

Notes from 11/20/2019 Evaluators Workshops

11:00 am session

The discussion focused on the Attendance Roster. Jonathan began by reviewing the main purposes of the roster, which include:

- Helping to determine whether or not programs are meeting their participation goals

- The data is also essential for conducting analysis of program performance—for example, by linking data about attendance hours to State SIRS data such as state tests, attendance, etc., we will be able to explore whether dosage makes a difference. If it does, that provides support to the theory that any improvements in those outcomes might be at least partly attributable to the program.

Jonathan then asked the group to discuss their experiences using the Attendance Roster over the past year and any suggestions for improvement. Following are the main points of discussion:

- Suggestion to add sorting and filtering capability to the roster. This would make it much easier to identify missing information, etc.

- The roster didn't allow for easy editing. Couldn't just delete something within a cell, and also Couldn't insert rows into the roster.

- . Had to recreate the data elsewhere and then copy and paste it into the roster.

- Suggestion to have one local evaluator "test" the next iteration of the Attendance Roster to identify issues/problems and correct them before distributing it broadly.

- Before and After School are still lumped together on the roster, and this is problematic in terms of calculating average daily hours. For example, a student who only attends the program before school will have low average daily hours that will be "red flagged."

- Jonathan indicated that before and after school hours can be separated on the roster to help resolve this issue. Or more spaces could be added for comments to explain outliers.

- Nora shared that it may be useful to combine the Program Information Tab and student data Tab so that people would see the comments box(es) as they were completing it and would be more likely to remember to fill it out; also the target achievement would be displayed more prominently and less likely to be overlooked (I received a side suggestion after the session from an evaluator that we could have a comment column at the end, so that programs would have the ability to enter comments about each individual student. She thought this would be helpful, and they could just copy and paste for all students that were affected by a certain issue so it wouldn't be that much work, although I'm not sure all programs would agree with this! But we could still have it as an option, in addition to a general comments box, and it could be used by those who wanted to add more detail but not be a requirement.)

--Accounting for time spent on Field Trips

--Field trips aren't listed. How should we code them?

--What about field trips that occur on the weekend or on a holiday?

Response: given the currently proposed structure for the Y3 roster, field trips would be included in the appropriate column (weekends, summer, after school, etc.) depending on when they occur.

--However, concerns were expressed that including long field trips skews overall data in categories that would normally otherwise reflect short activities. This is especially true for trips that occur during the school day, and for after-school field trips.

--For this reason, Suggestion to perhaps create a separate category for field trips. And in cases where there is a field trip held during the school day, ask for additional information. Or create a sub question that asks if field trip was held during an official school day. (Since only ELTs can count field trip time that occurs during the school day.)

--However, concerns were expressed that Further disaggregation of data (including separating before/after school, and adding field trips) may not necessarily help improve data quality (reduction of red flags); it could result in just as many problems, because requesting a change in record keeping now that we are well into the program year could itself cause data errors.

--Suggestion to allow rounding up of student attendance hours, so that students who are close to 30 hours are counted. (This seemed reasonable at the time, but upon review I'm not sure that we should round up – if the requirement is 30 hours, then shouldn't a student be counted when they have attended 30 hours, not 29.45 hours? Is this a question for Elizabeth? If we do decide to round, need to clearly define the lowest acceptable number to be included.)

--If changes are going to be made to the roster, they should be implemented on June 30 rather than in the middle of the program year. It's difficult to adjust when changes are made mid-year. Programs and local evaluators need to be able to plan for data collection and have consistency during the program year.

In response, MI expressed appreciation for this concern, but also pointed out that programs are already clearly required by NYSED to document attendance "by hours *by activity*". However, it was acknowledged that even though this means that it's reasonable to expect that programs already have the disaggregated data, it is often on paper, and they haven't necessarily been entering it that way into their databases.

--Accessing student level data

--Local evaluators in NYC are no longer required to get IRB approval in order to access student level data. They just need to sign a confidentiality agreement; however, they do still need to go through the Research and Policy Support Group (RPSG).

--However, one local evaluator of NYC programs said it's still difficult for her to obtain student level data on attendance and behavior, because these types of data need to be obtained from the district, not the school.--Several participants expressed the sentiment that it shouldn't be a local evaluator's responsibility to secure cooperation from districts with regard to data sharing. The state department of education should establish what data is needed at the various levels, how the data will be obtained, etc. They have the leverage to compel cooperation from districts.

MI reminded them that Elizabeth had already distributed an official memorandum stating that all participating programs are required to share data with the State, and that notice of such requirements from the State can and should be used as leverage to obtain the district's cooperation, if necessary. [Nb: In NYC, this should be easier now that programs no longer have to go through the IRB; but this may still need to be resolved in other districts.] However, it isn't realistic for the State to negotiate with each of the 150 or more districts that are participating, with each program using different data. We are not asking programs to provide the standardized data (such as attendance) that will be used for Statewide analyses – that is being provided to MI by NYSED, or once the statewide data system is in place, through the data warehouse.

-- Regarding the stated plans to use state data to look at trends, one evaluator expressed concern that there is “no way to demonstrate that gains are attributable to program.” Jonathan pointed out that [post hoc, quasi-experimental] strategies can be used to bolster evidence about attribution – such as demonstrating whether there are associations between outcomes and dosage.

--Question: Should we include students with less than 30 hours on the roster? Answer: Yes. This enables us to determine the proportion of all participants who are receiving at least a minimum dosage. However, “participants” with 0 hours should *not* be included!

--Suggestion that program directors be included in decision making around what is measured and how, to get their perspective on what will work. It would also be helpful to include program directors in the emails sent to local evaluators, and also have a few program directors on the Advisory Committee.

Jonathan indicated that he would add program directors to evaluators listserv and welcomed participation of program directors on the Advisory Committee.

2:00 session

Jonathan began by reviewing the purposes of the AER template, reminding them that there are specific expectations from NYSED in terms of the data we collect and report. In Year 1 (before the report

template was implemented) many reports were missing a lot of the required information. Having a template helps ensure that all of the required information is included.

Each table was then given time for “table talk” about their experiences using the AER template and asked to note any challenges they encountered or what they liked/appreciated about the template. After table discussion, each table reported out to the full group. Following are the main points of discussion:

--SMART Criteria for Goals and Performance Indicators:

--Some evaluators felt that some of the performance indicators don't map into the SMART criteria

--Sometimes, programs have established objectives that are not realistic or measurable. Are we allowed to change or update an objective? Answer: Yes, you can (and should). You need to contact Elizabeth Whipple to request a program modification. You'll need to provide justification for the request.

--For some PI's, it's unclear if evaluators should be reporting annual or over 5 years. Answer: It depends on how the objective is written. If it's ambiguous, that means the PI doesn't meet the “time-bound” aspect of SMART criteria and might need to be re-written. It's quite reasonable to have 5-year objectives/indicators; if you do, you should still report progress towards meeting them each year. Similarly, even if you didn't meet a particular indicator, it's still useful to report how much progress was made and how close you came.

--Use of appendices:

--Some local evaluators would like more clear guidance on the appropriate use of appendices. Some found themselves using appendices heavily for things like student survey data tables, since they are not embedded in the report anymore.

Jonathan responded that the intention is for the template to be open-ended enough to allow inclusion of charts, etc. It's up to local evaluators to decide where to put data tables, etc. either in the body of the report or in the appendices.

--Others wondered what was enough or too much, in terms of elaboration and evidence included in the appendices. Some expressed concern that the readers/reviewers may be overwhelmed by the amount of information included in the appendices.

Answer: the template intentionally leaves considerable flexibility on this; however, there are some specifications, for example that results for all surveys, observation protocols, etc. need to be reported. The intention of allowing such detail to be included in appendices is so that the narrative section can be made more user-friendly.

--Question: Are we supposed to report at the consortium level or individual site level? Unclear about how much detail to provide. Answer: This will need to be clarified in the Y3 template.

--Most of the AER content seems appropriate to ask of local evaluators. However, the Engagement and Communication section seemed more like self-reporting. It may be more appropriate to ask program directors to provide this information. Response: this section does not only address communication with the evaluator, it is also an assessment of communications with/among all stakeholders. However, evaluators still felt that they wouldn't necessarily have good information about that and that it should be asked of program directors.

--Target participation number should be added. It was suggested to add it to the tables in Section II (Eval Plan & Results)

--Mixed opinions about the usefulness of the AER:

--Some expressed that the AERs are not useful/appropriate for local program stakeholders, and lamented the fact that they need to create a second (more simplified and less jargony) report for those audiences.

--One participant expressed that she doesn't like the idea of having to use a template. Wants more flexibility.

--Can some research be done to find out how other states handle AERs?

--Some of the data being requested in the AER exists elsewhere. Why does it have to be included in the AER? (The speaker did not specify which data they were referring to.)

--How are you using the data from these reports? I want to know that someone is reading them and making use of the information. It would be helpful to know, for example, if a program is providing test prep, whether or not that test prep is resulting in improved test scores.

Jonathan and Lily responded that MI can use the AER data to show trends across the state, to summarize impacts, etc. For example EW had requested that we track statewide trends in meeting objectives. Can roll up individual program data on things like # of students being served, comparison of results for objectives vs. dosage, etc. However, they noted that a State-wide analysis would not be granular enough to show impacts of program-specific activities like test prep.

--Concern that meaningful, qualitative information is not being captured. Some of the programs' most important successes/impacts are not captured by quantitative data. APR is useless for measuring academic outcomes. Response: the shortcomings with the APR is the reason we are collecting detailed local attendance data which will be compared with state records on academics, attendance, etc. Qualitative results may be more appropriate as evidence for program level outcomes, but this will be considered in the exploration of how other states approach AERs.

--Others were more positive. One participant expressed that the AER gives MI and NYSED what they need, but also useful to her programs. Using the AER information, she was able to customize a shorter report for each of the programs that she's evaluating, per their requests, tailored to their specific wants and needs. She also noted the efficiencies that this reporting structure will allow in subsequent years, now that she has retrieved all the baseline data. It will be simple and easy to just plug in the latest results.

Another participant indicated that she feels the AER is more effective for local programs than a typical evaluation report. Since programs know that the report is being shared with the state, they pay closer attention to the content and what it's telling them. The programs are using it for continuous improvement purposes. She shared that she requires the programs to read the AER and sign off on it, then follows up by sharing a data placemat which contains lots of graphics.

--Formatting of the AER (using template) was challenging. But now that everything is set up, updating the reports this year will be fast and easy.

--A program administrator in attendance felt that the AER was extremely useful for her, in terms of fundraising, grant writing, and marketing. She appreciates that it is comprehensive, allowing her to access whatever information she might need. It helps her illustrate/communicate the value of the program to funders.

At my table there was a team of evaluators who wrote down their comments and shared them with me. Regarding the AER Template:

"As a result of the AER Template we changed what/how was reported. Program Director found it easier to understand and forced them to hone in on areas needing improvement/ indicators not met. We created a one page document based on objectives and which sites met each of the indicators. We would be happy to share this created document."

State Expectations Outline

JT walked through/introduced the Outline components.

One evaluator informed the group about components of SMV Tool. Suggested it may be worthwhile making changes to the student satisfaction surveys (because SMV requires programs to collect this info) – We have a good example that was shared by one of the local evaluators. Also noted the requirement for a parent survey/needs assessment (LC notes that this is NOT necessarily a requirement of the local evaluator to help with, it is explicitly a requirement for grantees to conduct an annual needs assessment/ survey of parents.)

JT acknowledged that the Outline does not include anything about parent surveys at this time. But he'll take a look at that.

Some discussion around local evaluators role in continuous improvement efforts and the QSA specifically.

JT proceeded to clarify more of the expectations in the Outline doc, citing the origin docs. Explains how it reinforces the purpose and function of the new AER Template to help gather outcome data and report it.

One evaluator said there were other crossover areas between services local evaluators may be providing and the requirements specified in the SMV. For example SMV requires “Action Plan” and documentation of communications with evaluators. So, just be prepared to collect and provide this to clients, re-name formative/ recommendation reports as “action Plans” if they seem to fit that purpose.

Another evaluator commented on the structure of the document. Good as a reference for programs and evaluators, but seems like a justification because it emphasizes the policy language/ jargon. Too technical, perhaps.

LC commented that this was a first draft to show the “bones” of the expectations and the originals. But there is intention from NYSED to refine and create user-friendly version.

An evaluator at LC’s table reiterated a suggestion raised at last spring’s meeting – someone needs to create a Program Guide to clarify expectations in a concise way for grantees. Especially with such high turnover.

One evaluator suggested that a document should be created for grantees clarifying the expectations for the Advisory Board. (It should be noted that NYSED circulated a memorandum through the EMSC with guidance for the Advisory Board on 6/20/19.)