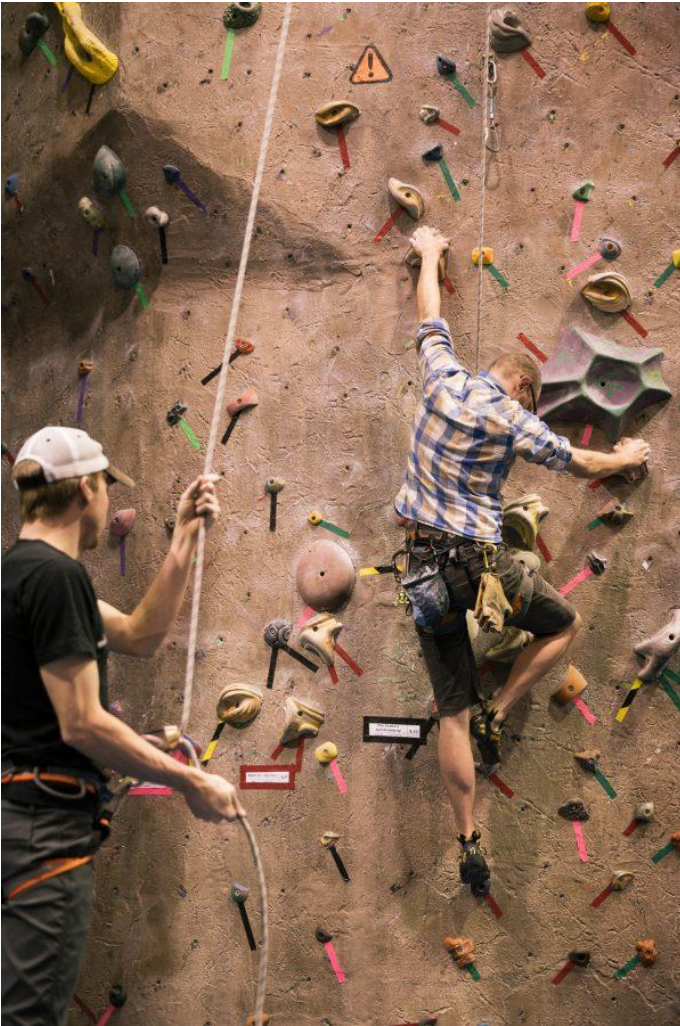
You are in charge of assessing a program through your university's recreation program intended to improve students' knowledge of proper safety techniques in belaying rock climbers. 200 students were randomly assigned to either the treatment group (an hour-long workshop each week for six weeks) or the control group (which received no intervention). At the end of the six weeks, knowledge of essential safety techniques was assessed for all students in both groups. Students who participated in the workshop displayed significantly larger increases in knowledge than students who did not participate.

What is the design?

single time point
no comparison group
no random assignment

or multiple time points
or comparison group
or random assignment

Scenario 1



You are in charge of assessing a program through your university's recreation program intended to improve students' knowledge of proper safety techniques in belaying rock climbers. 200 students were randomly assigned to either the treatment group (an hour-long workshop each week for six weeks) or the control group (which received no intervention). At the end of the six weeks, knowledge of essential safety techniques was assessed for all students in both groups. Students who participated in the workshop displayed significantly larger increases in knowledge than students who did not participate.

What is the design?

single time point

no comparison group

no random assignment

or

or

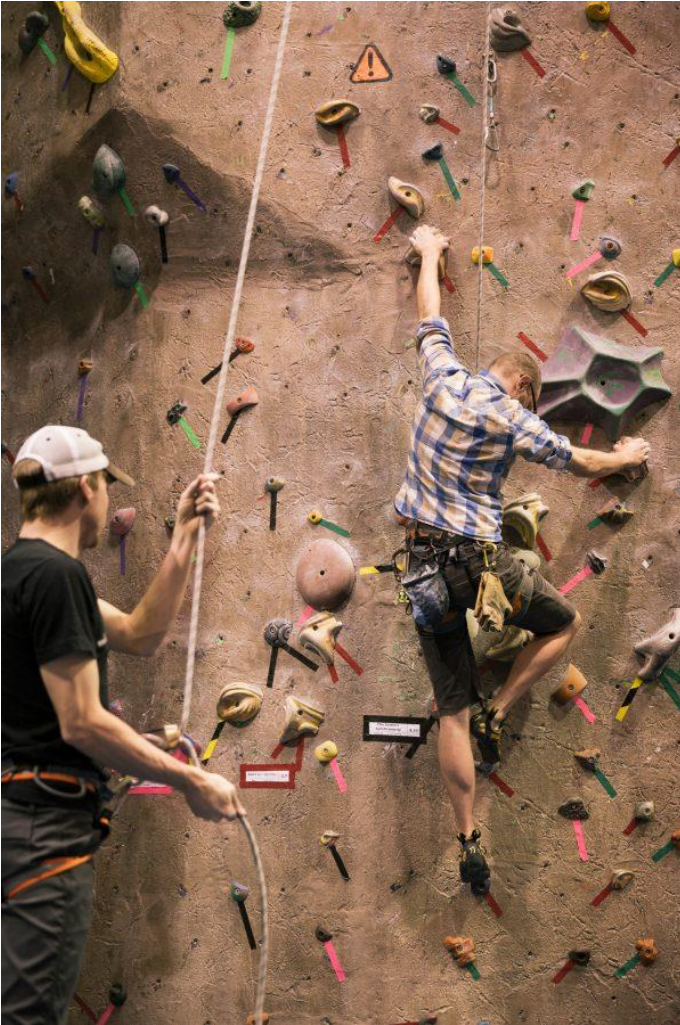
or

multiple time points

comparison group

random assignment

Scenario 1



You are in charge of assessing a program through your university's recreation program intended to improve students' knowledge of proper safety techniques in belaying rock climbers. 200 students were randomly assigned to either the treatment group (an hour-long workshop each week for six weeks) or the control group (which received no intervention). At the end of the six weeks, knowledge of essential safety techniques was assessed for all students in both groups. Students who participated in the workshop displayed significantly larger increases in knowledge than students who did not participate.

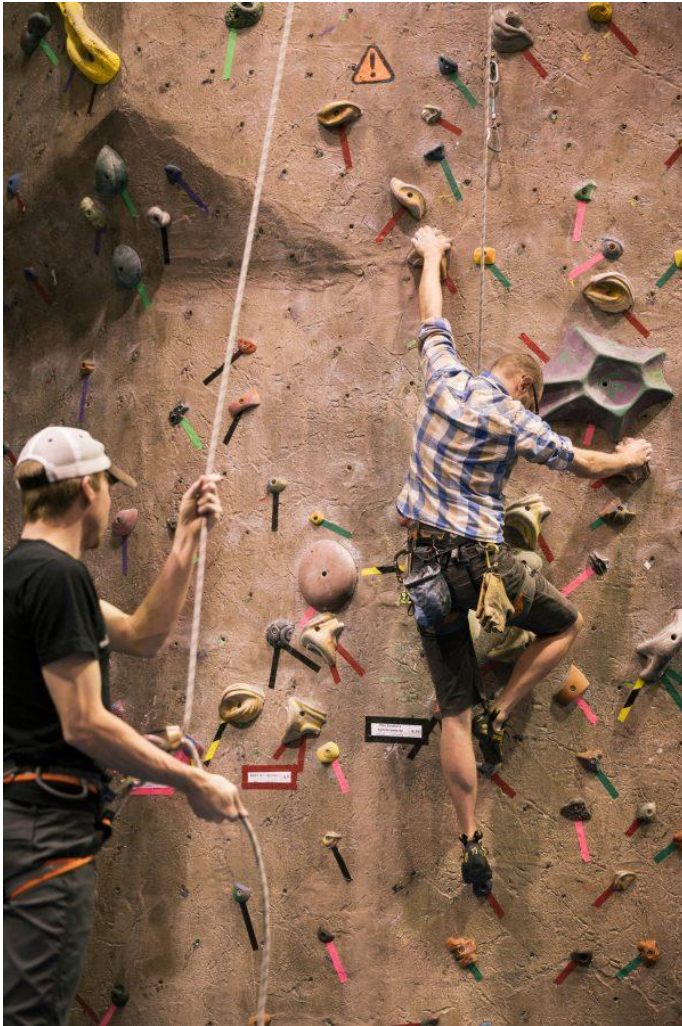
Which of these problems accompany the single time point with control group and random assignment design?

maturation

self-selection

history

Scenario 1



You are in charge of assessing a program through your university's recreation program intended to improve students' knowledge of proper safety techniques in belaying rock climbers. 200 students were randomly assigned to either the treatment group (an hour-long workshop each week for six weeks) or the control group (which received no intervention). At the end of the six weeks, knowledge of essential safety techniques was assessed for all students in both groups. Students who participated in the workshop displayed significantly larger increases in knowledge than students who did not participate.

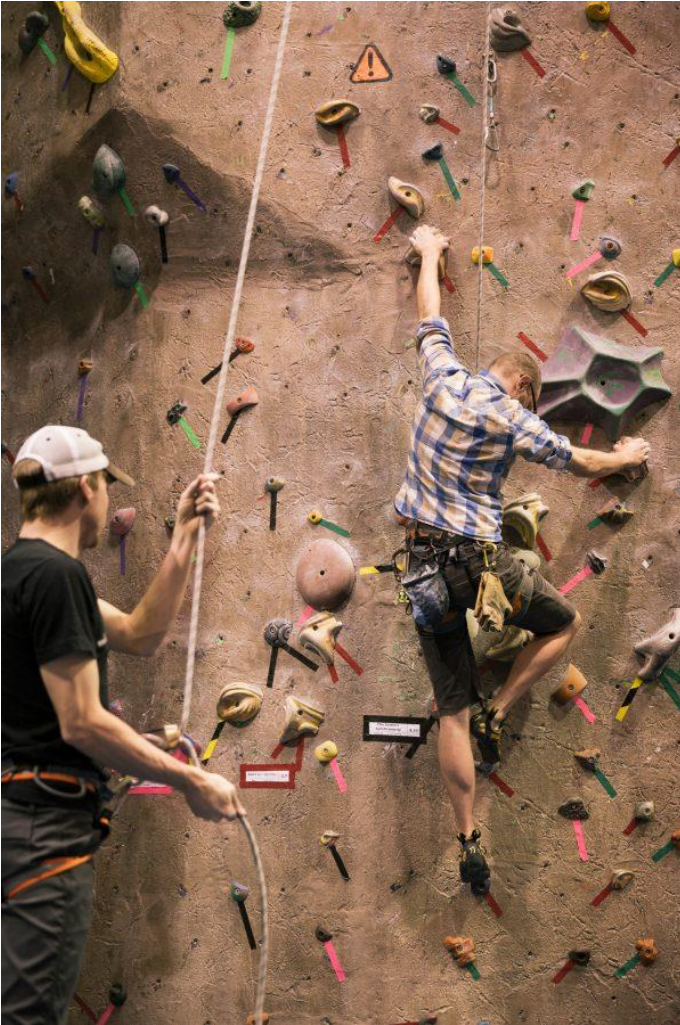
Which of these problems accompany the single time point with control group and random assignment design?

maturation

self-selection

history

Scenario 1

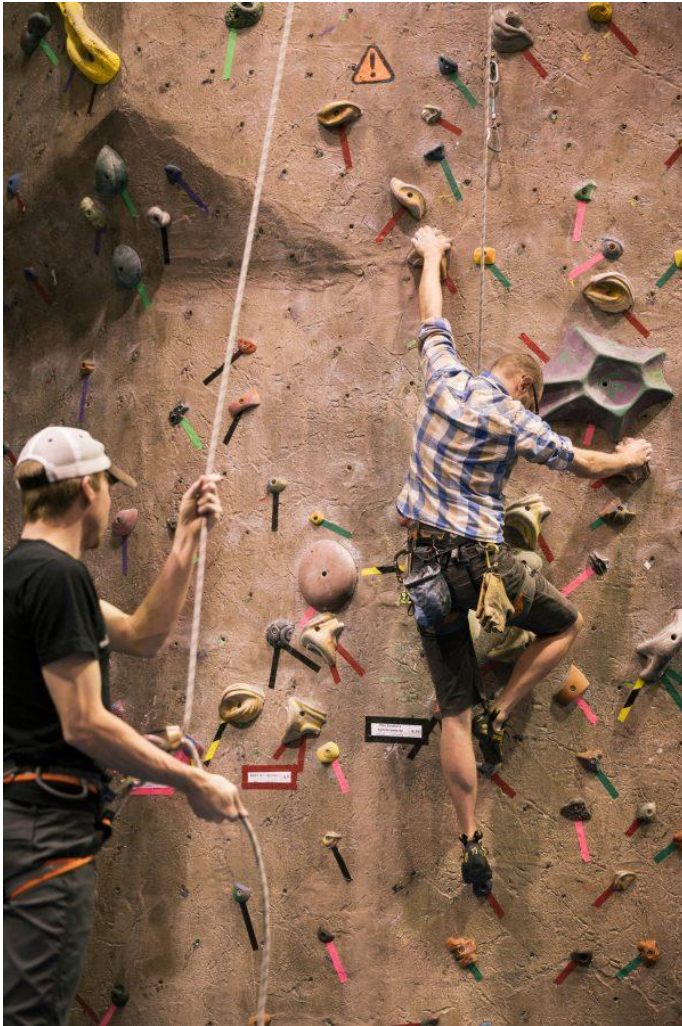


You are in charge of assessing a program through your university's recreation program intended to improve students' knowledge of proper safety techniques in belaying rock climbers. 200 students were randomly assigned to either the treatment group (an hour-long workshop each week for six weeks) or the control group (which received no intervention). At the end of the six weeks, knowledge of essential safety techniques was assessed for all students in both groups. Students who participated in the workshop displayed significantly larger increases in knowledge than students who did not participate.

Do **maturaton** effects seem plausible here?

Why or why not?

Scenario 1



You are in charge of assessing a program through your university's recreation program intended to improve students' knowledge of proper safety techniques in belaying rock climbers. 200 students were randomly assigned to either the treatment group (an hour-long workshop each week for six weeks) or the control group (which received no intervention). At the end of the six weeks, knowledge of essential safety techniques was assessed for all students in both groups. Students who participated in the workshop displayed significantly larger increases in knowledge than students who did not participate.

Do maturation effects seem plausible here?

Do **history** effects seem plausible here?

Why or why not?

Scenario 2



Your university decides to add a new program targeting knowledge about campus alcohol policies to the week-long freshman orientation. All freshman students will complete the new program. All students complete a pre-test prior to the program and a post-test after the program ends. You find that scores on a measure of knowledge about campus alcohol policies rise considerably from pre-test to post-test.

What is the design?

single time point	or	multiple time points
no comparison group	or	comparison group
no random assignment	or	random assignment

Scenario 2



Your university decides to add a new program targeting knowledge about campus alcohol policies to the week-long freshman orientation. All freshman students will complete the new program. All students complete a pre-test prior to the program and a post-test after the program ends. You find that scores on a measure of knowledge about campus alcohol policies rise considerably from pre-test to post-test.

What is the design?

single time point

no comparison group

no random assignment

or

multiple time points

or

comparison group

or

random assignment

Scenario 2



Your university decides to add a new program targeting knowledge about campus alcohol policies to the week-long freshman orientation. All freshman students will complete the new program. All students complete a pre-test prior to the program and a post-test after the program ends. You find that scores on a measure of knowledge about campus alcohol policies rise considerably from pre-test to post-test.

Which of these problems accompany multiple time point designs without randomization/comparison groups?

maturation

self-selection

history

Scenario 2



Your university decides to add a new program targeting knowledge about campus alcohol policies to the week-long freshman orientation. All freshman students will complete the new program. All students complete a pre-test prior to the program and a post-test after the program ends. You find that scores on a measure of knowledge about campus alcohol policies rise considerably from pre-test to post-test.

Which of these problems accompany multiple time point designs without randomization/comparison groups?

maturational

self-selection

history

Scenario 2



Your university decides to add a new program targeting knowledge about campus alcohol policies to the week-long freshman orientation. All freshman students will complete the new program. All students complete a pre-test prior to the program and a post-test after the program ends. You find that scores on a measure of knowledge about campus alcohol policies rise considerably from pre-test to post-test.

Do **maturati**on effects seem plausible here?

Why or why not?

Scenario 2



Your university decides to add a new program targeting knowledge about campus alcohol policies to the week-long freshman orientation. All freshman students will complete the new program. All students complete a pre-test prior to the program and a post-test after the program ends. You find that scores on a measure of knowledge about campus alcohol policies rise considerably from pre-test to post-test.

Do maturation effects seem plausible here?

Do **self-selection** effects seem plausible here?

Why or why not?

Scenario 2



Your university decides to add a new program targeting knowledge about campus alcohol policies to the week-long freshman orientation. All freshman students will complete the new program. All students complete a pre-test prior to the program and a post-test after the program ends. You find that scores on a measure of knowledge about campus alcohol policies rise considerably from pre-test to post-test.

Do maturation effects seem plausible here?

Do self selection effects seem plausible here?

Do **history** effects seem plausible here?

Why or why not?

Let's Get Real Activity

Morning Handout #3

Morning Handout #3: Let's Get Real Activity

An assessment coordinator sent the following report to program faculty, requesting feedback on the report before disseminating it to a wider audience. Two faculty members could not come to consensus about the conclusions drawn from the report. One faculty member believes it is great and should be submitted with the annual assessment report as is. The other faculty member is extremely critical and believes the assessment report has no merit. On the next page is a brief excerpt from the report, along with a summary of the comments from the two faculty members. Your job is to break the tie.

Let's Get Real Activity: Scenario

An assessment coordinator sent the following report to program faculty, requesting feedback on the report before disseminating it to a wider audience. Two faculty members could not come to consensus about the conclusions drawn from the report.

One faculty member believes it is great and should be submitted with the annual assessment report as is. The other faculty member is extremely critical and believes the assessment report has no merit. On the next page is a brief excerpt from the report, along with a summary of the comments from the two faculty members. Your job is to break the tie.

Let's Get Real Activity: Scenario

Think through the excerpt, reviewer comments, and questions individually (~5 minutes)

Share your thoughts with others at your university (~10 minutes)

- Choose a representative from your university

Each university will share their thoughts about their assigned question (~5 minutes)



What research design was employed?



What research design was employed?

Single time point with random assignment and a comparison group





OLD DOMINION UNIVERSITY

What inferences were made?

What inferences were made?

Neuroscience Bootcamp is effective for increasing knowledge

Neuroscience Bootcamp causes a 16 point growth



Are there any *plausible* threats to the validity of those inferences?

What makes those threats plausible?

Are there any *plausible* threats to the validity of those inferences?

Threat to the validity of the growth inference

What makes those threats plausible?

They didn't measure the same students pre-post, so they can't make inferences about the amount of growth



How would you respond to each of the reviewers?
What changes would you make to the assessment methods or the assessment report?

How would you respond to each of the reviewers?

Neither are completely correct—we can confidently make inferences about program effectiveness but not about the specific amount of growth

What changes would you make to the assessment methods or the assessment report?

To the assessment methods—include pre- and post-test for the same students if we want to make inferences about growth

To the assessment report—remove the inference about growth for now

A solid teal square is positioned on the left side of the slide.A dashed teal rectangular border frames the central text area.

Practical Challenges

Things Just Got Even More Real

A solid grey square is positioned on the right side of the slide.

The “Real” → “Ideal” Continuum



Practical Challenges

- Small sample sizes (small N)
- Attrition (from school, our program, our measures)
- Implementation
- Rubric data
- Course-embedded data

Breakout Activity

Morning Handout #4

Morning Handout #4: Breakout Activity on Practical Challenges

<i>Practical Challenge</i>	<i>Location</i>
Small sample size (small <i>N</i>)	UVA
Attrition (from school, program, measure)	JMU
Implementation	VCU/ Cisco

At your location:

- How does this challenge complicate assessment and/or the inferences that can be made about the program?
- What does this challenge look like at your institution/within your program or department?
- When faced with this challenge, what do you do? How do you draw meaningful claims about your programs?
- Offer suggestions for the rest of us.
- We will report back after 20 minutes
- If you want to send some things to add to the presentation, email horstsj@jmu.edu. We will add your information to our ending slides.

Practical Challenge	Location
Small sample size (small N)	UVA
Attrition (from school, program, measure)	JMU
Implementation	VCU/ Cisco
Rubric Data	GMU
Course-embedded data	VT/ODU

If you finish your assigned challenge, pick a challenge of your choosing to work with.

Small Sample Sizes



Example: Some majors may only have few graduates per year. The challenge is how to present credible evidence of their program's impact on student learning. What do you do in this situation?

Let's Talk: Small Sample Sizes

Example: Some majors may only have few graduates per year. The challenge is how to present credible evidence of their program's impact on student learning.

Ways to approach this challenge

- Qualitative data collection, mixed methods, focus groups, case studies
- Frequency counts
- When able, combine data from cohorts
- Put less emphasis on "power"; report raw metric effect sizes

If you have a benchmark for competency, doesn't matter how small N is

Positives: can do more intensive assessment like performance assessments. More feasible to do some types of assessment with small N

Small Sample Sizes (UVA notes)



Problems

- Limits statistical possibilities
- Limits statistical power
- Confidentiality is difficult to ensure when you have small programs
- Representativeness of the sample is difficult. What is “representative” may change radically from year to year.

Solutions

- Mixed methods: Add qualitative approaches.
- Aggregation across years or across like programs. However, make sure that the method of aggregation is logical/sensible.

Attrition

Example: When assessing one of our small programs ($N = 30$), despite several on-site requests for post-test data, students did not respond. Our sample size went from $N = 30$ to $N = 7$. How do we draw accurate conclusions when we're missing data from over half of our original sample?

Let's Talk: Attrition

Example: When assessing one of our small programs ($N = 30$), despite several on-site requests for post-test data, students did not respond. Our sample size went from $N = 30$ to $N = 7$. How do we draw accurate conclusions about our program's impact, when we're missing data from over half of our original sample?

Ways to approach this challenge

- Accurately report the attrition as a limitation (be up front about it)
- For next time, think through how to guard against it
- Compare the data at the first time-point. Are the two groups different, in any way?

Is the attrition due to dropping out of treatment (i.e., treatment attrition) vs. not responding to our measures (i.e., measurement attrition)?



Attrition (JMU notes)

Problems

- Figuring out when you'll be able to access a sample of students
- If SLOs change across months/semesters/years, you might have attrition problems as assessments (or items on assessments) become irrelevant.

Solutions

- Collect data when you have a "captive audience."
- Keep your learning outcomes stable so that you can compare across years.
- Determine whether or not you actually do need pre-/post-test data
- What's the purpose of the assessment? Say we have an engineering program where a lot of students leave because the program is difficult. Are you trying to gather information about the students who start the program? Or are you actually only interested in the students who finish the program?
- Individualizing programs based on student needs and then assessing that flexible program. This might prevent attrition from happening in the first place because students may identify more closely with the program or benefit more from its content.

Implementation



Example: We want to show evidence of program effectiveness, for example, in general education. Yet, there are multiple courses all taught to the same student learning outcomes. We respect faculty autonomy and don't have a standardized curriculum. We are making claims about program effectiveness, yet we don't always know what's going on in the classroom.

Let's Talk: Implementation

Example: We want to show evidence of program effectiveness, for example, in general education. Yet, there are multiple courses all taught to the same student learning outcomes. We respect faculty autonomy and don't have a standardized curriculum. We are making claims about program effectiveness, yet we don't always know what's going on in the classroom.

Implementation Fidelity

- Systematic approach to seeing whether what's taught in the classroom is aligned to the student learning outcomes.
- Is what we say we're teaching, really what we're teaching?

Implementation (VCU notes)

Problems

- Faculty participation: What is our relationship, as assessment professionals, to faculty members?
 - If we don't acknowledge assessment activity in faculty promotion and tenure, it's unsurprising that faculty might resent involvement in assessment
 - How do we work with part time/adjunct faculty?
- Faculty already have a lot on their plates.
- Assessment as a "stick" instead of a "carrot"

Solutions

- It's important to try our best not to be vague, trying to standardize/control as much as possible ***without infringing on faculty autonomy.***
- Professional development activities about assignment design could be helpful in producing transparency
- "Autonomy" and "standardization" are often the binary options. We should think about how to give faculty opportunities to construct a different understanding of how to participate in a program that has shared expectations for student learning, and how to embed in a course elements that don't restrict autonomy but *do* support a shared goal.

Let's Talk: Implementation

Example: We want to show evidence of program effectiveness, for example, in general education. Yet, there are multiple courses all taught to the same student learning outcomes. We respect faculty autonomy and don't have a standardized curriculum. We are making claims about program effectiveness, yet we don't always know what's going on in the classroom.

Ways to approach this challenge - Ideal world:

- Recruit someone external to observe the classroom/program
- Ask faculty/staff to reflect on the alignment of their programming/curriculum to the SLOs

Ways to approach this challenge - Real world:

- There are good reasons faculty don't want too much standardization (e.g., common assignments). We want to respect faculty autonomy.
- Takes secure faculty who *get* assessment. Develop relationships with faculty and have conversations. Be on the same page. This takes time.

Rubric Data



Example: Many of us are in the situation where we are using rubric data. What is the best way to report rubric data?

Let's Talk: Rubric Data

Example: Many of us are in the situation where we are using rubric data. What is the best way to report rubric data?

Ways to approach this challenge (or opportunity!)

- A lot to consider. How many score levels do you have?
- If it's ordinal data. How do we deal with the data? Averages not appropriate.
- Dimensionality: Is it appropriate to create a total score when there are different dimensions or traits?
- How do we deal with rubric data from different assignments?
- How do we evaluate reliability of the ratings?

Rubric Data (GMU notes)



Problems

- Reviewers of artifacts might have disparate grading practices
- Measuring the unmeasurable: how do we measure things like creativity in a rubric?
- Reporting: what do the numbers we produce actually mean?

Solutions

- Norming: It's an arduous process and we can't always do it, but we should do it whenever possible.
- Create targeted rubrics: It's important to make sure that the rubric actually fits the artifact you're trying to assess, even though common rubrics are often desired
- Alignment: Do the rubrics match the artifacts? Do both the rubric and the artifacts match the instruction?

Example

Proportion at each level from rubric on previous slide. Helps with usefulness of results.

Purpose

Generally refers to conveying a message appropriate to its audience. Features may include a thesis or central idea, topic selection, relevance, clarity, and focus.

Where are they?

- *Developing* (27%) = Occasionally appropriate for the audience or intended audience somewhat clear. Central idea, thesis, or goal emerges but may lack focus or consistency.
- *Competent* (54%) = Mostly appropriate for a defined audience. Exhibits a generally clear and consistent central idea, thesis, or goal.
- *Advanced* (19%) = Clearly appropriate for a well-defined audience. Consistently exhibits a focused central idea, thesis or goal.

Why might they be here?

How can we help ALL students get to the *Competent* level or better?

Assignment Strategies

This is the competency question we mentioned at the beginning. We cannot draw strong causal claims, but we can demonstrate our students' competence.

Course-Embedded Data



Example: Many of us are in the situation where we are using course-embedded data. How do we plan, collect and manage the course-embedded data in order to draw causal claims about the effectiveness of our programs?

Let's Talk: Course-Embedded Data

Example: Many of us are in the situation where using course-embedded data. How do we plan, collect and manage the course-embedded data in order to draw causal claims about the effectiveness of our programs?

Ways to approach this challenge (or opportunity!)

- A lot to consider. How many courses? How different can the courses be and still be comparable?
- There are specific designs that can help make causal claims. These are borrowed from the educational research (Scholarship of Teaching and Learning; SOTL) literature. Each have their associated plausible threats to validity of the inferences about causality.
 - Example: Compare similar course with and without the intervention (pre- and post-test)

Course Embedded (ODU notes)

Problems (or opportunities)

- Ensuring that the embedded assignments are well-aligned to learning outcomes. If they aren't representative of the learning outcomes, we get results that can't be generalized and can't lead to improvement.
- We can't force anyone to collaborate on common assignments/items. If we don't get full compliance, then we can't generalize.
- Calibration: if you have courses with multiple sections and faculty aren't calibrated on how they assess the assignments (or how they present the assignments to students), our results won't be generalizable.
- If the course content isn't standardized across modalities (e.g. in-class or online classes), we might have significant differences across sections due to the modality.
- If courses are undergoing changes during data collection, we might not be measuring the same thing across different classes.

Solutions

- Identify which outcomes we are assessing with our assignments
- Distinguish between assessment that happens within a course (is the student doing well in the class?) and course-embedded assessment (collecting data about student achievement for the purposes of program-level assessment). These levels might not always agree.

Course Embedded (VT notes)



Problems (or opportunities)

- Often a problem in general education assessment.
- Academic freedom vs. standardization: how do we assess without infringing on academic freedom?

Solutions

- Standardize rubrics (when appropriate)
- Standardizing small elements of courses, but allow flexibility in remaining areas (so we aren't removing faculty autonomy across the entire course).

See next several slides, submitted during conference.
Apologies for missing these during the presentation (jh).

Course embedded Problems Question 1

- Assessment tools may not be aligned to student learning outcomes of the course or program.
- Assessment instruments might be poorly designed, invalid and inconsistent.
- Institutions that are new to assessment may not know how to effectively use the broad spectrum of data they receive.

Submitted by:



Question 2

- There is wide variation among course types. Not every course can be assessed the same way.
 - Large institutions can have complex problems.
 - Quality of data received.
- General Education student learning outcomes can be difficult to measure.
- Academic Freedom is a challenge to effective assessment practices.

Submitted by:



What do you do about it

- Assessment professionals should provide workshops to help educate instructors and faculty on how to create effective assessment tools.

Submitted by:



Can we Back Up that Claim?

Making Important Data Collection Design
Decisions

Andrea **Pope**, Caroline **Prendergast**, Morgan **Crewe**, & Jeanne **Horst** | James Madison University

Supporting High-Quality Teaching: A Day of Dialogue

- **Friday, June 7**
- **9:30 a.m.-3:00 p.m.**
- **Reynolds Community College (Richmond)**
- **Registration open until May 17**

<http://www.schev.edu/index/agency-info/additionalactivities/high-quality-teaching>

Or visit www.schev.edu and search for “high quality teaching”

Save the Date *2019 Annual Conference*

Theme: Inspiring Leadership for Innovations in Assessment

When: November 13-15, 2019

Where: Delta Hotel Marriott in Downtown Richmond

Join us for an opportunity to learn and network
with professionals in the field of assessment and beyond!



A Big Thanks to 4-VA!

This drive-in would not have happened without the support of 4-VA and Cisco. We have them to thank for the food, technology support, and Community College travel support. Many thanks to personnel for support at the six 4-VA locations.



Kelsey Tate

And a special shout-out to Kelsey Tate, who was the mastermind behind all of the technology that made today's drive-in possible. This would not have happened without you!

Morning Handout #1: And the Award Goes To...

The mission of the Center for Curricular and Co-Curricular Excellence (CCCE) is to recognize and support exemplary academic and student affairs programs that are contributing substantially to students' learning and development.

For the last five years, the CCCE has awarded one exemplary program a \$10,000 grant to support program expansion. As director of the Center, you are responsible for reviewing this year's nominations for the Program Excellence Award. The award criteria are as follows:

- Programs must target student learning/development outcomes.
- Programs must demonstrate an impact on student learning/development.

Your Task: For this first round of reviews, determine whether each of the programs below meets the award criteria. Only programs that meet both criteria will be considered for the Program Excellence Award.

Community Service Learning (CSL) – Alternative Spring Break Program (ASB)

Learning Outcome: Civic-Mindedness

Program Description: Each year, CSL offers a variety of week-long service breaks to locations in the United States and abroad. These alternative breaks provide opportunities for participants to engage directly with community members through hands-on projects and activities. There are three important elements of ASB:

- **Education:** Before the alternative break experience, participants attend three sessions to learn about the communities, organization/s, and projects with which they are working.
- **Reflection:** During the alternative break experience, structured time is set aside for participants to reflect upon the experience.
- **Re-Orientation:** Upon return from the alternative break experience, individuals are encouraged to transfer lessons learned by engaging in continued education, service, and/or advocacy.

Summary of Assessment Results: After completing ASB, students were required to complete a measure of civic-mindedness. Civic-mindedness is defined as displaying a concern for the public good or humanity as a whole. According to the survey, 87% of ASB participants demonstrated high levels of civic-mindedness, 9% of ASB participants demonstrated moderate levels of civic-mindedness, and only 4% of ASB participants demonstrated low levels of civic-mindedness.

Meets criteria for Program Excellence Award? ☐ Yes ☐ No ☐ Unclear

International Business, B. A.

Learning Outcome: Written Communication

Program Description: All students are required to take a 3-credit writing seminar the semester after enrolling in the major. Additionally, all students must complete two writing-intensive major courses. As part of these courses, students complete a semester-long writing project. Multiple drafts of the project are submitted throughout the semester and students receive detailed, formative feedback between each draft.

Summary of Assessment Results: Assessment results from the 2017-2018 academic year indicate that students entered the major with poor writing skills (1.5 out of 5 on writing rubric). After two years of coursework, including the writing seminar and writing-intensive courses, students demonstrated great improvement (scored 3.5 out of 5 on writing rubric).

Meets criteria for Program Excellence Award? ☐ Yes ☐ No ☐ Unclear

Office of Student Success – Student Academic Success Program (SAS)

Learning Outcome: Academic Self-Efficacy

Program Description: SAS is an voluntary eight-week course designed to help students on academic probation return to good academic standing (i.e., achieve a cumulative grade point average of 2.0 or higher). This is accomplished by targeting knowledge, attitudes, and skills empirically related to academic success, such as knowledge of academic support resources, organizational skills, and academic self-efficacy (belief in one's ability to achieve in an academic setting).

Summary of Assessment Results: At the end of the fall semester, a measure of academic self-efficacy was sent to all students on academic probation. Results were presented for two groups: students who chose to complete the program (SAS participants), and students who did not. Notably, SAS participants scored two standard deviations higher than their peers who did not complete the course. In other words, SAS participants had much greater confidence in their academic abilities than students who did not complete the course.

Meets criteria for Program Excellence Award? ☐ Yes ☐ No ☐ Unclear

Morning Handout #2: Threats to Internal Validity

Maturation Effect: The observed effect is due to normal developmental processes or changes over time, not the program.

- Program A (implemented during the first semester of college) claims to have increased students' sense of independence. However, studies show students naturally gain more independence during their first semester of college even without an intervention.

History Effect: The observed effect is not due to the program, but to some other unaccounted for event.

- Program B claims to have reduced instances of sexual assault on campus. However, sexual assault prevention is a university-wide initiative and upon further investigation, the facilitators of Program B realize their participants also received programming related to sexual assault prevention in their residence halls. Could the reduction in instances of sexual assault be due to this residence life program instead?

Self-Selection Bias: The observed difference between two groups at posttest is not due to the program, but to preexisting differences between the groups.

- Facilitators of Program C compare students who participated in their service learning program to students who did not and are pleased to find that their students are higher in civic engagement—clear evidence that the program works! Upon further investigation, however, they discover that students high in civic engagement were more likely to participate in their program in the first place. Thus, the difference between the groups was due to self-selection into the program, not the program's effectiveness.

Attrition: The observed effect may be biased due to a substantial amount of missing data (i.e., students failing to complete the program or take the posttest).

- Organization D finds that students' sense of belonging to their organization increased drastically from pretest to posttest—a major success! Upon further investigation, however, it becomes clear that students who felt lower sense of belonging dropped out of the organization and, thus, did not take the posttest. As such, the posttest results were artificially inflated.

Instrumentation Effect: The observed effect is due to changes in the instrument (or interpretation of scores), not the program.

- Program E recruits several raters to review ethical reasoning essays before and after a month-long ethical reasoning program. Shockingly, it seems students performed *worse* after the intervention. Upon further investigation, however, it becomes clear that the problem was with the raters. They became more critical over time, thus evaluating the posttest essays more harshly than the pretest essays.

Response Processes: Results cannot be trusted to reflect students' true ability because they are impacted by things like socially desirable responding and low motivation.

- After completing a 6-hour alcohol prevention workshop, students are fatigued and ready to leave. Unsurprisingly, when asked to complete a 100-item posttest (the only thing separating them from freedom) they speed through the test, responding randomly to the questions. Subsequent posttest results show students gained nothing from the workshop. Should these results be trusted?

For more information, consult: Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York, NY: Houghton Mifflin Company.

	Threats to Validity					
Data Collection Designs	Maturation Effect	History Effect	Selection Bias	Attrition	Instrumentation Effect	Response Processes
Posttest Only Design	X	X	N/A	N/A	N/A	X
Pretest-Posttest Design	X	X	N/A	X	X	X
Comparison Group Design (No Pretest)	?	?	X	X	N/A	?
Pretest-Posttest w/ Comparison Group Design	✓/?	✓/?	✓/?	X	✓	?
Pretest-Posttest w/ Comparison Group & Random Assignment Design	✓	✓	✓	X	✓	?

X = The design is highly susceptible to this threat.

? = This threat is a *possible* source of concern for the design. The design *may* provide partial protection against this threat.

✓ = The design provides strong protection against this threat.

N/A = The threat is not applicable for this design.

Morning Handout #3: Let's Get Real Activity

An assessment coordinator sent the following report to program faculty requesting feedback on the report before disseminating it to a wider audience. Two faculty members could not come to consensus about the conclusions drawn from the report. One faculty member believes it is great and should be submitted with the annual assessment report as is. The other faculty member is extremely critical and believes the assessment report has no merit. Below is a brief excerpt from the report, along with a summary of the comments from the two faculty members. Your job is to break the tie.

You work through the following questions to prepare your review:

(Work through each of these questions, focusing on the one your site will answer aloud.)

- What research design was employed? (GMU will answer)
- What inferences were made? (ODU will answer)
- Are there any *plausible* threats to the validity of those inferences? What makes those threats plausible? (VCU and VT will answer)
- How would you respond to each of the reviewers? What changes would you make to the assessment methods or the assessment report? (UVA and JMU will answer)

Assessment Methods

The purpose of this assessment report was to examine the impact of a week-long bootcamp on knowledge of techniques used in Cellular Neuroscience. Specifically, the bootcamp addressed the SLO that “after completing Neuroscience Bootcamp, students will demonstrate a 10 point growth in knowledge of techniques used in Cellular Neuroscience.” Neuroscience Bootcamp is offered during the first two weeks of August to incoming first-year students who intend to pursue a degree in neuroscience. Since this bootcamp is offered prior to the start of the semester, program faculty expect this bootcamp is students’ first exposure to Cellular Neuroscience techniques. This year, 104 students participated in Neuroscience bootcamp. Half of the students were randomly assigned to attend the bootcamp during the first week of August, and the other half are randomly assigned to attend the bootcamp during the second week of August. Due to time and resource constraints, students were only assessed at one time point. Specifically, both groups of students were assessed at the end of the first week of August, which was *after* the first group completed the bootcamp and *before* the second group started the bootcamp. See timeline portrayed in Figure 1.

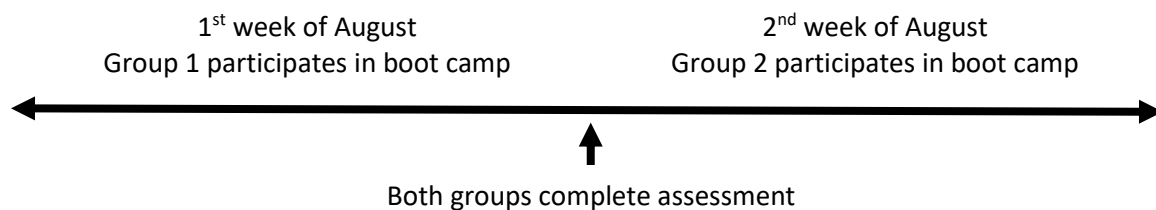


Figure 1. Assessment design

Knowledge of techniques used in Cellular Neuroscience was assessed using an online assessment. The assessment consisted of several demographic questions including age, race, and gender, and a series of questions about Cellular Neuroscience Methods. Knowledge of these methods was measured using twenty multiple choice items. These items were combined to produce a total score that ranged from 1-20, with lower scores indicating less knowledge of Cellular Neuroscience methods and higher scores indicating more knowledge of Cellular Neuroscience methods.

Results

All students who participated in the program completed the online assessment. The assessment results are presented below.

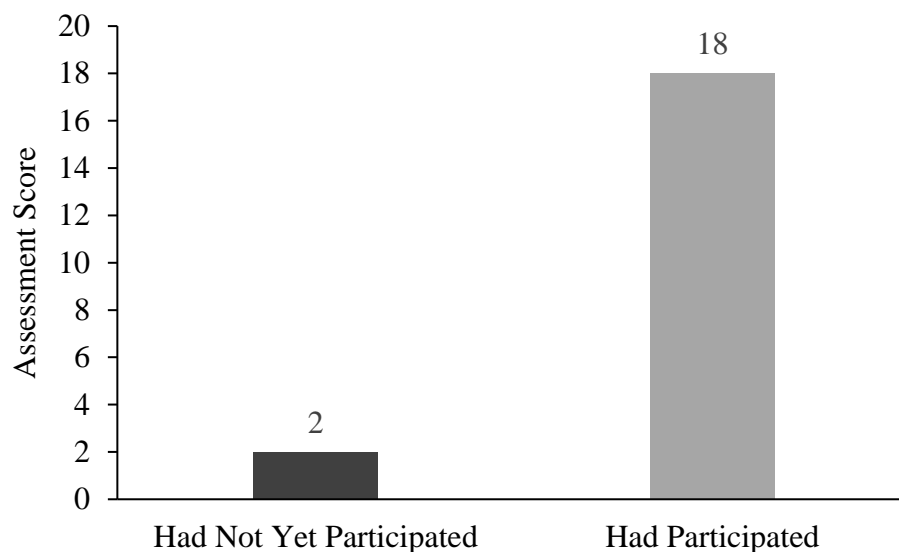


Figure 2. Average assessment scores for individuals who had not yet participated in Neuroscience Bootcamp versus students who had participated in Neuroscience Bootcamp

Conclusions

Students who participated in Neuroscience Bootcamp scored higher on knowledge of Cellular Neuroscience methods than students who had not yet participated in Neuroscience Bootcamp. These results show that Neuroscience Bootcamp is effective for increasing knowledge of Cellular Neuroscience methods in first year students, and results in a 16 point growth in knowledge of Cellular Neuroscience methods measured by this assessment. These results provide compelling evidence that universities that wish to produce large, measurable growth in their students' knowledge of Cellular Neuroscience methods should implement Neuroscience Bootcamp at their school.

Faculty Member 1 Comments: The report portrays a well-designed assessment plan that clearly shows their program is effective for not only increasing knowledge of Cellular Neuroscience methods in first year students, but producing a 16 point growth in students! Well done! We need to make sure that all incoming students attend the neuroscience bootcamp. The design of this assessment is infallible due to the use of random assignment with a comparison group. This assessment report is ready to be submitted for our annual report.

Faculty Member 2 Comments: The assessment has major design flaws and issues that call into question the integrity of the conclusions. Without a pre and post-test design you cannot make any claims about the effectiveness of your program or the ability of your program to produce growth in students' knowledge. This report should not be submitted as is, as the results have no merit.

Morning Handout #4: Breakout Activity on Practical Challenges

<i>Practical Challenge</i>	<i>Location</i>
Small sample size (small <i>N</i>)	UVA
Attrition (from school, program, measure)	JMU
Implementation	VCU/ Cisco
Rubric Data	GMU
Course-embedded data	VT/ODU

Discuss the following questions:

- How does this challenge complicate assessment and/or the inferences that can be made about the program?
- What does this challenge look like at your institution/within your program or department? Some examples are on the reverse side of this sheet. Feel free to use them or to come up with your own.
- When faced with this challenge, what do you do? How do you draw meaningful claims about your programs?

Report Out: Offer suggestions (or lessons learned) to address these practical challenges.

Two Formats: Verbally or email to horstsj@jmu.edu/. We will add your information to our ending slides.

Time: 20 minutes

If you finish your assigned challenge, pick a challenge of your choosing to work with.

Examples of Practical Challenges

Small Sample Sizes: Some majors may only have few graduates per year. The challenge is how to present credible evidence of their program's impact on student learning. What do you do in this situation?

Attrition: When assessing one of our small programs ($N = 30$), despite several on-site requests for post-test data, students did not respond. Our sample size went from $N = 30$ to $N = 7$. How do we draw accurate conclusions when we're missing data from over half of our original sample?

Implementation: We want to show evidence of program effectiveness, for example, in general education. Yet, there are multiple courses all taught to the same student learning outcomes. We respect faculty autonomy and don't have a standardized curriculum. We are making claims about program effectiveness, yet we don't always know what's going on in the classroom.

Rubric data: Many of us are in the situation where we are using rubric data. What is the best way to report rubric data?

Course-embedded data: Many of us are in the situation where we are using course-embedded data. How do we plan, collect and manage the course-embedded data in order to draw causal claims about the effectiveness of our programs?

Facilitator Guide

Morning Handout #4: Breakout Activity on Practical Challenges

Purpose: This short facilitator guide explains Morning Handout #4: *Breakout Activity on Practical Challenges*. The purpose of the activity is to address the learning outcome “Recognize common assessment data collection challenges.”

Length: 75 minutes

Outline

- 10 minutes: Jeanne will introduce the activity and will assign practical challenges (Jeanne will give these instructions to the group via *telepresence*).
- 20 minutes: Group work *at individual location (site host leads)*
 - 1 minute: Assign (or ask for volunteer) a spokesperson, timekeeper, and a scribe or moderator, if desired, for your location.
 - 2 minutes: Explain to participants that this is a pair-share activity. Participants will work in small groups and then share with others at your location. At the end of this activity, when called upon (*via telepresence*), your location will be asked to report out. We will ask your location’s spokesperson to report out to the larger group (*via telepresence*). If you wish, you may add the information to a PowerPoint slide or Word doc and send to Jeanne at horstsj@jmu.edu to show when it’s your turn to share. You may use the examples on the back of this sheet, or create your own example that is more relevant to your location. Respond to the questions and provide suggestions for how to work with your assigned challenge.
 - 7 minutes: Small groups of 2-3 discuss their thoughts to the following questions, related your location’s assigned topic.
 - How does this challenge complicate assessment and/or the inferences that can be made about the program?
 - What does this challenge look like at your institution/within your program or department? Some examples are on the reverse side of this sheet. Feel free to use them or to come up with your own.
 - When faced with this challenge, what do you do? How do you draw meaningful claims about your programs?
 - 10 minutes: Small groups share their thoughts with the larger group at your location. Someone moderates or serves as scribe. Prepare a 5-minute message

that you want to share via telepresence to the entire group, when called upon. The message can include a brief description of what this challenge might look like on your campuses, and tips or recommendations for working with the challenge. The purpose for this activity is that we can all learn from the ways in which others deal with practical challenges.

- 45 minutes: Each individual location will be called upon to report back on their topic (5-minutes per topic), we will add some thoughts and wrap-up (*telepresence; Jeanne leads*)

Elements of Evidence-Based Storytelling

Gianina Baker, Assistant Director
National Institute for Learning Outcomes Assessment
(NILOA)

Evidence-Based Storytelling

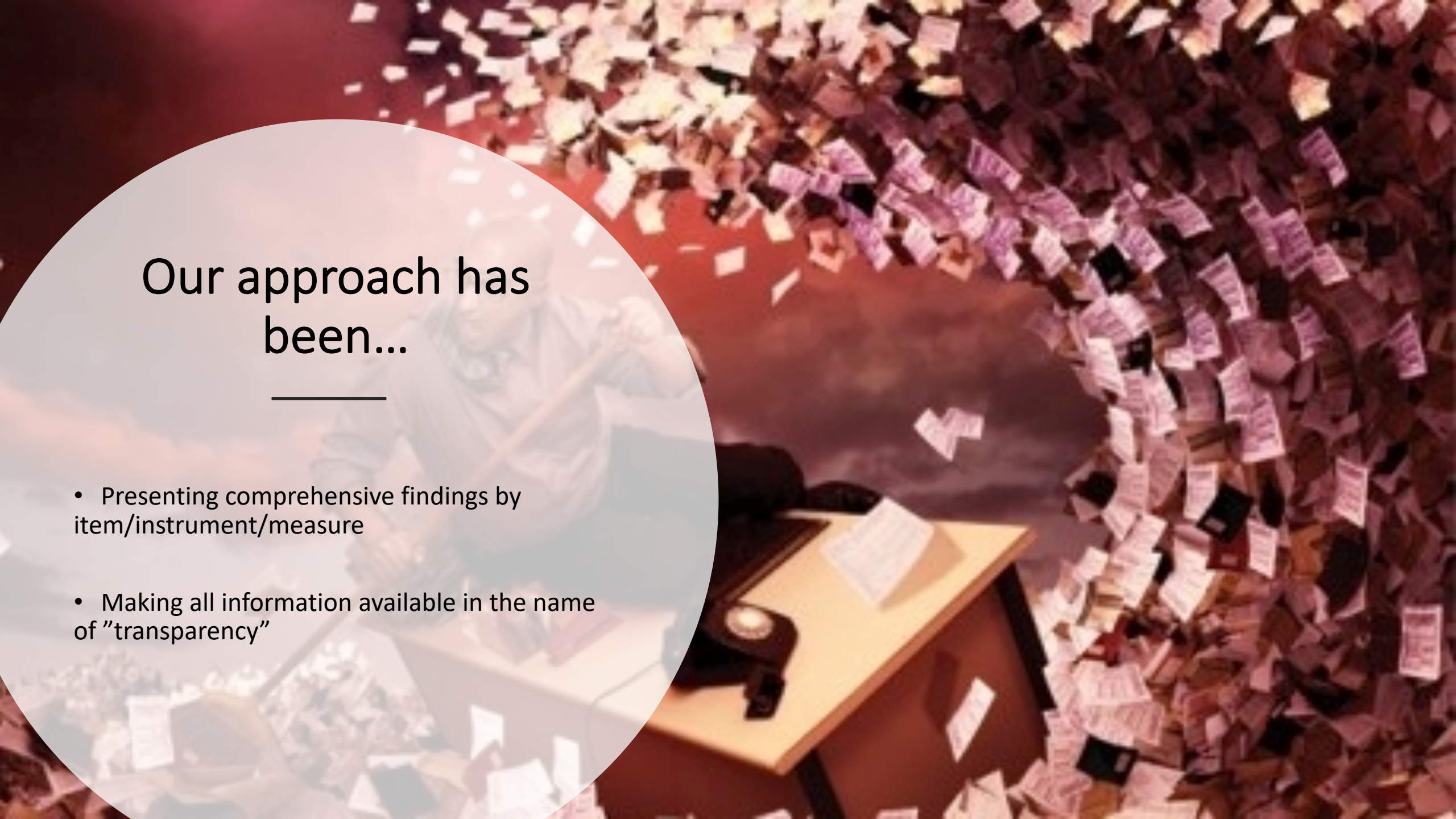


Evidence of student learning is used in **support of claims or arguments** about **improvement and accountability** told through **stories** to **persuade a specific audience**.

Need to tell our story and help students tell theirs.

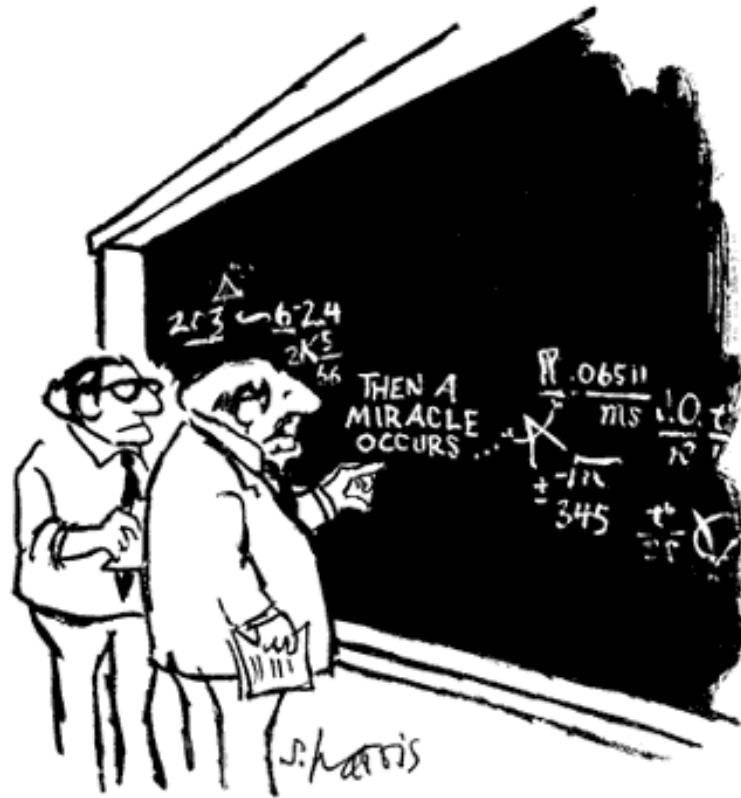
Why the group activity?

- ◆ Opportunity to practice putting together an evidence-based story for various audiences *using a report you didn't have to write* 😊
- ◆ Paying particular attention to audience(s) and purpose
- ◆ Traditional dimensions of storytelling

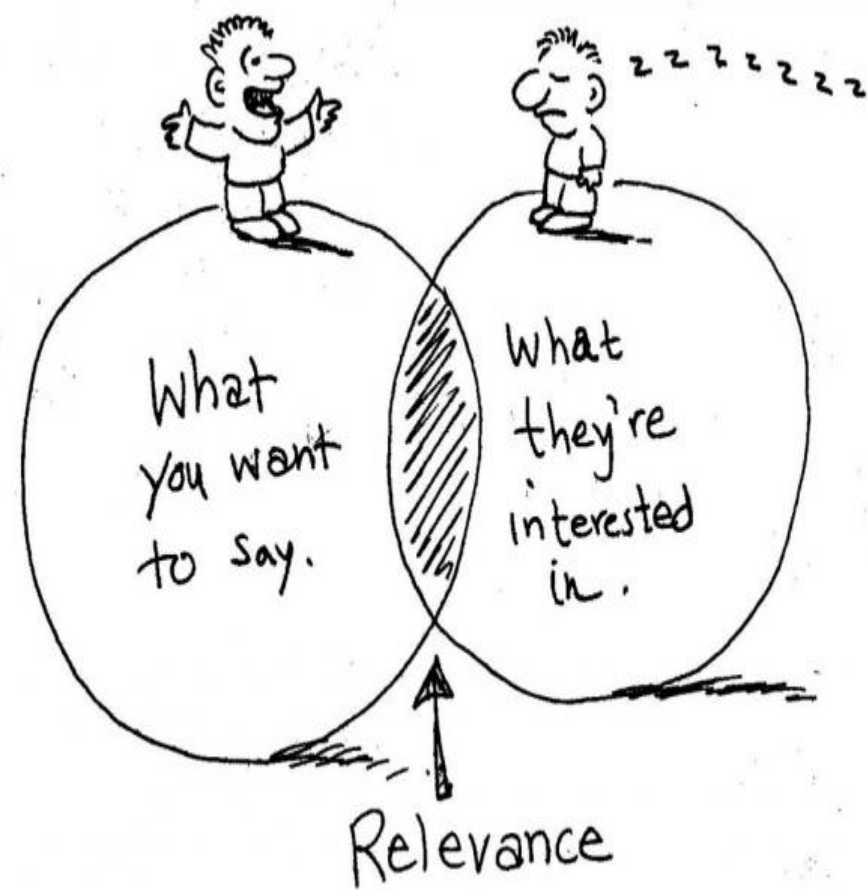


Our approach has been...

- Presenting comprehensive findings by item/instrument/measure
- Making all information available in the name of "transparency"



"I think you should be more explicit here in step two."



Current Approaches

- ◆ Overwhelm with access to data without meaning making
- ◆ Scatter shot bullet lists of processes attempting to guess what people want
- ◆ Archives of reports that document our processes – but provide a history of data collection or changes made
- ◆ Individual student stories of success (internships)



Why Stories?

Shadiow (2013) presents a process to see stories “as something other than **sentimental anecdotes** with thinly veiled lessons” but instead as a mechanism by which we may reflect on our practices and teaching (p. viii).

For instance...

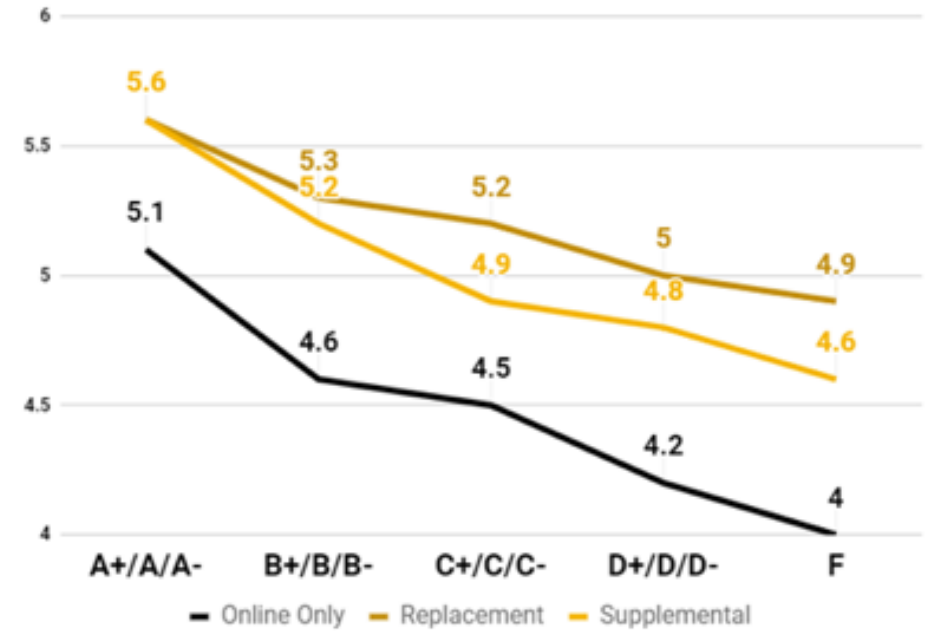
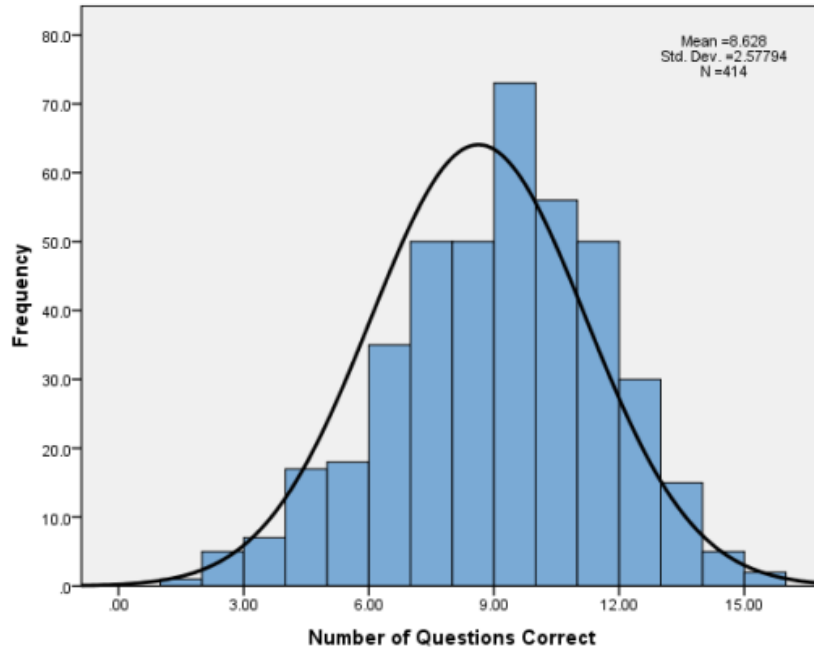
“Educational assessment is at heart an exercise in evidentiary reasoning. From a handful of things that students say, do, or make, we want to draw inferences about what they know, can do, or have accomplished more broadly.” (Mislevy & Riconscente, 2005, p. iv).

A Good Story

A good story is easy to read, introduces a problem, and shares how the problem was solved, highlighting the role of the institution in addressing the problem. **We need context and a story, because evidence gives stories substance, but stories give evidence meaning.** Our stories can be our context, our histories, our missions, our organizational saga, it is how we see the world and why we do what we do.

What do we want to know about students?

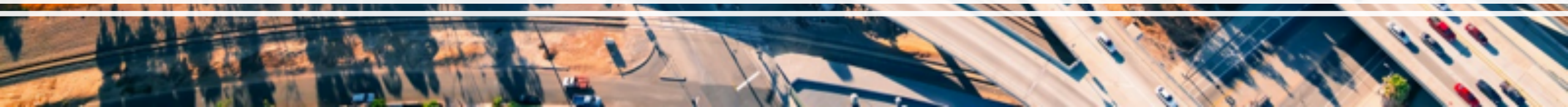
- ◆ What argument do you want to make about your students' learning?
- ◆ What type of evidence would be necessary to make the argument?



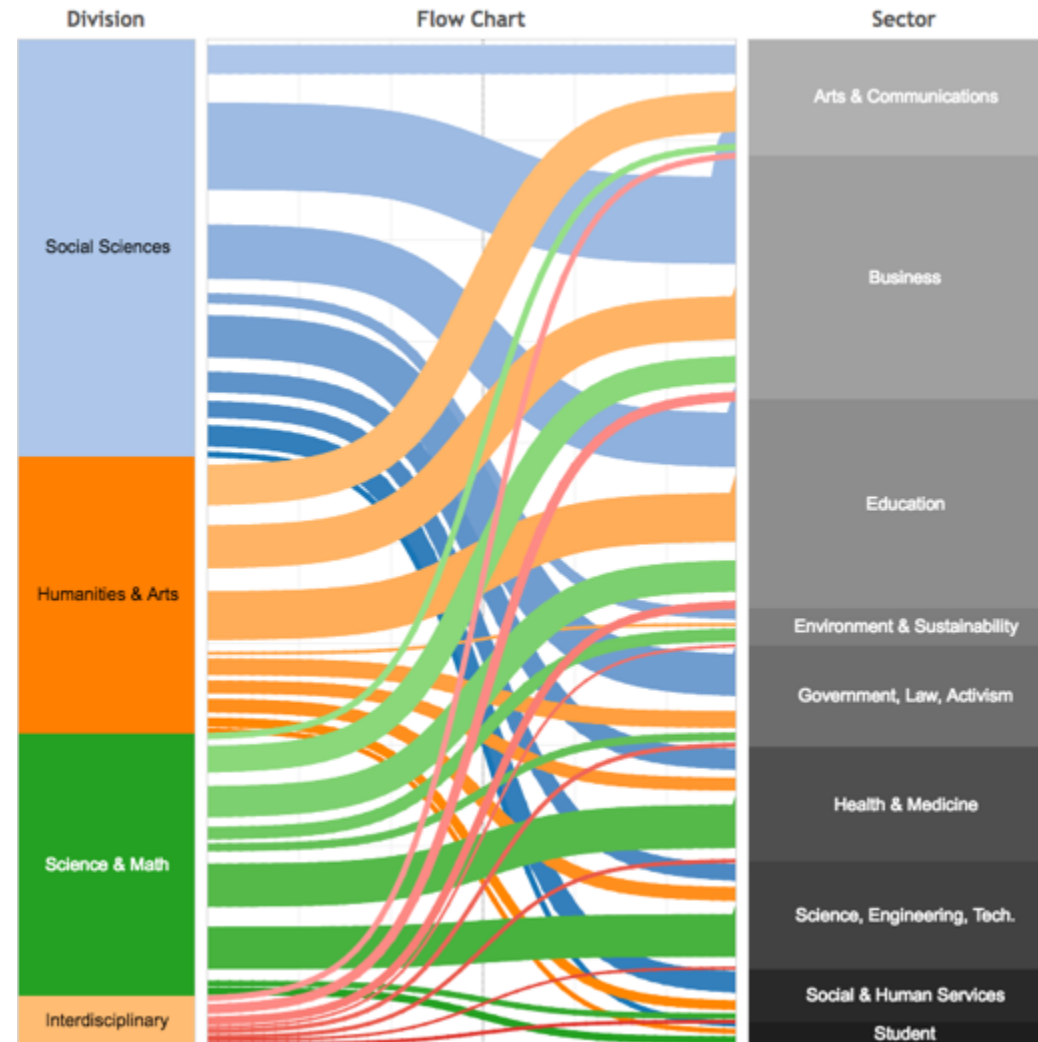
Data Story or Person Story?



Intersections of Data Visualization and Narratives



Visualizations on their own...



Kapi'olani Community College Student Journey

Plan Your College Experience and Adventure



UNIVERSITY OF HAWAII
KAPI'OLANI
COMMUNITY COLLEGE

Toolkits for Reviewing Stories

1. Audience: For whom is this narrative written? What counts as evidence for the different audiences of the report?
2. **What kind of story are you telling?** (i.e., compliance, improvement, loss, struggle, quest, tragedy, fantasy, etc.) What context is needed for readers to understand the story? What is the setting?
3. **Who are the character(s) in your story?** (Is there a protagonist in your story—someone who is driving the action and/or someone with whom your audience is likely to identify? What are the motivations of the characters?)



Toolkits for Reviewing Stories

1. What is the plot? (The plot is the causal sequence of events and includes setting and conflict.)
2. What evidence do you have to assert your claims?
3. Based on the story you crafted, what is the best medium through which to share it? Video, written narrative, shorter visual image pieces, a combination, others?
4. If you are using visuals in your narrative – are they appropriate? Do they support the story you are trying to share or detract from them?
5. How will you make your target audience(s) aware of the story?



Handouts on Storytelling

Group Activity Time!



Storytime!

Reflections

- ◆ Thank you JMU for the letting us use your report for the exercise!
- ◆ Loved incorporating student voice into the presentation
- ◆ Be mindful of direct and indirect measures of assessment and how they can provide strong evidence when used together
- ◆ Helpful to hear other voices/stakeholders in the conversation, not only did they help with interpretation of the results but also in thinking about the various audiences of the presentation
- ◆ Highlighting the importance of substantiating the need by connecting through institutional priorities
- ◆ Consider primary and possibly secondary audiences (i.e., student orientation presentation where students are primary and parents are secondary)

Some Common Pitfalls of Reports

- ◆ Use of unnecessary visualizations
- ◆ Overuse of bullets
- ◆ Too much information on data collection and analysis and not enough making sense of the data and implications for the reader
- ◆ Data point without context

Thoughts on *how* to share...

- ◆ We need to provide reports that begin the meaning making process for readers - we need to synthesize information and present findings
- ◆ We need to connect the data to real people – making the information lived and tied to a story or persona, allows readers to connect with the information leading to more likely use of the data to inform practice
- ◆ Mix up the presentation of results: While questions might be in a particular order on a survey – don't have to stick to that for findings – can group things together around shared points that make sense for the reader and enhance the argument
- ◆ Note on the use of bullets: Bullet points aren't a narrative and if everything is a bullet, no point to have them! They also take up additional space you could use for text

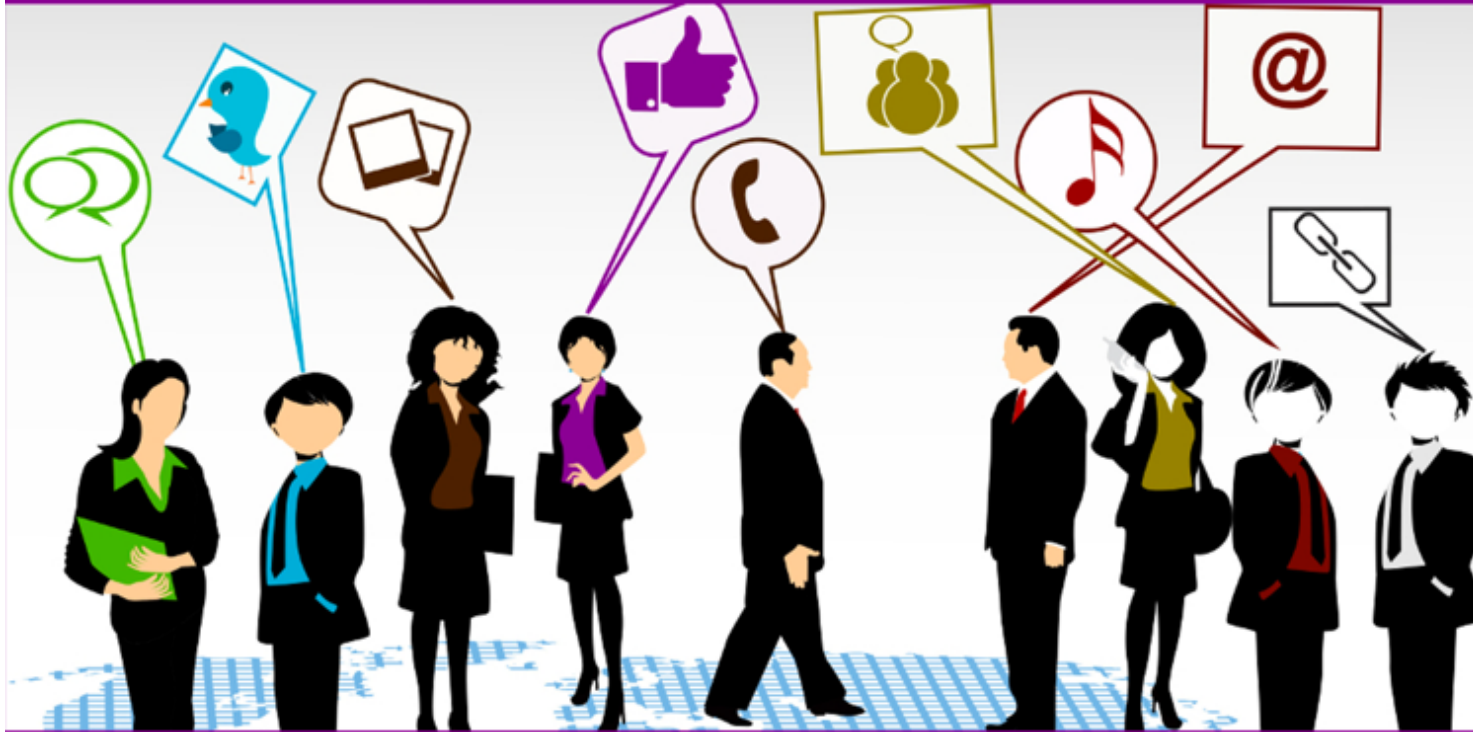
Additional Considerations

- ◆ Remember to answer the “so what” question
- ◆ We tend to report by instrument, not topic, missing opportunities to pull data together across various sources
- ◆ Target report to institutional priorities: mission; strategic plan; state mandates; initiatives; take a look at committee titles to get an idea of current areas of work
- ◆ Add contact information for any follow-up questions
- ◆ Use **bold**, color, images, to draw attention – and headers based on findings – give your readers a path through the report

Final Points

- ◆ Ask before including: is it something you can do anything about? Can we impact this number? Or is this contextual information that needs to be included regardless?
- ◆ Present “less” and “more” findings together – don’t jump back and forth between them. To a reader it is like changing the range on a survey scale from 1 – 5 with 5 being strongly agree to 1 – 5 with 1 being strongly agree.
- ◆ Remember: It’s about the argument being presented and less about methodology. If people attack the methods its because they don’t want to believe the data. They need a compelling narrative and enough information to trust the data being used in an argument.

COMMUNICATE EFFECTIVELY



Organizational Digital Narratives

- What information is currently shared about your program's organizational narrative?
- Is your story public? Discoverable? By whom?
- Who updates it?
- What is the digital narrative your students made about you?



...Evidence- Based Storytelling Toolkit

Questions



Email us: niloa@education.illinois.edu

National Institute for Learning Outcomes Assessment

University of Illinois at Urbana-Champaign
51 Gerty Drive, Suite 196 CRC, MC-672
Champaign, IL 61820

Assignment Library: www.assignmentlibrary.org

Degree Qualifications Profile: www.degreeprofile.org

NILOA Website: www.learningoutcomesassessment.org

Join Our Email List: goo.gl/DDU56q



@NILOA_web



@LearningOutcomesAssessment

Afternoon Handout #1: Evidence-based Storytelling Activity

Assessment Report: Global Citizenship Development Program

Program overview: The Global Citizenship Development Program (GCDP) focuses on global awareness, local community engagement, and global citizenship-related attitudes (e.g., empathy, value for diversity, and responsibility to act). To foster global citizenship development, GCDP uses a multifaceted programming approach throughout the (typical) four-year education at DMU. The program was designed to align with our SLOs, which are listed below. Many program elements are planned and implemented by peer facilitators, who receive training and support from our trained professional staff.

Note: Our main assessment efforts were focused on assessing student learning over the course of their four-year undergraduate experience. Except when noted otherwise, measures were administered at three time points:

- Time 1: Beginning of freshman year
- Time 2: End of sophomore year
- Time 3: End of senior year

Outcomes:

Global Awareness (GA) Outcomes: As a result of completing the GCDP, 85% of participants will...

1. Cite at least 3 current global issues.
2. Apply ethical reasoning to at least one example of global injustice.
3. Experience a 25% increase in personal concern for humanitarian issues.
4. Experience a 25% increase in self-reported interpersonal social interactions with individuals outside of their immediate in-group.

Local Community Engagement (LCE) Outcomes: As a result of completing the Global Citizenship Development Program (GCDP), 85% of participants will...

1. Choose to voluntarily engage in at least 10 hours per year of community service.
2. Cite at least 1 direct impact on the community resulting from their reported community service hours.

Global Citizenship Attitude (GCA) Outcomes: As a result of completing the Global Citizenship Development Program (GCDP), 85% of participants will...

1. Experience a 25% increase in knowledge and use of global perspective-taking
2. Experience a 25% increase in personal value for cultural diversity
3. Experience a 25% increase in feelings of responsibility to act for the betterment of others

Assessment Report: Global Citizenship Development Program

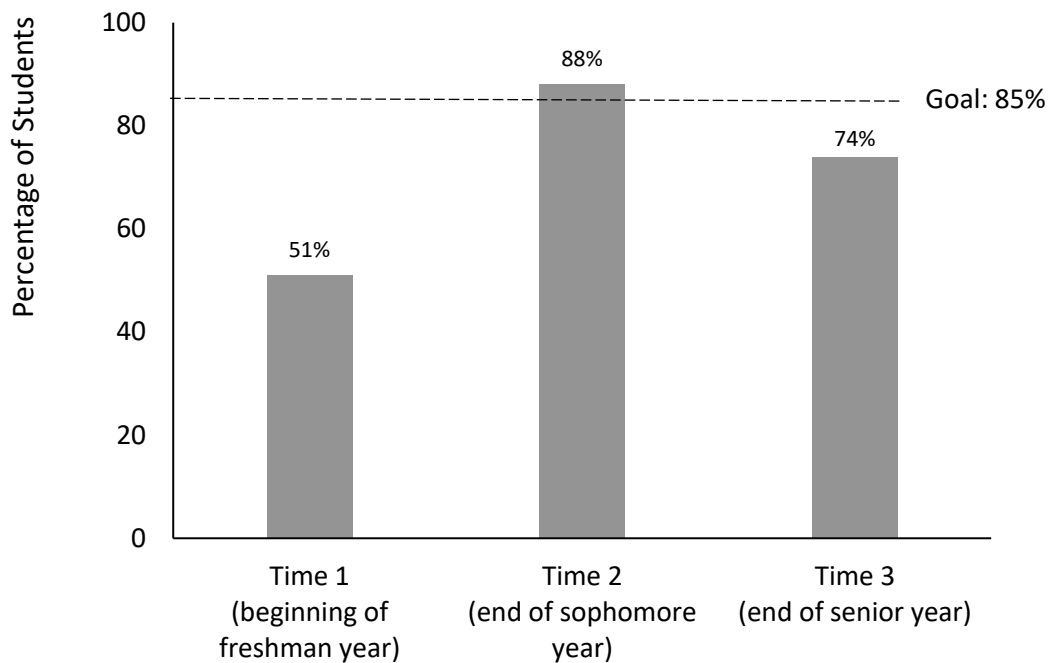
Global Awareness (GA) Outcomes

Outcome GA1: Cite at least 3 current global issues

In order to meet Outcome 1, students needed to be able to cite at least 3 current global issues. Figure 1 shows the percent of students who met Outcome 1 at each time point. At Time 1, slightly over half (51%) of the students were able to cite at least 3 current global issues. By Time 2, 88% of students were able to cite at least 3 current global issues, slightly exceeding our department's goal of 85% of students meeting Outcome 1. However, by Time 3, only 74% of graduating seniors could cite at least 3 current global issues. Therefore, we did not meet our SLO. Future work will investigate why the decrease between Times 2 and 3 occurred so we can try to maintain high levels of global awareness through the latter half of students' careers.

Figure 1

Percent of Students able to Cite 3 Global Issues by Time Point

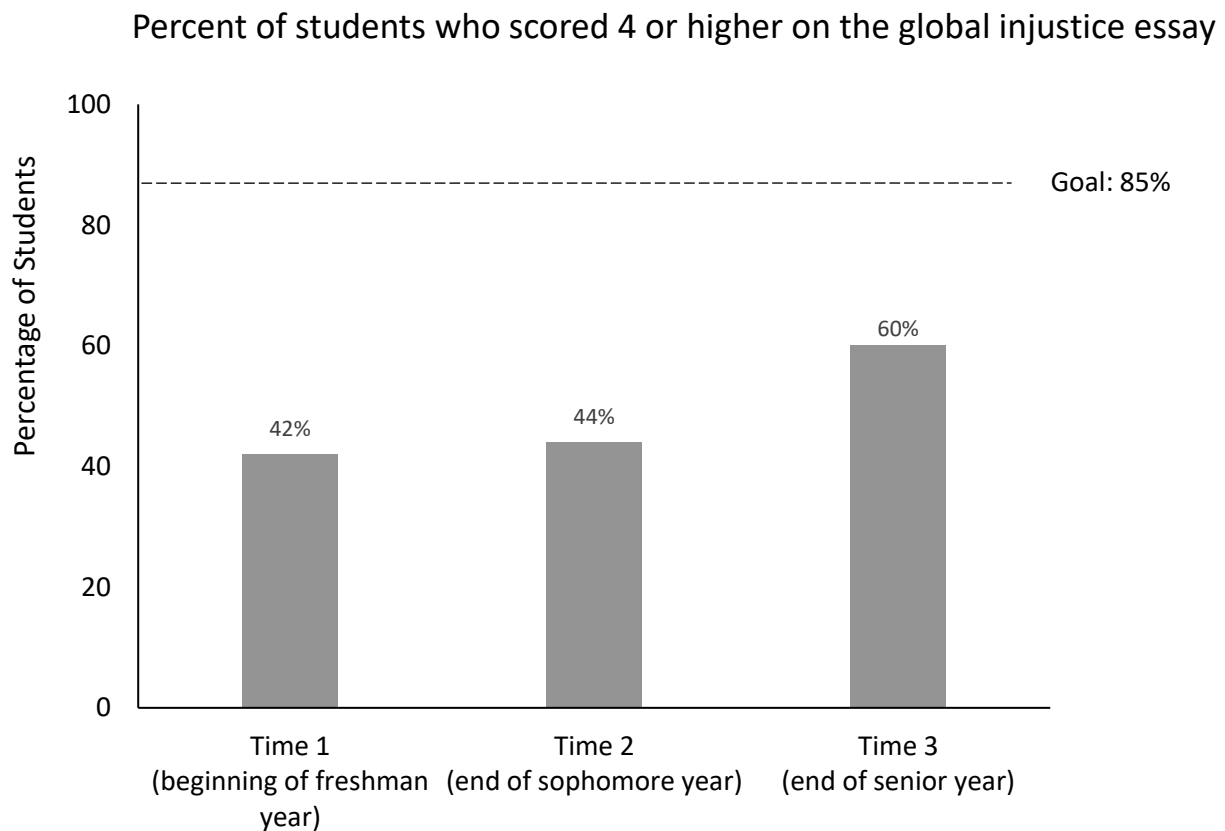


Assessment Report: Global Citizenship Development Program

Outcome GA2: Apply ethical reasoning to at least one example of global injustice

A sample of students completes a global injustice essay at each time point. The essay is rated by trained raters using a high-quality rubric, and the essay prompt is developed in collaboration with ethical reasoning experts across DMU's campus. Students needed to score at least 4 points on the global injustice essay in order to meet Outcome 2. Figure 2 displays the percentage of students who met Outcome 2 at each of the three time points. Results from the Time 1 assessment showed that less than half of first-year students at DMU (42%) scored a 4 or higher on the global injustice essay. By Time 2, 44% of students scored a 4 or higher, showing minimal improvement from Time 1. Results from the Time 3 assessment demonstrate that outcome 2 was not met by the graduating seniors, only 60% of whom scored a 4 or higher on the essay. However, this is an improvement over the Time 1 and Time 2 scores.

Figure 2



Assessment Report: Global Citizenship Development Program

Outcome GA3: Experience a 25% increase in personal concern for humanitarian issues.

A sample of 1000 students was asked to report the number of peer-facilitated sessions they attended throughout the GCDP. At Time 1 and Time 3, students completed a measure of their personal concern for humanitarian issues. Table 1 illustrates the differences between student scores on the measure between students who, at Time 3, reported attending 0, 1-2, or 3 or more peer-facilitated sessions throughout their time at DMU. The group that attended 0 sessions showed a very slight decrease in personal concern for humanitarian issues between Time 1 and Time 3. The students who attended 1-2 sessions showed a slight increase between Time 1 and Time 3. The most dramatic difference between Time 1 and Time 3 scores is seen in the group of students who reported attending 3 or more sessions: on average, this group displayed an increase of 4.5 points between Time 1 and Time 3. These visuals also demonstrate the relative similarity of the groups at Time 1. Students who went on to attend the peer-facilitated sessions at the three frequency levels were generally similar on their personal concern for humanitarian issues scores at Time 1.

Among the students who attended no sessions, 15% demonstrated a score increase of 25% or more between Time 1 and Time 3. Among students who attended 1-2 sessions, 65% demonstrated an increase of at least 25%. Among students who attended 3 or more sessions, 92% demonstrated an increase of at least 25%. Overall, 69% of surveyed students demonstrated a score increase of at least 25% between Time 1 and Time 3. We therefore did not meet SLO GA3 overall, although we did meet this SLO within the group of students who attended peer-facilitated sessions at the highest rate.

Table 1: Personal concern for humanitarian issues across time points by peer-facilitated session attendance

	Number of students	Time 1 average	Time 3 average	Time 1 to Time 3 change	Percent in each group demonstrating increase of $\geq 25\%$
0 sessions	200	3.5	3.4	-0.1	15%
1-2 sessions	300	3.3	5.1	1.8	65%
3+ sessions	500	3.4	7.9	4.5	92%
Average		3.4	6.2	2.8	69%

Outcome GA4: Experience a 25% increase in self-reported interpersonal social interactions with individuals outside of their immediate in-group.

At each of the three time points, a sample of students (N=556) completed a measure of interpersonal interactions with diverse groups. Scores on this measure range from 1 (no interactions with diverse groups) to 27 (daily interactions with diverse groups). A small number of respondents (7%) demonstrated decreases in scores from Time 1 to Time 3, while most students demonstrated increases in scores. In total, 87% of students displayed an increase of 25% or more between Time 1 and Time 3. The SLO—that 85% of students would experience a 25% increase in self-reported interpersonal social interactions with individuals outside of their immediate in-group—was therefore met.

Assessment Report: Global Citizenship Development Program

Local Community Engagement (LCE) Outcomes

Outcome LCE1: Choose to voluntarily engage in at least 10 hours per year of community service.

At the end of the spring 2018 semester, we asked 800 students (200 freshman, 200 sophomores, 200 juniors, and 200 seniors) to report the number of hours they spent on community service in the past year, excluding hours required by student organizations, classes, or disciplinary interventions. Table 2 displays the reported community service hours for the students as a whole, as well as for each class standing group (freshmen, sophomores, juniors, and seniors). Among all respondents, 71% reported spending 10 hours or more engaging in voluntary community service in the past year.

Generally, higher class standing was associated with larger percentages of students reporting high levels of volunteering. The proportion of students reporting no hours of community service was highest for freshmen (15%), followed by sophomores (11%). Juniors and seniors displayed very low rates of non-volunteerism (3% and 1%, respectively). Seniors were the most likely to report spending more than 10 hours volunteering (95%), followed by juniors (81%). 68% of sophomores and 40% of freshmen reported spending more than 10 hours volunteering. Outcome LCE1 identifies that, after participating in the GCDP, 85% of students should choose to voluntarily engage in at least 10 hours of community service per year. Therefore, the key group to examine is seniors, as they had received the entire GCDP program by the time of the survey. Among this group of students, we reached our goal of more than 85% of seniors volunteering for more than 10 hours per year.

Table 2: Community service by class standing

	Freshmen	Sophomores	Juniors	Seniors	Total
No hours	15%	11%	3%	1%	8%
9 hours or fewer	45%	21%	16%	4%	22%
10 hours or more	40%	68%	81%	95%	71%

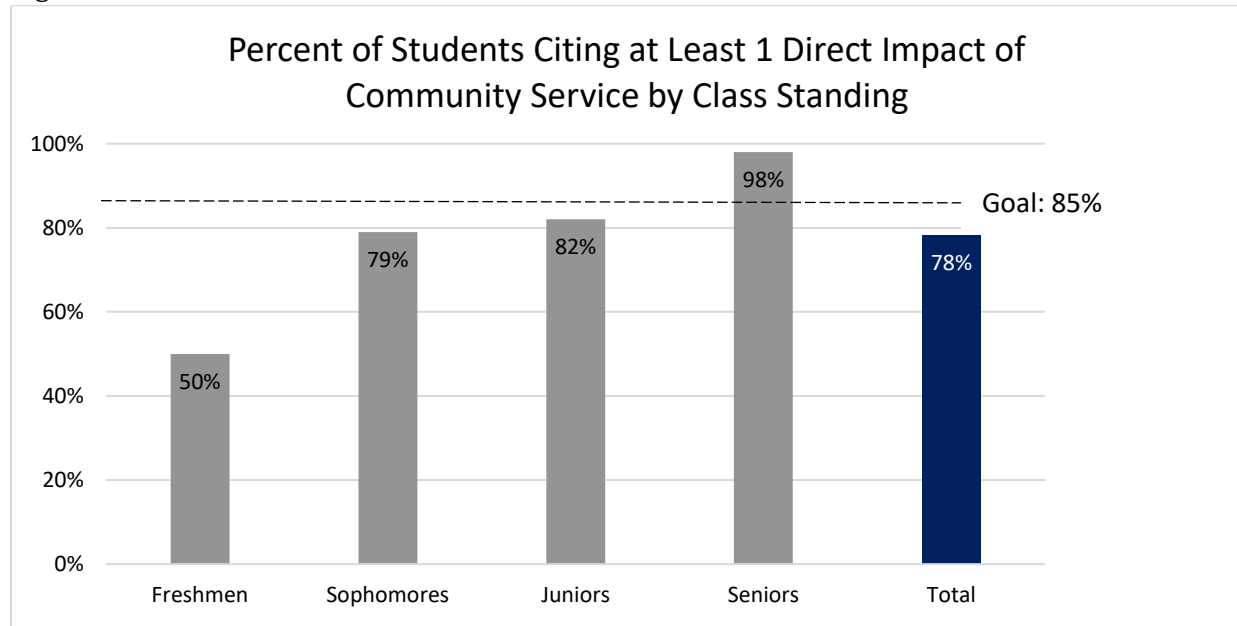
Assessment Report: Global Citizenship Development Program

Outcome LCE2: Cite at least 1 direct impact on the community resulting from their reported community service hours.

Assessment of Outcome LCE2 took place alongside assessment of Outcome LCE1. Students who reported participating in community service hours (e.g., all students who did not report spending “no hours” participating in community service in the past year) were asked a follow-up question: “Please list any direct impacts of your community service work on the community in which you volunteered.” As shown in Figure 3, 98% of seniors who had participated in community service were able to identify at least one direct impact of their community service. As reported in Table 2, 99% of seniors participated in community service last year. Together, this means that 97% of surveyed seniors were able to identify at least one direct impact of their service.

Respondents were also asked to identify the community (or communities) impacted by their service. Table 3 contains excerpts of students’ responses to this question.

Figure 3



Assessment Report: Global Citizenship Development Program

Table 3

Class standing	Community impacted	Quote
Freshman	Local school	"Because of our volunteer hours, the school's cafeteria has a bright new mural. Our corresponding teacher told us the students love looking at it during their lunchtime."
Freshman	Youth organization	"By volunteering as coaches, we provided local youth with a safe and healthy after-school activity. I got to use my love of soccer to help kids have a good time, build their self-confidence, and get some fresh air and good exercise along the way."
Sophomore	LGBTQIA+ community, youth organization	"Our time was spent putting together decorations for a local LGBT youth organization's prom celebration. Our work helped create an awesome, fun night for local LGBT students. It helped me, too: I feel a lot more connected to the town now than I did my freshman year, when I felt like I was kind of trapped inside the campus bubble. But now I know there are people outside of the university who are doing good work and who care about other people."
Sophomore	Religious community, elderly populations	"I helped out with a Hanukkah celebration at a nearby nursing home. Some of the folks living there don't have family nearby, so I think having some college kids around helped them feel less alone during the holiday."
Junior	Housing-insecure populations, food-insecure populations	"We were responsible for cooking meals at a homeless shelter. I think the homeless population definitely benefited from my volunteer hours, because one lady told me she hadn't had a warm meal all week. But I think what I liked the most about the experience was getting to work with this amazing organization that does so much good in the community. I learned so much from the other volunteers. I can't wait to go back."
Junior	Refugee community, youth organization	"I think the most obvious group that was impacted by my volunteer hours was the refugee kids because they're the ones who might not have had access structured after-school activities. But I think all of the kids—even the ones who aren't refugees—were impacted, because the organization brings all these kids together to learn new skills."
Senior	Local school	"Our volunteer hours helped the school raise money for new science equipment, so the school was the community we impacted. Now, students will be able to do more hands-on stuff during science class. And the teachers won't have to pay for as much on their own anymore."
Senior	Food-insecure populations, elderly populations	"My hours included a 'food rescue' mission—where we collected food that was still fresh but was going to be thrown out anyway—and a 'chef mission'—where we took the food and turned it into single-serving frozen meals. Then we delivered to elderly folks around town, especially those who might have trouble preparing their own meals. Too often, I think, these folks are forgotten in a busy college town full of young adults. But they're real people with real needs and they have so much to contribute. I think a lot of my classmates could learn a lot from volunteering with this organization."

Assessment Report: Global Citizenship Development Program

Global Citizenship Attitude (GCA) Outcomes

Outcome GCA1: Experience a 25% increase in knowledge and use of global perspective-taking

The GCA outcomes were targeted by peer-facilitated sessions that sought to introduce DMU students to issues of global citizenship. In particular, global perspective-taking was defined as students' abilities to recognize that people from different cultural backgrounds have different ways of understanding truth and, therefore, of interacting with the world around them. Students who are high in global perspective-taking are able to use this knowledge to predict different ways that people from varying backgrounds might view the same situation.

Global perspective taking was assessed via the Pertinent Rational Scenario Perspective Consideration Test (PeRSPeCT). The measure provides both an estimate of students' knowledge of global perspective-taking (on a scale from 1-70) and an estimate of the frequency with which students use these skills (infrequently, moderately, or frequently). Students were sent a link to the measure to complete as a pre-test during the week prior to a three-hour peer -facilitated workshop. They also completed the measure as a post-test a month after the workshop. Figure 4 displays the mean difference between the pre-test and the post-test. Overall, 95% of students displayed an increase of 25% or more on the PeRSPeCT: on average, students' scores increased 32% from the week before the workshop to the month after the workshop. Therefore, we have achieved SLO GCA1.

Figure 5 displays frequency of global perspective taking skill use prior to the workshop and following the workshop. Students reported more frequent use of global perspective taking skills following the workshop than they did prior to the workshop. Prior to the workshop, 65% of students "infrequently" used global perspective taking skills, and only 15% "frequently" used these skills. However, following the workshop, only 12% reported "infrequently" using these skills, while 60% reported frequently using these skills.

Assessment Report: Global Citizenship Development Program

Figure 4

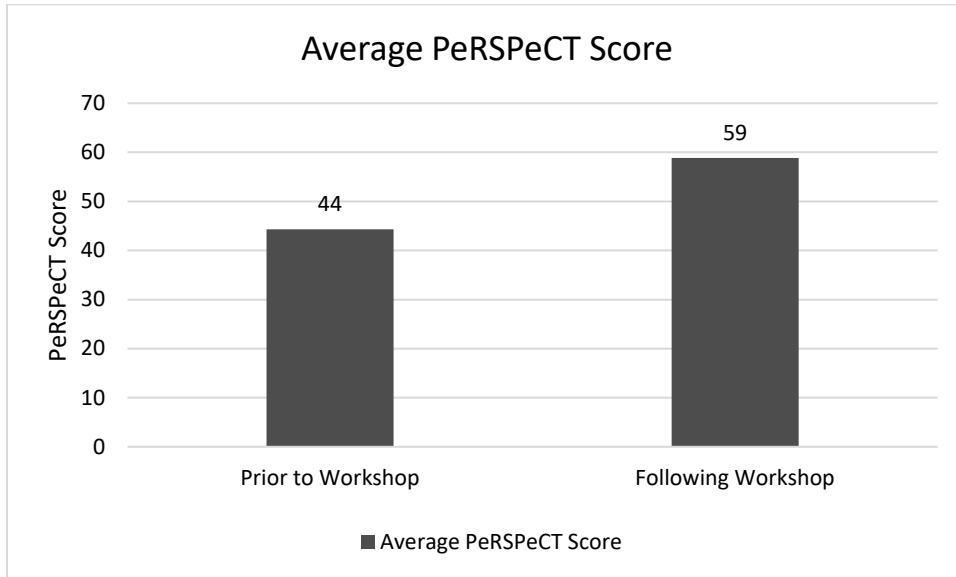
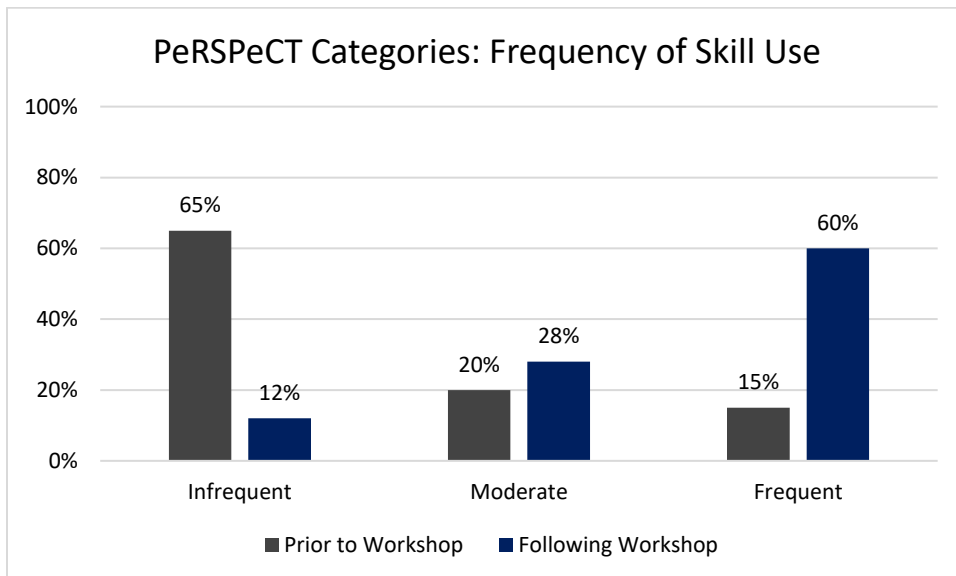


Figure 5



Assessment Report: Global Citizenship Development Program

Outcome GCA2: Experience a 25% increase in personal value for cultural diversity

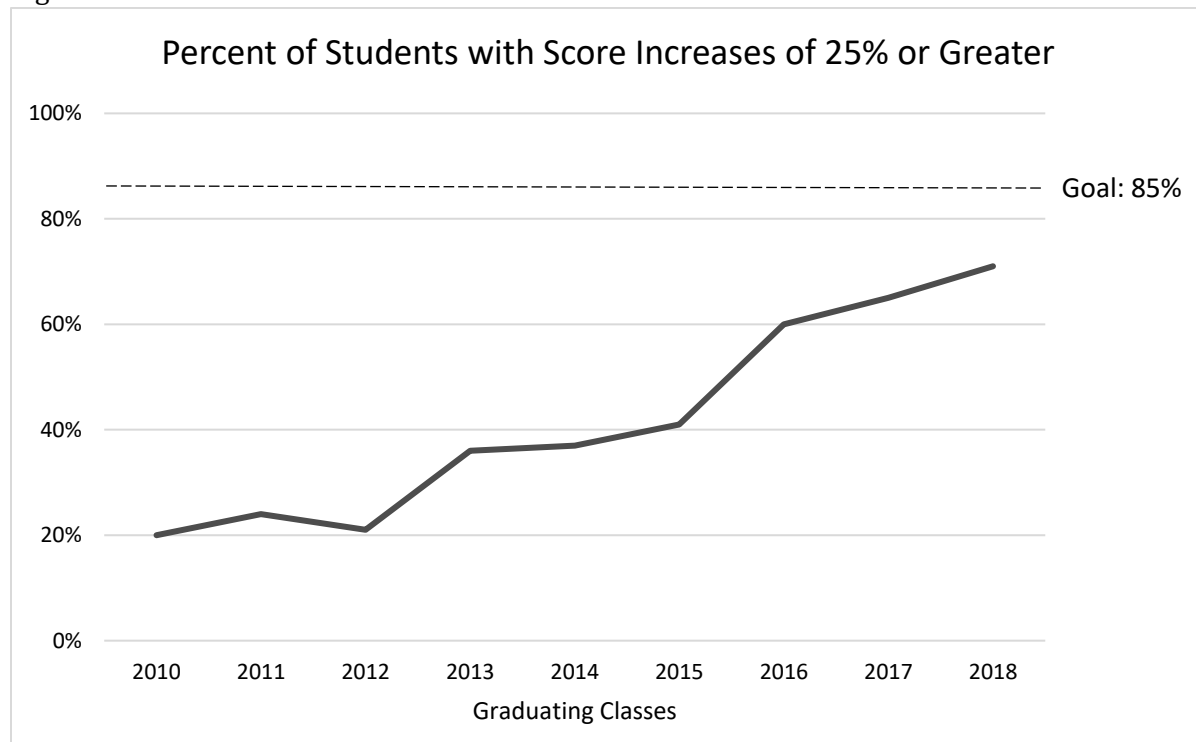
Personal value for cultural diversity was measured via a survey sent to a sample of 200 students at each of the three time points. The scale ranged from 1 to 20, with lower scores indicating lower value for cultural diversity. Each sampled student completed the measure all three times. Students' scores increased slightly between Time 1 and Time 2, but then declined slightly by Time 3. From Time 1 to Time 3, only 30% of sampled students displayed an increase in personal value for cultural diversity of 25% or higher. Therefore, we have not yet achieved SLO GCA2.

Outcome GCA3: Experience a 25% increase in feelings of responsibility to act for the betterment of others

For each entering class of students, 100 students are sampled to evaluate SLO GCA3. Their feelings of responsibility to act for the betterment of others are measured via the Responsibility for Change and Improvement (RCI) scale at Time 1, Time 2, and Time 3. The percentage change in RCI scores between Time 1 and Time 3 is then calculated for each student. Figure 5 displays the percent of students displaying a 25% or greater increase in RCI scores from Time 1 to Time 3 for the nine most recent graduating classes.

As seen in Figure 6, we have been making steady progress toward SLO GCA3 over the past decade. Although we have not met our goal of 85% of students displaying an increase of 25% or greater, the trend indicates that we are making good progress.

Figure 6



Evidence-based Storytelling Activity

Afternoon Handout #2

Scenario 1

You are reporting to the general student body. You are trying to convince them to participate in your program during a student activities fair during New Student Orientation. They want to know how your program will help them reach their goals after college.

Scenario 2

You have been asked to speak to your institution's governing board at their annual retreat. The members want to know about the state of student learning at the institution, which has recently been criticized in local media. (Yours is only one of several presentations the board will hear.)

Scenario 3

You are reporting to a committee at your institution that has been tasked with making recommendations about future resource allocations. The committee includes academic deans, the VP for student affairs, the VP for advancement and alumni relations, the chair of the Faculty Senate, and the president of the student government, among others. They want information that will help them determine where this program should fall in the list of funding priorities.

	Message for...	Audience for...
GMU	Scenario 2	Scenario 1
JMU	Scenario 1	Scenario 3
ODU	Scenario 3	Scenario 2
UVA	Scenario 1	Scenario 3
VCU	Scenario 2	Scenario 1
VT	Scenario 3	Scenario 2

Evidence-based Storytelling Activity: Facilitator Guide

Purpose: The purpose of the activity is to address the learning outcome “*Apply [evidence-based storytelling] skills to a variety of scenarios and for a variety of audiences.*”

Length: 1 hour 45 minutes

Outline

- 5 minutes: Gianina will introduce the activity, Jodi will explain the logistics.
- 30 minutes: Group work *at individual location (site host leads; time allocations below are suggested – use your judgment)*
 - 5-7 minutes: Group members read the assessment report and scenario provided.
 - 13-15 minutes: Participants work in small groups (3-5 people) to craft a message in response to the assigned scenario.
 - 10-12 minutes: Groups come together to compare messages and decide which message to present to the entire assembly (via telepresence). They may choose one group’s message or combine elements of different groups’ messages.
 - 1 minute: Assign (or ask for volunteer) a spokesperson for your location.
- 15 minutes: After the 10-minute comfort break, one site will be called on to present their message for Scenario 1 via telepresence. The other site with the same scenario will be asked if they have anything to add. Then, participants at the sites assigned as the audience for that scenario will be invited to ask questions (appropriate to their particular role as students, board members, or committee members). Either spokesperson may answer the questions, and they may ask other group members to chime in if they wish. In answering questions, participants may embellish their responses with information not included in the report, but they may not change the basic findings of the assessment. (*Jodi moderates*)
- 30 minutes: Repeat the previous procedure for Scenarios 2 and 3.
- 25 minutes: Gianina will offer thoughts on what she heard during the presentations. Participants at all sites will be invited to use Slido to ask questions and/or respond to reflection questions from Gianina. Jodi will monitor Slido and present questions/reflections that generate high numbers of “likes.” Gianina will respond/comment as appropriate.