**When a Triangle Becomes a Three-Sided Square:**
**Measuring Students' Science, Technology, Engineering and Math (STEM) Interest and Learning through Data Triangulation**

Ginger Fitzhugh
gfitzhugh@eraeval.org
Evaluation & Research Associates

Paper presented at the American Evaluation Association Conference
Orlando, FL
November 13, 2009

## Abstract

Although triangulation or the use of multiple methods is widely considered to be good practice in evaluation, there is relatively little guidance in the evaluation literature about how to design a multi-method evaluation to triangulate data appropriately, or how to interpret incongruent findings. Measuring whether and how students' interest in science, technology, engineering and math (STEM) changes in response to program interventions can be particularly challenging, and triangulation is one approach to try improve the confidence and consistency of findings. In this paper, we describe how multiple methods were used in an evaluation of TechREACH, an out-of-school-time STEM program serving low-income and underrepresented middle school students in Washington State. We present examples of data about students' STEM interests that were congruent (i.e., the data from all sources generally agreed in direction and magnitude), and examples of data that were incongruent (i.e., one or more sources did not agree with the others), and discuss how to interpret the lack of congruence. We recommend steps to avoid problems with incongruity prior to beginning data collection and methods to manage incongruity following data collection.

## Introduction

Triangulation is the collection and synthesis of multiple sources of data to investigate the same phenomenon. The term is borrowed from a navigation technique in which multiple reference points are used to locate an object's position. The purpose of triangulation is typically to increase the validity and credibility of findings by comparing results across multiple sources. If the results are consistent across all sources, the evaluator or researcher may be more confident in formulating a particular conclusion. Using multiple independent assessments of the same phenomenon can offset bias and measurement error (Campbell and Fiske, 1959), and strengthen the validity of findings through their congruence and/or complementarity (Greene & McClintock, 1985). Alternatively, triangulation can provide a richer picture of the phenomenon being studied, and can uncover variance that would otherwise be ignored if only single methods were used (Jick, 1979).

Although triangulation or the use of multiple methods is widely considered to be good practice in evaluation, there is relatively little guidance in the evaluation literature about how to design a multi-method evaluation to triangulate data appropriately, or how to interpret incongruent findings. Measuring whether and how students' interest in science, technology, engineering and math (STEM) changes in response to program interventions can be particularly challenging, and triangulation is one approach to try improve the confidence and consistency of findings. In this paper, we describe how multiple methods were used in an evaluation of TechREACH, an out-of-school-time STEM program serving low-income and underrepresented middle school students in Washington State. We present examples of data about students' STEM interests that were congruent (i.e., the data from all sources generally agreed in direction and magnitude), and examples of data that were incongruent (i.e., one or more sources did not agree with the others), and discuss how to interpret the lack of congruence. We recommend steps to avoid problems with incongruity prior to beginning data collection and methods to manage incongruity following data collection.

## TechREACH Program and Evaluation Description

TechREACH is funded in part by the National Science Foundation's (NSF) Innovative Experiences for Students and Teachers (ITEST) program, which seeks to engage students from diverse backgrounds in STEM education opportunities with the goal of increasing the number of youth who pursue STEM careers. TechREACH program staff train middle school teachers to lead after school clubs and summer workshops providing hands-on technology-integrated curriculum. Students also go on field trips to community sites that are related to the TechREACH curriculum (e.g., a wind farm), and STEM professionals visit the clubs as guest speakers to discuss their job responsibilities and career pathways. The goals of TechREACH are to: (1) increase middle school underrepresented students' interest, confidence, and achievement in STEM; (2) build the capacity of schools and communities to support students' STEM learning and interest; and (3) increase the number of underrepresented middle school students who pursue STEM classes, degrees, and careers. In 2008-09, TechREACH served more than 200 students in 14 clubs from six middle schools located in Eastern Washington.

From fall 2008 through summer 2009, Evaluation & Research Associates conducted the evaluation of TechREACH's third year of activities in Eastern Washington. The previous two years of the evaluation were conducted by another research firm. As much as possible, the original evaluation design and instruments were preserved to facilitate comparison of findings across all three years of implementation and to minimize disruption to TechREACH staff and club leaders. Multiple methods were used. The principal sources of data were student, club leader, teacher, parent, and guest speaker surveys. In addition, student focus groups and interviews of club leaders, program staff, and school administrators were also conducted. Students' grades, club attendance records, and subsequent involvement in STEM activities in high school were collected. Site visits to after school clubs, club leader trainings, and special program events were also documented. The evaluation examined students' interest in STEM, achievement in STEM coursework, proficiency in computer technology, knowledge of and interest in STEM careers, and likelihood to pursue future STEM courses.

This paper focuses on selected results from survey measures regarding students' interest in technology and science to illustrate how data were triangulated in the evaluation. Students in the TechREACH program completed pre- and post-surveys which asked a set of identical questions regarding their interest and confidence in STEM. The student post-survey included an additional set of questions which were only asked at the end of the year. Club leaders, parents, and students' math and science teachers also completed surveys about students' interest in STEM at the end of the program year.

**Congruent Data Example**

Table 1 shows an example in which the data from different measures were congruent, in this case regarding students' interest in technology. Data from students, club leaders and parents all suggest that TechREACH had a positive impact on student interest in technology. Between 79% and 83% of students, club leaders and parents agreed that TechREACH had increased students' interest in technology and/or computers. These findings are also consistent with qualitative interview results from club leader interviews and student focus groups, which suggested that although many students already had a high interest in technology when they entered the program (as also evidenced by the fact that 91% of students indicated that they like using computers on the pre-survey), TechREACH successfully introduced them to new technologies, bolstered their confidence in using technology, and deepened their interest in using technology in the future.

**Table 1. Student Interest in Technology**

| *Data Source* | **Student Pre/Post Survey** | **Student Post Survey** | **Club Leader Survey** | **Parent Survey** |
|---|---|---|---|---|
| *Survey Item* | "I like using computers and other technology for my school work." | "TechREACH has increased my interest in technology." | "Because of TechREACH, most of the students in my club are more interested in computers and/or technology." | "Because of TechREACH, my child is more interested in computers and/or technology." |
| *Results* | 91% pre<br>93% post<br>agree or strongly agree | 83%<br>agree or strongly agree | 79%<br>agree or strongly agree | 91%<br>agree or strongly agree |
| *Conclusion* | Students, club leaders and parents all report positive program impact on student interest in technology | | | |

**Incongruent Data Example**

Table 2 shows an example in which the data were not congruent, in this case regarding students' interest in science.

**Table 2. Student Interest in Science**

| *Data Source* | **Student Pre/Post Survey** | **Student Post Survey** | **Teacher Survey** | **Parent Survey** | **Club Leader Survey** |
|---|---|---|---|---|---|
| *Survey Item* | "I like studying science." | "TechREACH has increased my interest in science." | "This student has exhibited an increased interested in the subject I teach [science]." | "Because of TechREACH, my child is more interested in math and/or science." | "Because of TechREACH, most of the students in my club are more interested in math and/or science." |
| *Results* | 69% pre<br>65% on post<br>agree or strongly agree | 66%<br>agree or strongly agree | 55%<br>agree or strongly agree | 73%<br>agree or strongly agree | 93%<br>agree or strongly agree |
| *Conclusion* | Student pre/post survey results show a small decline in students' science interest, while post-only survey results from students, teachers, parents, and club leaders suggest an increase in interest (but to different degrees) | | | | |

While student pre/post survey results show a small *decline* in students' general science interest (which was not statistically significant), post-only survey results from students, teachers, parents, and club

leaders suggest an *increase* in students' science interest. Approximately two thirds of students said that they like studying science on both the pre and post survey (i.e., there was essentially no change following participation in TechREACH), while at the same time two thirds of the students also said that TechREACH had *increased* their interest in science. Just over half of students' science teachers' also said that students' interest in science had increased, about three out of four parents said that their child's interest in science *and/or math* (italics added for emphasis) had increased because of TechREACH, and nine out of ten club leaders said that *most of the students* in their club had become more interested in science *and/or math*. What's going on here?

There may be a few reasons that the results are incongruent. First, it is likely that the results diverge at least in part due to differences in the format of the items. The survey questions do not ask about exactly the same phenomenon. While students, parents and club leaders were asked about TechREACH's impact on students' interest in science in general, teachers were asked about changes in students' interest within the classroom context and without any reference to TechREACH's impact on their interest. Data from club leader interviews and student focus groups (additional sources used to triangulate the findings) suggest that students greatly valued the hands-on science activities they did in TechREACH, and that these often differed from their experiences in science classes during the school day. The qualitative data might suggest that students increased their interest in science generally, although not necessarily in a classroom setting. As an aside, the fact that two thirds of the students responded they became more interested in science, but did not report more interest in *studying* science is a potential cause for concern because one of the program's goals is to increase the likelihood that students consider pursuing STEM careers, which would require them to take science coursework.

Another possible reason for the lack of congruence is that the data from each measure do not necessarily refer to the same students. The student pre/post results and post-only results are from the same group of students, while the science teacher, parent and club leader surveys refer to a slightly different student pool (including students who did complete both a pre- and post-survey). It was not possible to report results across all measures for the same group of students since the student and parent surveys were anonymous (i.e., did not name the student participant), and the club leader surveys were not individually completed for each student.

**How to Interpret Lack of Congruence**

As described in the above example, there may be several reasons for the lack of agreement across measures. These are summarized below along with some additional possible reasons for lack of data congruity:

- **Measuring different concepts.** The measures might not be consistent in scope or content.
- **Measuring different samples within the population.** If each data source is from or about a different sample of the population and these samples differ from each other on the measure of interest, results may be incongruent.

- **Measuring the same concept but at different times.** If the data from the independent measures are collected at different times and the measures are sensitive to immediacy effects, proximal events that occur near the time of data collection could influence the results differentially. For example, if immediately before completing a post-survey, students learn what science courses they need to take to pursue certain STEM careers, they might respond differently to questions regarding their intentions to enroll in science courses in high school than if they had taken the survey prior to the coursework presentation. Furthermore, parents who complete a survey with similar questions a few weeks or months before this presentation and student post-survey administration might respond differently than their children.
- **Subjective difference of opinion.** Constructivists might argue that there is no objective reality of social constructs and that differences in results may simply reflect different viewpoints. Students, parents and their teachers each view students' attitudes and intentions from a different perspective.

**Lessons Learned**

A number of lessons emerged from our TechREACH evaluation experience as potentially useful strategies to employ prior to data collection and following data collection to increase the likelihood that the data triangulate.

<u>Before Data Collection</u>

- **Clearly define concepts.** Determine what concepts you wish to measure and ensure that they are similarly defined across methods.
- **Consider your "congruence threshold."** Determine a priori what results will be considered sufficient agreement across measures. Also consider whether some data sources receive more weight than others. For example, self-report data regarding attitudes may receive more consideration than data from third parties, who may be less likely to know and report another person's attitudes accurately. Likewise, data from pre/post surveys with a comparison group might be given greater weight than retrospective post-only measures.
- **Weigh time/budget considerations to determine the appropriate number of data points.** Consider the costs and benefits to determine the number of data points needed. Is a triangle sufficient? Or do you need a dodecahedron? The more data points that you collect, the more confident that you can be about your conclusions, assuming the data collection methods are valid and reliable. However, many evaluations are limited in resources and time. It may make sense to concentrate on triangulating concepts that are the most important to stakeholders, or that are controversial, particularly uncertain, or that might be the most variable.

<u>After Data Collection</u>

- **Collect additional data to help interpret incongruent findings.** If time and budget permits, it may be helpful to collect additional data to help interpret incongruent findings and explain the

reasons for their lack of agreement. For example, you might wish to conduct additional focus groups or interviews with participants.

- **Ask stakeholders to help interpret findings.** Similarly, you might present preliminary findings to your client or a group of stakeholders and ask for their assistance in making sense of the findings. Are they surprised by the differences? How might they explain them?
- **Refine measures.** If it is a longitudinal study, make changes to the measures as needed to capture the concepts you wish to evaluate.
- **Congruity does not mean that the findings are unquestionable (Bryman, n.d.).** Finally, a cautionary note. Because findings appear to be congruent does not necessarily mean that they are. It is possible the measures are all biased in the same direction.

**References**


Bryman, A. (no date). Types of triangulation. Retrieved October 12, 2009. http://www.referenceworld.com/sage/socialscience/triangulation.pdf

Campell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin 56,* 81-106.

Greene, J. & McClintock, C. (1985). Triangulation in evaluation: Design and analysis issues. *Evaluation Review 9* (5), 523-545.

Jick, T.D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly 24*, 602-611.