THREATS TO VALIDITY

Performance Standard 13A.H

Students will apply scientific habits of mind to investigations in the sciences, accordingly:

- *Knowledge*: Define the basis of scientific habits of mind necessary in scientific investigations.
- *Application*: Identify basic strategies to evaluate evidence, question sources of information and analyze validity of scientific investigations.
- *Communication*: Correlate how scientific habits of mind validate science and invalidate pseudo-science examples of various media (literature, advertisements, movies, etc.)

Procedures

1. ***In order to know and apply accepted practices of science (13A),*** students should experience sufficient learning opportunities to develop the following:
   - Apply scientific habits of mind from questioned media examples, especially in curricular contexts,
   - Evaluate evidence presented as scientific,
   - Make inferences and deductions from data and observations,
   - Analyze choice of variables selected in investigations and how they are varied,
   - Review experimental procedures or explanations for possible faulty reasoning or unproven statements,
   - Distinguish relationships of scientific theories, models, hypotheses, experiments and methodologies, and
   - Distinguish fact from opinion and science from pseudoscience.

   Note to teacher: This activity is based on knowledge associated with standard 13A, and should be considered for inclusion in all standards-based investigations as a scientific habits of mind.
2. Have students review and discuss the assessment task and how the rubric will be used to evaluate their work.
3. This activity can be introduced by listing examples of science (whether factual or nearly factual or far-from-factual) found in current movies, books, advertisements, etc. Generate questions about how the public can distinguish facts, opinions, real science and pseudoscience. Role-play the "Quick Guide to Validity" script. Ask students to explore the www.badastronomy.com website for recent examples of movies or television or printed media science references. Allow students to generate questions about the reference and express ideas about concepts that may affect the validity of the reference. Students should select specific examples and provide evidence which debunks the pseudoscience, evaluates evidence, provides inferences or deductions, analyzes the choice of variables, examines procedures or explanations for faulty reasoning or unproven statements, etc. Students should present their examples and evidence in an "Expose the Real Science for Inquiring Minds" news show as feature reporters.
4. Evaluate each student's work using the Science Rubric as follows and add the scores to determine the performance level:

   - *Knowledge*:. Scientific habits of mind are defined using clear examples and appropriate terminology.
   - *Application*: Strategies for testing scientific evidence, information and investigations are clearly stated and practiced with multiple examples.
   - *Communication*: Scientific habits of mind are correlated to media examples appropriately and correctly.

Examples of Student Work not available

Time Requirements

- One-three class periods for introduction and regular reinforcement.

Resources

- Copies of the "Threats to Validity" script and notes
- Access to internet for introduction to www.badastronomy.com for initial information
- Science Rubric

QUICK GUIDE TO VALIDITY

Here are some of the basics to get you up to speed with validity. Validity can be thought of as a set of standards by which you judge things. It is the closest thing to the truth for a conclusion or an inference. Measures, samples, and designs DON'T have validity. They may lead to valid conclusions and inferences, but in and of themselves they are not valid. They are more likely to lead to valid conclusions and inferences if they are free from bias.

FOUR TYPES OF VALIDITY

There are four types of validity and they build on each other.

- *Is there a relationship between two variables in a given study?* A causal relationship means that one variable causes the other (e.g., applying high heat to a metal caused it to melt). A correlation means that two variables are changing, but one did not cause the other to change – a third variable may have caused both to change (e.g., many mushrooms appear in a lawn, the grass is growing faster, but the mushrooms did not cause the grass to grow faster – there has been increased rainfall).
- *Assuming there is a relationship between variables, is the relationship a causal one?* For years people believed that stress and spicy foods were the cause of ulcers. Recent research has found that while stress and spicy foods may aggravate ulcers, it is bacteria that cause most ulcers.
- *Did the study adequately measure the construct or idea it was supposed to measure?* Which do you think is a more adequate measure of student grades: Students telling you their grades or recording information from their report cards? Sometimes researchers use “proxies” as a measure for something else. Do you think, for example, that using persons’ annual earnings is a good proxy for their intelligence?
- *Can you generalize the findings to other situations and people?* If it is not generalizable, then it may lead to invalid conclusions about other situations and people.

THREATS TO VALIDITY

Threats to validity are the possible reasons that the inferences and conclusions made from research may be wrong. Threats to validity raise “what about” and “what if” questions that people often ask about a research study. They can also ask "how come you didn't think of this" questions. The following story illustrates threats to validity.

An enthusiastic student (we'll call her Susie) tells her class about her ideas for a research study.

"I am planning on doing a study which looks at increasing the level of student involvement in school activities, particularly extracurricular activities. I figure I will talk to a couple of classrooms to recruit students to participate in extracurricular activities. The topics of the recruitment program could include things like: playing sports makes you a better team member, science clubs improves your thinking skills, and foreign language clubs makes you more aware of other cultures. I could measure their involvement in extracurricular activities both before and after the recruitment program and see if the talks made a difference."

The students in the class glanced at each other with furrowed brows. The teacher was the first to comment: "What about validity, Susie? How do you plan on addressing validity issues in your study?"

"Validity? What do you mean exactly?" said Susie, slumping down like a deflated balloon.

"Well," said the teacher, "Maybe we should start at the beginning. Validity is the best available approximation to the truth of a given inference or conclusion. Let's start with "construct validity." Construct validity refers to generalizing from your research to the concept of you are trying to measure. In other words, did you measure what you thought you measured? There are many issues and topics to consider when addressing the concept of validity. It is a complicated topic and can be overwhelming for someone who was first starting out.

"Threats refer to questions and issues that may be raised by critics (both of the friendly and non-friendly variety) of your research. There are several threats we will address today. The first one is; did you do a good job of defining what you are going to measure? In your case, you will need to think through concepts such as 'student involvement'. What do you mean by student involvement? Do others share your view? In order to address this threat, it often helps to get expert opinions."

Another threat to validity is selection bias. This means that maybe the classes that volunteered to have you present the program already were interested in becoming involved in extracurricular activities. You can never be completely sure of your results unless you have a sufficient number of students who are randomly assigned to your recruitment program and to a control group that doesn't participate in your program. This adds validity to a conclusion that your recruitment program increased participation in extracurricular activities.

Another threat to validity is using only one method to get results. In your case this would refer to having only one version of a recruitment program; critics may argue that any conclusions drawn may refer to your particular version of motivating students. In order to avoid this threat, you could use multiple versions of your recruitment program.

On the other hand, your critics may ask you how you can be sure you are measuring student involvement, if you are only using one measure of it. In order to avoid this threat, use several methods!"

"That doesn't sound too bad," said Susie, "but I have a feeling that's not the whole story."

"That's not even the half of it!" exclaimed the teacher. "Oh no indeed, next there are the *interaction threats*. The first of these refers to interactions between different student activities. You need to be sure that the results you are obtaining are a result of your recruitment program and not some combination of activities in which the students are involved. For example, maybe the students in your study are also participating in activities at church or through a neighborhood organization. These activities may encourage more involvement in extracurricular activities as well. It may be that the reinforcement of both programs working in conjunction with each other is what prompts them to increase their involvement in extracurricular activities.

Another type of interaction that may occur is an interaction of the testing and the recruitment program. Simply, by giving the students a questionnaire beforehand about participating in extracurricular activities may increase their involvement. Because you have made them aware of the time they spend in those activities may cause students to reflect and increase involvement levels. When you label your recruitment program "the treatment" (i.e., what you are doing to cause an increase in student involvement) you are leaving out the questionnaire, which also may be influencing the student involvement."

Finally, the next threat is limitations of the recruitment program. Imagine that you carried out your project, and the data analysis revealed that the recruitment program really did nothing to increase involvement. You'd be bummed and your first reaction would probably be to chalk it up to another research study that proved nothing. However, it may not be that recruitment program is useless; it may be that you didn't conduct ENOUGH sessions of the recruitment program to see the desired effect. Therefore, it is not appropriate to label the recruitment program as a "waste". Get it?"

"I got it, or at least I think I am starting to get it. More importantly, did I hear you say FINALLY and did that mean we are close to the end of this for today?" Susie queried.

"We're close. The last three threats we will cover are known as 'social threats'. These threats all represent the joy of doing research with human beings. The first of these is HYPOTHESIS GUESSING."

"Let me guess," joked Susie. "It is when the people in your study guess what you are looking at and their actions reflect their guess. In my case, that would mean that students guess I am trying to measure involvement in extracurricular activities and they purposefully get more involved, not because of my recruitment program, but because of their inference."

"Excellent!" cried the teacher. "Why don't I finish up with the last two threats?" Without waiting for an answer, she proceeded.

"EVALUATION APPREHENSION is next. Many people get really anxious about being in a study -- they are afraid they won't look good or smart, or in this case that they won't appear to be good students. In their desire to look like the model participant, their behavior and actions may not reflect reality. Again, this is a labeling problem because your recruitment program could make the students think they won't look good if they don't participate in extracurricular activities.

FINALLY, for real this time, one must address the RESEARCHER'S EXPECTANCIES threat. Without knowing it, you may bias your study. For example, you may become really enthusiastic when you discuss student involvement. This may send a message to students that you think involvement is a "good thing" and may prompt them to act accordingly. Again, this means you will label the involvement as an effect of your recruitment program, when in truth it is your overwhelming enthusiasm. SO...what do you think Susie?"

"I think I am overwhelmed."

"Overwhelmed is understandable, but don't panic!" cautioned the teacher.

"Understanding validity is a long process and there are many sources out there to help you get a grip on things. As a matter of fact, Professor William Trochim has quite a bit of information on his web site that will be useful (see Source below).

Source: Adapted from Trochim, William M. (2000) *The Research Methods Knowledge Base, 2nd Edition*. Cincinnati, OH: Atomic Dog Publishing. Internet URL: http://trochim.human.cornell.edu/kb/index.htm (version current as of August 02, 2000).