
activity-based assessment

integrating evaluation into prevention curricula



TOOLS FOR CHANGE

a publication of the Texas Association Against Sexual Assault & the Texas Council on Family Violence

Written and developed by

Morgan J Curtis, LMSW on behalf of the Texas Association Against Sexual Assault

Surabhi Kukké, MPH on behalf of the Texas Council on Family Violence

TABLE OF CONTENTS

- 4 What is Activity-Based Assessment? (And What isn't It?)
- 5 The Improvement Loop
- 6 Step 1: Establish Learning Objectives
- 7 Step 2: Identify Learning Activities
Step 3: Identify Application Activities
- 8 Step 4: Data Collection and Analysis
- 20 Steps 5 & 6: Revise Program Plan or Curriculum
- 22 Appendix A: Program Evaluation Resources
- 23 Appendix B: Sample Process Notes
- 25 Appendix C: Case Study
- 29 Appendix D: Worksheet for Creating Your First Scoring Tool

A Note to the Reader

Activity-based assessment is an evaluation methodology that focuses on curriculum-level program improvement. This approach is designed especially for programs educational curricula to achieve social change goals. Social-change efforts are invariably long-term and require a variety of components to be successful. Interactive, educational curricula that invite participants to reflect on their values and perspectives and consider alternatives to norms that are harmful are one important component that is widely used. However, evaluation of these types of programs is often at the macro level (i.e., measuring the impact of the entire program), missing a rich opportunity for assessing learner development and program improvement.

As a facilitator of social- and behavioral-change-oriented curricula, you have the opportunity to notice and document growth and change like few others at your organization. This toolkit will help you take a fresh look at your curricula and educational materials to develop simple strategies to assess your participants' understanding of the material and their ability to apply the knowledge they have gained.

What Is Activity-Based Assessment? (And What Isn't It?)

It may help to begin by clearing up what we are not talking about in this toolkit. This is not a guide to outcomes evaluation. You will not find evaluation information or tools based on traditional research design using pre-tests and post-tests and comparative data analysis. We also do not address measures of participant satisfaction.²

Activity-based assessment is a method that integrates evaluation into educational curricula. It is a simple methodology that places at its center activities that are already designed for educational purposes and draws from them useful information for facilitation guidance and program improvement.

This toolkit uses the architecture of the curriculum and weaves into it an improvement loop to assess both participant knowledge and comprehension, as well as to improve facilitator practice. In this way, evaluation becomes an organic component of curriculum design and implementation. And most importantly, you will not have to do extra work to evaluate your process; it will be built into your practice.

In order to make evaluation practice meaningful, we want to keep it simple and systematic and have it result in a story that we want to tell. Taking the time to pay attention to each educational session can be very valuable but also has the potential to create mountains of data. Making it meaningful means asking yourself strategic questions:

What do I need to know?

What matters?

What do I have to collect to help me do this work better or differently?

Answering questions like this at the outset will help you collect the minimum amount of data necessary. This process also invites your own subjectivity as a facilitator into the evaluation practice. We can all recognize that evaluation that is requested by outside parties does not typically feel meaningful or useful. A meaningful evaluation practice will help you tell a story about your work that you want to see amplified and make decisions about how to adjust your own practices as a facilitator.

¹The term “participants” will be used throughout this toolkit to describe the people who are participating in educational sessions. This term is meant to apply to people of all ages and is not specific to young people or adults.

²For more information on outcomes evaluation, see the resources listed in Appendix A.

Thinking Like an Evaluator

Facilitators are evaluating all the time. When you gauge the interest and attentiveness in the room, you are assessing whether you need to do something differently to get the energy moving. When you ask review questions or involve the participants in recapping earlier lessons, you are evaluating comprehension. Using evaluative thinking is a powerful way to enhance your facilitation and program outcomes.

With all of this in mind, activity-based assessment methodology invites the educator to think like an evaluator. When selecting activities for a curriculum, she might ask herself, "How will I know if my students have really understood what I am trying to teach? What will I see if they have integrated the concepts or ideas? What will success look like?"

These are critical first questions for a facilitator to ask.

Starting with an idea of what success looks like will help you know it when you see it! But just knowing it will not be enough; it is critical to have a method of recording this success as well. Scoring tools are one of the most useful tools for recording the assessments that you make during a session. These simple rubrics, which are developed before the curriculum is implemented, help the facilitator see if the activity has achieved its desired intention -- namely whether the participants have learned what they were supposed to learn.

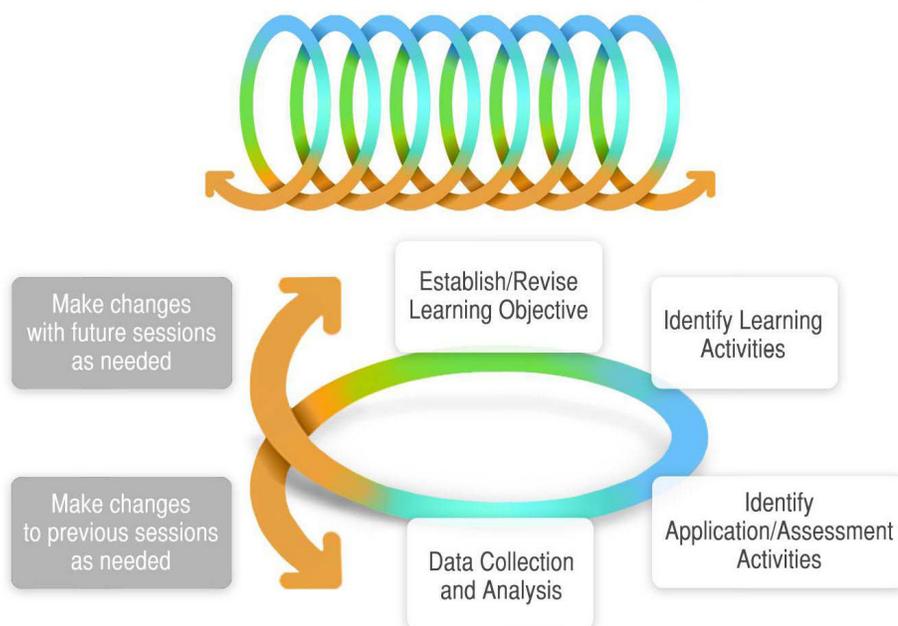
The other important aspect of thinking like an evaluator is knowing in advance how the information you gather will be used. As with any evaluation, we never want to collect more information than we need. As you build your scoring tools, consider how the information you collect will be used to improve your facilitation and curriculum. If you won't use it, don't collect it!

A final point about thinking like an evaluator: when planning the flow of the curriculum, specific learning objectives are critical to supporting the methodical progression of learning. Learning objectives are also the backbone of activity-based assessment. They are an essential touchstone while planning and developing scoring tools and other methods of monitoring progress. The development of strong learning objectives will be discussed throughout this toolkit.

The Improvement Loop

The improvement loop that follows walks us through this model of assessment and is much like any model for program improvement - the difference being the micro-level at which this feedback loop is operating. With this model, we are examining our curricula at the activity level to assess whether or not each successive activity in a session is taking us in the direction of learning and application. The loop highlights several important steps in planning and using data to improve your practice.

The Improvement Loop



Step 1: Establish Learning Objectives

Learning objectives³ are the core of any curriculum; each session should be built around action-oriented learning objectives that clearly delineate the change participants should experience as a result of participation in that session. That is, learning objectives put words on what success will look like in a given session.

As part of thinking like an evaluator, these learning objectives will guide not only your selection of activities but also your approach to teaching/facilitating those activities.

The following is an example of a learning objective:

At the end of this session, participants can identify ways gender norms operate in society.

Notice that “identify” is underlined. With all learning objectives, specific focus should be paid to the verb⁴, especially for the purposes of evaluation. In this case, we’ve stated that success for this educational session includes participants being able to identify how gender norms operate in society. Since that’s the case, we don’t want to try to measure whether or not they can explain what gender norms are (for example).

Getting Started

As you go through this process, you might find that it will reveal a lot of possibilities for how to collect information through your facilitation.

We encourage you to go slowly, perhaps just starting with one scoring tool for one activity so that you can practice this before you try to apply it curriculum-wide.

³For more information on learning objectives (also called learning outcomes), check out this resource: <http://www.humber.ca/centreforteachingandlearning/instructional-strategies/teaching-methods/course-development-tools/writing-learning-outcomes.html>

⁴For more action-oriented verbs, check out Bloom’s taxonomy: <http://www.taasa.org/wp-content/uploads/2012/04/Working-on-the-Wow-Side-Handout-31.pdf>

Step 2: Identify Learning Activities

For the purposes of this model, the term “learning activity” applies to any educational activity that is meant to teach to the learning objectives. Generally speaking, you might have one learning activity to meet one or two learning objectives. Here are some examples of learning activities:

Lecture

Guided discussion

Video

Pre-scripted role plays

Previously, we used the learning objective “identify ways gender norms operate in society.” One relevant learning activity that teaches to this objective is the Gender Box (also known as Man Box/Lady Box)⁵ activity that explores expectations that society places on men and women and implications/consequences of those expectations.

There are many ways to facilitate the Man Box/Lady Box activity, depending on the desired outcome of the activity and conversation. Since we have a specific learning objective tied to this activity, the way the activity is facilitated should be geared toward that learning objective. When facilitating this activity, you might focus more on the *places* we learn these stereotypes/expectations and *how* we learn them than on the societal consequences are for stepping outside of those expectations. This will guide students toward identifying these norms in context, which fits with the objective. (Whereas, if the objective was to help them identify the consequences, you might spend more time on that piece.)

Step 3: Identify Application Activities

In order to increase the likelihood that learning will be integrated, participants need an opportunity to practice or apply their new knowledge or skills. Application activities are the activities that give participants opportunities to apply what they learned in learning activities. (There is not always a clear distinction between learning activities and application activities, as sometimes both learning and application will happen in the same activity. The more critical point is to understand where application is happening, because that is a key point for assessment.) Application activities could also be called assessment activities since these are the activities that allow us to assess participant learning. Again, it's important to revisit to the learning objective to see what exactly participants should be applying in the activity. Remember to refer back to the verb in the objective and make sure you're assessing the action that it indicates.

⁵For information about this activity, see the description: <https://www.pcc.edu/resources/illumination/documents/gender-role-boxes-glbtcq-and-sexism-exercise.pdf>

The following are examples of application activities:

Role plays

Guided discussions

Worksheets

Journaling or other writing exercises

Note, again, that the same kinds of activities can be either learning activities or application activities (or both) depending on their focus.

For the learning objective and learning activity examples listed above, an appropriate application activity might be to have participants watch a movie and pick out moments where gender norms are reinforced. (Notice that this lines up with “identify,” which was the verb from the learning objective.) For the purposes of this example, let’s say that the application activity will involve watching specific clips from the movie *Twilight* and identifying examples of gender-role socialization.

Step 4: Data Collection and Analysis

Once you’ve decided what you want participants to learn, how they’ll learn it, and how they’ll apply it, it’s time to consider how you’ll assess their learning. In order to do that, you’ll need to develop a scoring tool that will be used with the application activity. There are different types of scoring tools that you can create, and each type gives slightly different kinds of data.

Participants As Data Collectors

Another option in data collection is to involve the participants themselves in a peer review process. This can increase their own integration of the concepts by helping them identify the points of learning as other people demonstrate them, and it might also keep them engaged during moments when it would be otherwise easy to tune out. As with involving non-participant data collectors, you’ll want to establish very clear, collective ideas about how to complete the scoring tools. Since training them on the scoring tools does also give them an added dose of learning around the concepts at hand, it might improve their own likelihood of success for applying the behavior or skills. If those participants will also be engaging in and scored on their performance in the application activity, this might skew your data. You can consider having a few students involved in scoring who don’t also engage in the application activity. An additional option is to make sure that every participant is given an opportunity to rate another person or group. As with any other process you’re facilitating, this can impact and be impacted by group dynamics (e.g., level of trust, power dynamics), so use your discretion about whether or not this will be a productive/useful process for all parties involved.

Develop Scoring Tools

The top of the scoring tool should include any information about the class and participants that will be relevant context for the data. You definitely want to include the following:

Class/location/school

Day

Session #

Date

Additionally, it might be helpful to include some general demographic information about the participants. This kind of information will help contextualize your experience and your analysis of process notes (more about those in Data Collection). For example, a discussion about gender roles may go very differently if there are all girls in the room as opposed to a mixed-gender class. Such observations can help inform how you might approach facilitation given the composition of your class, and this information might also help you understand how to use the data from your scoring tool to make changes. Some demographic information that is useful to collect:

Gender breakdown

Racial composition

Grade levels

Setting (e.g., classroom, lunchroom, auditorium, gymnasium)

Extenuating circumstances (e.g., if participants are dealing with a big issue outside of class)

Any other outstanding feature of the day's session

The following chart describes three types of scoring tools you might create and lists some of the pros and cons to consider when deciding which tool is right for your needs.

	Description	Pros	Cons
Checklist	A list of desired responses, marked with yes or no; often the basis for counts and scales answers the question: did (something) happen?	Easy to collect	Only yes/no assessments Lacks detail
Count	The number of times a desired response was shared and/or The number of participants who share a desired response Answers the question: how many times did (something) happen? or how many participants did (something)?	Easy to collect Frequencies are easy to calculate and understand	Does not allow for a nuanced analysis of participant learning
Scale	Categorizes a range of possible responses (e.g., from low to high or bad to great); answers can fall on either end or somewhere in between Answers the question: how well did participants do (something)?	Allows for more subjectivity More nuanced results than the checklist or count options	More detail to manage, more time consuming data analysis More complicated collection might require additional raters Needs rubric for setting up criteria More difficult to compare between sessions

Checklist Template

If students are doing a role play, you might develop a checklist to indicate whether or not they demonstrated specific skills or ideas that they were taught. You would simply check the appropriate box indicating whether or not the item on the sheet was observed. A scoring tool for this might look like the following example (next page).

Group 1 - Role play	Yes	No
Demonstrated Skill 1		
Demonstrated Skill 2		
Used Language/Terms from the Lesson		
Successfully Solved the Scenario		

(You can eliminate the “no” box, of course, but it is sometimes helpful to be able to add check marks for all items just to be sure that they were all assessed.)

Count Template

If you want to know how many times a behavior or skill is demonstrated and/or how many students demonstrate it, then a count-based scoring tool is what you need. You simply list the behaviors of interest and make a hash mark when the behavior is observed.

Incident	# of students/groups
Demonstrated Skill 1	
Demonstrated Skill 2	
Used Language/Terms from the Lesson	
Successfully Solved the Scenario	

Scale Template

If you’d like to gather data that indicates how well participants demonstrated the skills or information they learned, you could use a scale that looks more like the example below. As mentioned in the chart, this scoring tool gives you slightly more nuance than does the checklist above but might also make the data collection a little more taxing.

Group 1 - Role play	Yay 4	Sure 3	Eh 2	Boo 1
Demonstrated Skill 1				
Demonstrated Skill 2				
Used Language/Terms from the Lesson				
Successfully Solved the Scenario				

Twilight Example (Count)

For the Twilight activity example, a count-based scoring tool is a great option for assessing the application activity. For the application activity, you would select certain clips that the participants would view, and you would identify in advance what the examples of gender-role socialization are in those clips. Creating a scoring tool with those examples in the left-hand column, you can then use the right-hand column to make hash marks for each student that identifies the incident. That identification can happen through written worksheets, through a show of hands when the incident occurs, or by any other means that would allow the students to indicate that they have identified the incidents. Keep in mind that the options that involve learners identifying something through hand raising or other public means will allow for other students to follow along even if they haven't identified the moment themselves. (Since the verb we are working with is identify, the two most appropriate scoring tools to assess it are checklists and counts. Measuring how well the participants identified these instances of gender-role socialization wouldn't add much to your understanding of the success of this given activity.)

Incident	# of students/groups
Specific time when Bella does something completely irrational/emotional	
Edward refuses to let Bella do something for herself	
Bella is seen doing housework or grocery shopping for her father	
Jacob and Edward fight over what is best for Bella	

Data Collection

When thinking about data collection, there are three key factors to consider:

Who will collect the data?

How will the data be collected, namely what tools or instruments will be used?

Where or how will the data be stored?

Who: Once the scoring tools are developed, who collects the data will largely be a pragmatic question. Are you on your own in the classroom, or is there a teacher in there who could help you? Do you conduct your sessions in pairs or solo? Is there an option of bringing an intern with you?

Collecting data and facilitating at the same time can be tricky.

Trying to assess participant engagement while you are facilitating can introduce bias; we all tend to see what we want to see and ultimately want to feel like we have achieved our objectives. This tendency can skew our assessment of how things went.

Not only is it challenging to do these things at the same time, more importantly, it can be hard to do them both well. Being distracted with a scoring tool while participants are presenting their work can affect your facilitation as well as the dynamic in the room.

If you know you will be the data collector, factor this into the scoring tool development; you want it to be as simple as possible to fill out because you will be collecting the data while facilitating.

If you have someone who can help you with data collection, be sure to discuss the scoring tools with them in advance so there is clarity about the criteria you have developed and the data can be as reliable as possible. Share some examples with them about the types of attitudes or behaviors that would fit what you're hoping to see (or not see). Review the sheets together afterward, too, to continue refining your understanding of the conditions that make each session successful.

Subjectivity

While evaluation is a form of research, activity-based assessment is not focused on getting to some objective truth about the participants. Because this is geared toward program improvement, the assessment method is inherently subjective. We can manage the subjectivity by having other people collect the data, but there is very little we can do to completely control for it.

So, with that in mind, accept the subjectivity in your assessment! In fact, use your perspective strategically to inform your assessments. You know your populations best; you may be able to notice subtle shifts in behavior, attitude, or response during implementation, and, begin documenting them.

How: Alongside the scoring tool, process notes are a key data collection instrument for activity based assessment. This is a simple log or journal that you keep as a facilitator on how things went during each session of the curriculum.

At minimum, a basic activity log will capture data like the number of participants, notable absentees, the date, and session number – information necessary to document the outputs of your session.

Process notes are also an opportunity to capture a variety of factors that may have affected the outcome of the session:

Did you get through the activity as planned, or did you need more time?

If things went especially well (or not), what do you think contributed to that?

What kinds of activities keep the participants most interested? Does your class gravitate toward drawing/art, role play/skits, writing?

What kinds of questions are participants asking?

How can you tell that they are engaging with the material?

The answers to these kinds of questions can directly inform program improvement one session at a time and provide important context for the data collected with the scoring tool. You can simply create a sheet with these questions (or others that you deem important) that allows you to jot down answers after the session. Appendix B shows a more detailed way to set up process notes.

Data Organization and Storage

Now that you have all this collected data what do you do with it? Keep it!

If you are using paper and pen to collect data, labeling and filing data sets correctly will be very helpful for analysis and use (see the next section). When labeling your data, be sure to keep sets separated by class or session as opposed to day. For instance, if you facilitate three sessions a day, clustering all the data from the sessions on that day will not help you address the needs or improvements for each class.

Depending on the complexity of your scoring tool, you may find that you can scan the results and quickly glean the information that guides next steps you need to take. In cases where the tool is simple, you are encouraged to incorporate this kind of evaluative thinking into your review of the data. When the tool is more complicated, such as ones involving scales, do your best to enter the data into a spreadsheet after each session. This will enable you to easily analyze and store the data. Additionally, entering it right after the session is better when you need to recall specific details (e.g., written examples included in your scoring tools). After you enter the data into a spreadsheet, you can hold on to the pages of raw data or shred them depending on storage space you have available.

If you are using a tablet to collect data, using a program like Survey Monkey, Survey Gizmo, or Google Forms, be sure to download and back up your data sets periodically.

To Use Technology or Not to Use Technology?

Using technology such as tablets or mini computers for data collection can seem like a great simplifier, but let's look at some of the pros and cons of going down this path:

Pros

- Ease of use
- Collects data in one place
- Paperless data storage
- No second step of data entry required
- Can do analyses fairly easily through data collection applications (e.g., Survey Monkey, Survey Gizmo, or Google Forms)

Cons

- Tablets can be expensive
- Need some familiarity with data collection applications
- Need wi-fi or data plan
- Technology can fail you
- Need to back up data
- Must keep tablet charged!

Take into account what resources your agency has to provide and also how you will use the data you collect. If you decide to use technology, always have a backup plan (perhaps hard copies of the tools) in case the technology fails you!

Analysis and Use

As suggested by the improvement loop, analyzing the data will help you answer two primary questions:

- 1. What, if anything, do you need to do differently the next time you facilitate these activities?**
- 2. What, if anything, do you need to do next with this group of people to make sure that their learning for the next session is on track?**

Sometimes the data we collect suggest that we are making the changes we hoped and intended to make; sometimes the data tell a different story. How do we know what the data are telling us? What kind of changes do we need to make based on the story the data tells? This section will walk you through several examples that highlight the process of analysis and use of activity-based assessment data.

Continuing with the *Twilight* example, let's say we collected the following data during one session:

Incident	# of student groups
Specific time when Bella does something completely irrational/emotional	18
Edward refuses to let Bella do something for herself	21
Bella is seen doing housework or grocery shopping for her father	16
Jacob and Edward fight over what is best for Bella	19

A cursory look at these data makes it difficult to tell how well things went, because you can't compare the numbers to the total number of students nor see any percentages or averages. Those types of numbers are a little bit easier to deal with for making statements about success or necessary changes.

To do the analysis on the data for just this session, you could create a spreadsheet in Excel or Google Docs that would look like this^{6,7}:

	A	B	C	D
1	Incident	# of students who identified it	Total Students	%
2	<i>Specific time when Bella does something completely irrational/emotional</i>	18	25	72
3	<i>Edward refuses to let Bella do something for herself</i>	21	25	84
4	<i>Bella is seen doing housework or grocery shopping for her father</i>	16	25	64
5	<i>Jacob and Edward fight over what is best for Bella</i>	19	25	76
7				
8	AVERAGE			74

So, what do these data tell us? It looks like all of the examples from the activity were identified by between 64-84% of students. If you have predetermined that you are hoping for at least 60% on each of them, then this means that you've met your goal. (Celebrate!) But, what if the data had looked like this chart instead?

	A	B	C	D
1	Incident	# of students who identified it	Total Students	%
2	<i>Specific time when Bella does something completely irrational/emotional</i>	18	25	72
3	<i>Edward refuses to let Bella do something for herself</i>	8	25	32
4	<i>Bella is seen doing housework or grocery shopping for her father</i>	19	25	76
5	<i>Jacob and Edward fight over what is best for Bella</i>	5	25	20
6				
7				
8	AVERAGE			50

⁶A formula to calculate the percentages is inserted in cells D2, D3, D4, and D5. The formula in D2 reads $=((B2/C2)*100)$ because you are dividing the actual number of responses by the potential maximum number of responses and multiplying the result by 100 to find out the percentage of responses received.

⁷This formula in D8 is $=AVERAGE(D2:D5)$. This averages the percentages from each individual item to give you an overall average for the activity.

⁸For most activity-based assessments, that will be a good measure of success to use.

In this instance, you have less than a 60% average for the participants in this application activity. At this point, you want to ask yourself some questions that will help you understand why the participants did not perform as expected.

Are there particular items on the instrument that received markedly low endorsement, or are all of the items lower than the established threshold for success?

Do the lowest items have anything in common? If so, what? (e.g., Do they point to the same concept or idea? Are they all about behavior? Or attitude?)

If we look at the responses, two items have very low scores. Also, we can see that the two items that received low scores both point to similar phenomena. Each of these questions point to a couple of things:

	A	B	C	D
1	Incident	# of students who identified it	Total Students	%
2	<i>Specific time when Bella does something completely irrational/emotional</i>	18	25	72
3	<i>Edward refuses to let Bella do something for herself</i>	8	25	32
4	<i>Bella is seen doing housework or grocery shopping for her father</i>	19	25	76
5	<i>Jacob and Edward fight over what is best for Bella</i>	5	25	20
7				
8	AVERAGE			50

1. Male behavior: These two questions focus on what Jacob and/or Edward are doing; the other two questions are about what Bella is doing. (Even though these behaviors are influenced by the behavior of men, they are about what she is doing.) As a society, we tend to focus more on women’s behaviors than we do on men’s around issues of both gender socialization and violence against women. Perhaps that’s part of what’s playing out here. Although the young people are beginning to identify gender-role socialization, it might take a slight shift in your learning activity or a little more time for them to focus on male behavior, too.

2. Potentially confusing behaviors: The two behaviors in these examples could be misconstrued as positive behaviors based on how we’ve been trained to think about gender. For example, Edward’s insistence on doing things for Bella could be seen as a kind, chivalrous gesture rather than as controlling. Additionally, men fighting over women is often seen as a desirable circumstance in our culture, even though the factor that is clearly missing from that dynamic is the desire of the woman herself. It might take time for people to reframe behaviors such as these, and it’s important to help them frame them within a larger context.

So, what do you do with this information? In order to figure out why that happened, you could consider reviewing the activity with the students at the beginning of the next session, specifically focusing on the parts of the movie that seemed to be less clear to them. You might want to show those clips and ask the participants what they see happening. Their responses should give you some idea about why they didn't endorse those moments as indicative of gender-role socialization. Jot down some information for yourself in your process notes so that you can think about how to modify your activities the next time you do them. For instance, write some ways that you might facilitate the learning activity differently next time based on what you learn through data collection and/or conversations with the participants. Also, think about what you are covering in the next few sessions. Will you need to find a way to reinforce the information the participants missed in order to set them up to learn and integrate the upcoming material?

For a more robust case study with several scaled scoring tool options, see Appendix C.

Steps 5 & 6: Revise Program Plan or Curriculum

Step 5 requires us to answer the question: What next?

Ideally, your curriculum is designed in such a way that each session builds off of the one before it in some way. If the assessment shows that the participants didn't successfully integrate the learning from one session, what will you need to do with them during the next session to reinforce the pieces they missed and set them up to successfully integrate the next pieces? The previous section gave some ideas for how to further contextualize the low scores from the *Twilight* example. Reviewing the pieces that got low scores and inviting dialogue about them can help reinforce the objectives that were not met and might be enough to prepare the participants for the next piece of learning.

Step 6 asks us to answer the question: What next time?

In addition to looking forward to the next session with this group of students, you also need to look back to the session that produced low scores and think about whether or not you need to modify it for future use. During and after the analysis process, it will be critical to take a look at what the data point to as opportunities for improvement. For example,

Do you need to select a different learning activity to meet the learning objective? Or keep the learning activity and facilitate it differently?

Is the application activity actually assessing the learning activity as you had hoped?

Perhaps the learning objectives are what need to be revised resulting a revision of multiple parts of the curriculum.

Were there particular aspects of this group or group experience that were actually the critical factor that

Tips as you move forward

- **Start slowly.** Even if you're really excited about this methodology, remember that it's okay to start slowly. Implement this practice in one session and get comfortable with it before you use it more frequently.
- **Only collect what you need!** Don't collect data or information just because you can; focus on what is useful.
- **Involve others.** Try to find someone who can do the scoring for or with you - (e.g., a teacher, intern, colleague, or volunteer).
- **Reach out.** Contact people at TAASA or TCFV for support in any of the steps above.
- **Share lessons/ideas.** Share your ideas and experiences with us! We learn through your process as well. Trying something new? We'd love to incorporate your learning into our efforts to help people adopt this method and collect useful data.

prohibited integration of learning? If so, do you need to make notes about how to modify this session if later groups have similar dynamics?

Remember that this is not necessarily a complete overhaul of your program. Activity-based assessment allows you to take a microscopic view of your curriculum and make very subtle adjustments to achieve your objectives and gather useful and meaningful information about your process.

And, finally, though evaluation can be intimidating at times, remember that you are the right person to do this. You don't need expertise or experience in formal evaluation to engage in activity-based assessment. The whole point is to tell a story. While there are more robust approaches, as long as you are systematic and create your rubric ahead of time by thinking through learning activities and application activities, whatever emerges will help improve what you're doing and help you talk about what's going well and why.

Appendix A

Program Evaluation Resources

Centers for Disease Control and Prevention: www.cdc.gov/EVAL/.

Jane Davidson, Actionable Evaluation: realevaluation.com

University of Wisconsin Extension: www.uwex.edu/ces/pdande/

Appendix B

Sample Process Notes

The following is an example of a process note sheet that can be used at the end of every session you facilitate and will be particularly important to use during any session you are assessing with this methodology. Certain sections might not be relevant so, as always, only collect the data you need and will use. Many of these questions will provide important context for you when compared to the data you collect through scoring tools, and there is space to consider the modifications you might need to make to activities or facilitation.

Date _____
Topic _____
Group _____
of attendees _____
Notable absences _____

Indicate with a check mark (☑) if the activity was completed.

Activities Planned	☑ If it was completed
	☐
	☐
	☐
	☐

For each of the following questions, please indicate your level of agreement.

	Strongly Agree	Agree	Unsure	Disagree	Strongly Disagree
The issues presented were relevant to the audience	<input type="checkbox"/>				
The activities were appropriate for the topic	<input type="checkbox"/>				
Skills development was effective	<input type="checkbox"/>				
The participants seemed engaged	<input type="checkbox"/>				

1. Were there any disruptions or interruptions that shortened the length of the lesson?

If yes, provide a brief explanation.

2. Was the time allotted for this session:

Too long Just right Too short

Explain:

3. Were there pieces that you had difficulty facilitating? Please explain the difficulty and how that might be dealt with differently the next time the activity or discussion is facilitated.

4. Was there a specific section or activity that was well received or that seemed to have a particular impact on participants? If so, please identify which section(s) and describe why you think it was well received or had a particular impact on participants.

5. Was there a specific section or activity that was problematic? If so, please identify which sections(s) and why it was problematic.

6. In what ways should this session or these activities be adapted to better meet the need of the participants (e.g., time, subject matter, materials)?

7. Please offer any other feedback or suggestions for improvement to the activities.

Appendix C

Case Study

Learning Objective: Demonstrate the key qualities of effective bystander intervention

Learning Activity: Participants learn the meaning and characteristics of effective bystander intervention and see different examples of effective and ineffective intervention. The methods used for this are lecture and scripted role plays to highlight the following key qualities of effective bystander intervention:

- Nonviolent
- Recognizes humanity of target and perpetrator
- Encourages pro-social behavior

Application Activity: Participants are given the following scenario and asked to role play potential interventions based on what they learned in the learning activity.

Scenario

In the hallway one day, you overhear some male students encouraging another young man to snap the bra of a female student who is at her locker. When he refuses, they start telling him to "be a man" and call him names like "pussy."

What will you do? Role play an intervention using the information from the first part of class.

Scoring Tool Options:

Rubric⁹

The following rubric is a multi-component scale that highlights several aspects that the role play would need to demonstrate in order to qualify for a particular score.

Benefits: Since each role-play only gets one score, this rubric could be an easy one to use if you don't have someone else in there to help with data collection.

Considerations: This type of rubric basically demands that you rank or prioritize certain components of what you're looking to assess. Although the rubric assesses, for multiple components, it assesses for

⁹For more information on this type of rubric development, see the work of Jane Davidson (<http://realevaluation.com/>), especially her book *Actionable Evaluation Basics*.

them collectively. This might make it difficult to pinpoint small adjustments that need to be made. In this example, a score of 2 for a particular group definitely means that there was room for improvement, but you might have to make some notes off to the side to help identify the exact pieces that require improvement. When developing this type of rubric, start with the best and worst examples and then fill in the middle options.

Score	Description
4 (Best)	Intervention in role play demonstrates the intervener doing all of the following: <ul style="list-style-type: none"> • Expressing solidarity with victimized person • Speaking for him/herself, expressing own discomfort • Addressing the problematic behavior • Appealing to good qualities or values known to be true about harasser
3	Intervention must involve <ul style="list-style-type: none"> • Intervener speaking for him/herself • Addressing behavior (as opposed to person) as problematic May involve any of the other characteristics of the “best” example but not all
2	Intervention can include any combination of desired behaviors above and unhelpful behaviors in “worst” category
1 (Worst)	Intervention demonstrates nothing from the Best list and any of the following: <ul style="list-style-type: none"> • Making accusations • Making assumptions about how everyone is feeling/thinking • Addressing the person (as opposed to the behavior) as problematic • Using physical violence/harassment • Escalating the situation

Scale

This scale allows you to not only assess for the presence of desired behaviors but also give a score that signifies the quality of the implementation of each.

Benefits: Knowing how well a behavior was implemented is a richer kind of data than just knowing whether or not the behavior was implemented. This gives you a better sense of how well the learning was integrated for each component that’s measured and therefore would allow you to make more specific adjustments to your facilitation or activities (e.g., if you discover that some aspects were implemented well whereas others were not, you can spend more time following up on only those specific aspects).

Considerations: This tool will require more attention to complete and is therefore one of the tools that really benefits from having a trained helper (i.e., someone other than the facilitator) complete it. This scale also takes more time to develop than does the rubric shown above or the next tool.

	4 (Best)	3	2	1 (Worst)
Did they use a non-violent means to intervene?	Intervention involved calm, verbal interventions	Intervention had no signs of verbal or physical abuse	Intervention involved verbal abuse or brought about a verbal altercation	Intervention involved physical violence or brought about a physical altercation
Did it recognize the humanity of the target?	Intervention was supportive to and uplifting to target	Intervention acknowledged the wrong but passively	Intervention denied agency of target	Intervention was demeaning to target
Did it recognize the humanity of the harasser?	Intervention was compassionate/supportive to harasser without excusing the harassing behavior	Intervention was neutral toward harasser and/or excused the harasser for the behavior	Intervention was judgmental or dismissive of the harasser	Intervention was attacking or demeaning to the harasser
Did it encourage pro-social behavior?	Intervention involved the use of clear communication, speaking for self, and/or appealing to the good qualities of parties involved	Intervention involved no unhelpful behaviors and also at least one or two pro-social behaviors	Intervention involved unhelpful behaviors	Intervention encouraged unhelpful behaviors such as accusations, name-calling, escalation of situation

*Continuum*¹⁰ (this is a form of a scale)

Continua like those that follow are a form of scale that asks you to put items that could be construed as opposites on two ends of a continuum and allows you to place a mark somewhere on that continuum based on what you've observed. So, for the first item on this tool, if the intervention contained verbal harassment, you'd probably put a mark to the right of center.

Benefits: These tools allow you to easily indicate a range of responses that give richer data than just saying if something was non-violent or not. It allows for the possibility of behavior going far away from the desired direction which will allow for corrections if that's happening.

¹⁰You might also see this methodology referred to as a semantic differential scale.

Considerations: This is one of the more subjective tools mentioned in this toolkit. Some of that can be controlled for by laying out in advance some clear indicators of each of the words in the differential. This is true both if you will be doing the rating and if someone else will be doing the rating. That clarity will at least help maintain consistency of scoring across the groups you're assessing. Scoring can also be a little trickier with a tool like this. It can be helpful to add a certain number of hash marks (e.g., 5 or 7) to allow for some consistency.



Appendix D

Worksheet for Creating Your First Scoring Tool

Curriculum Name _____

Session # _____

Learning objective
(circle the verb) _____

Learning activity: What activity will you facilitate to teach to the learning objective?

Application activity: What will participants do to apply the learning from the learning activity?

Measure of success: How will you know that they were successful in their application?

Scoring tool: Which of the following is the most important to know?

- Whether or not something was done (checklist)
- How many students did something (count)
- How many times something was demonstrated (count)
- How well something was demonstrated (scale)
- Some combination of the above

Data collection: Will you have someone available to help complete the scoring tools?

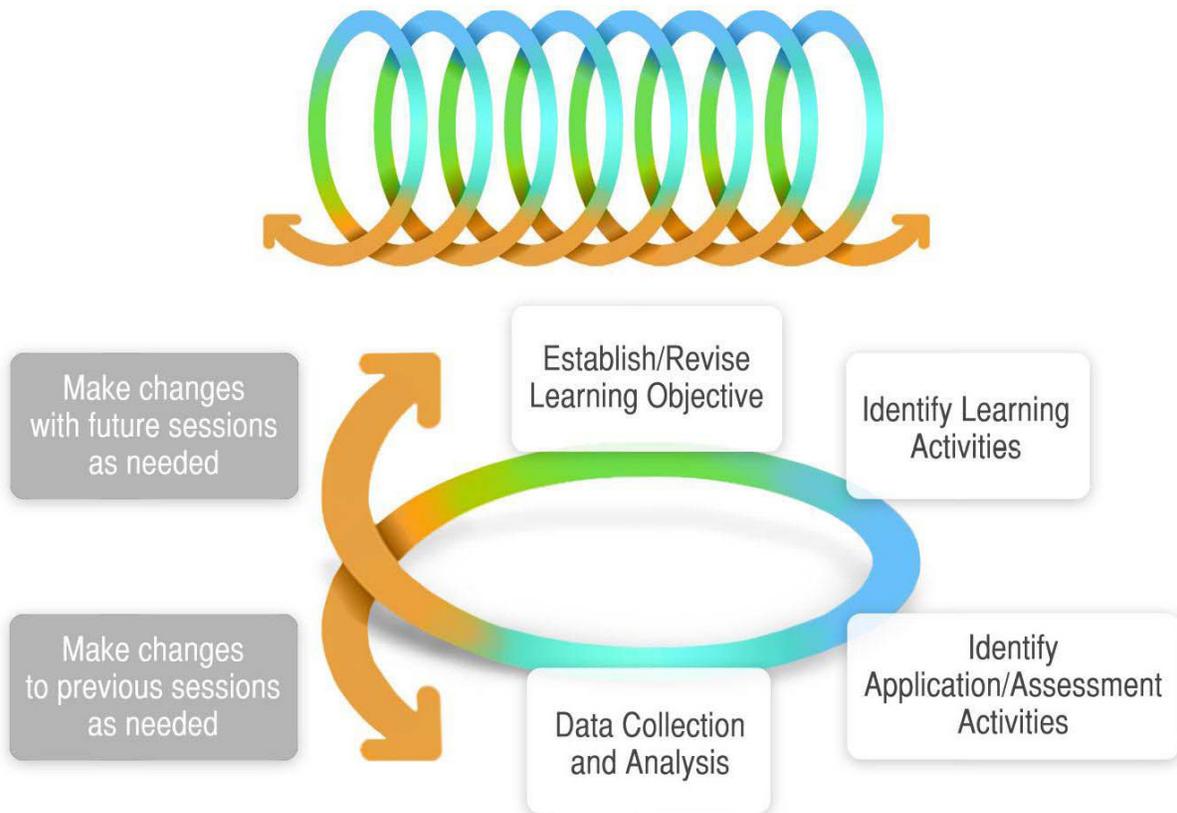
Yes

No

Appendix E

Improvement Loop

The Improvement Loop



developed by Morgan J Curtis, LMSW (Texas Association Against Sexual Assault) and Surabhi Kukke, MPH (Texas Council on Family Violence)

