Judging the Quality of Science

SETAC**TEACH***

* Topics for Exploration, Activities and Classroom Help

Purpose

SETAC**TEACH** articles are resources for teachers, students and others interested in environmental science. Articles are contributed by members of the Society of Environmental Toxicology and Chemistry (SETAC), a professional society with a worldwide membership from colleges, universities, government, businesses, and other organizations.

Introduction

Science is used to make many decisions in our world. We need to be able to evaluate scientific claims and findings so we can make good choices about actions and decisions facing our communities, states, nations, and world. All of us — leaders, members of the media, and the general public — need to be able to separate scientific facts and sound scientific claims from opinions and poorly supported claims.

This issue of SETAC**TEACH** describes what sound science means, and offers some suggestions for judging technical studies and what they mean. Additional resources are also provided.

A complementary document provides teachers with guidance regarding which United States National Science Education Standards may be met through use of this material, and with suggested classroom activities.

Why do we care about evaluating science?

We make decisions every day about the products we use, the foods we eat, and the medicines we take:

- Will this bottle of herbal medicine that's labeled "all natural" help my skin?
- Is this cheeseburger bad for me?
- Will loud music really damage my hearing?

Our governments use scientific findings, along with other information, to make policies on environmental, health, national security, and other issues. People with opposing viewpoints try to convince us that science is on their side. We need to be able to judge their competing claims:

- Some people think that nuclear power is a good way to provide clean abundant energy for the world's needs, while others are convinced that safety and waste disposal problems make this energy source too dangerous to use.
- Government rules say how much waste industry can put into streams, rivers, and the ocean. Industries may say the regulations are too strict, and environmental groups may claim our health is at risk.
- Genetically-modified foods are seen by some as a tool to improve the world food supply and reduce hunger, and by others as an environmental and health risk.

Even when scientific findings are not at issue, people will make statements that seem scientific but aren't. People may try to make their opinions seem more believable by adding scientific information to their argument when it isn't relevant (doesn't have anything to do with their argument).

Maybe the most important reason to care about quality science is that the more good science we learn, the more we can appreciate the wonder of the natural world about us.

What is not science?

- Generally accepted "facts" or information, things that "everybody knows" or that "they say" are true.
- Beliefs, religious or otherwise. Beliefs may be true, but they are not science.

When should scientific criteria be used to judge a claim?

- When it's a scientific claim
- When it's presented as a scientific claim. Something presented as science should be judged by scientific criteria.

What is sound science?

Sound science can be described as organized investigations and observations conducted by qualified personnel using documented methods and leading to verifiable results and conclusions. "Sound," used in this sense, means that the science is based on valid reasoning and procedures.

Sound scientific studies follow the high standards of the scientific method, resulting in organized investigations and observations. These standards include:

- A testable hypothesis (working assumption)
- Correct methods (for example: proper controls, a big enough sample)
- Enough clearly presented information to show what was done and how
- Use of the right data analysis tools (for example, using the right statistical methods)
- Conclusions that address (relate to) the hypothesis and are supported by the results

Qualified personnel conduct sound scientific studies. Qualified personnel may be formally trained in their subject, for example they may have advanced degrees such as a Ph.D., or they may have had enough experience to get the skills they need. Qualified people know the subject and will be able to:

- Design studies or experiments so that other hypotheses can be ruled out
- Describe and analyze material correctly
- Communicate results clearly and accurately

Documented methods mean that:

- Methods are described well enough that another researcher could:
  - repeat the study with the information given, and
  - find problems in the study (for example, problems with materials used)
- All needed data and calculations are given, so that:
  - others can check for errors,
  - data can be validated (confirmed),
  - the study can be well-reviewed by other scientists (peer-reviewed), and
  - the researchers can get credit for what they've done

Well-conducted scientific studies lead to verifiable results and conclusions:

- Conclusions are directly supported by the experiment
- Conclusions can be reached independently by other researchers in other places, using the same or other methods
- Conclusions should not go beyond the data or what the study was designed to explore. For example, if observations showed that there were fewer birds in a city park than in the previous year, you would not conclude that there were fewer birds in the whole city than in the previous year. Speculation (thoughts or guesses about the topic) is okay as long as it's identified as speculation.
- The absence of evidence is not the same as evidence of absence. Suppose there was a chemical spill in a lake, and you expected to find that the fish in the lake are harmed. Your study shows there are not fewer fish in the lake. You see no evidence that the fish are harmed. However, that does not prove they were not harmed; there could be effects that you have not measured. Maybe they will lay fewer eggs the next year.
- It's impossible to prove something doesn't exist. For example, prove that all crows are black. To do that, you have to show that other colors of crows don't exist.

How should scientific findings be reported?

Science should be reported clearly, completely, and accurately, with documented methods. The amount of detail in a report may depend on what kind it is. A technical paper will have enough detail to meet the criteria listed under documented methods above. A newspaper or magazine article would not be expected to contain a detailed method description or data tables. However, the facts, and the conclusions of the people

who did the study, should not be changed. Also, the writer should give enough information so the reader can look up the original technical report.

Things to watch out for

- Words like "always," "never," and "only" are suspect. Results are rarely absolute. For example:
  - "You kids only care about yourselves."
  - "It's always hot in July."
    A more informative statement might be, "In the last 5 years in our town, high temperatures in July ranged from 20 to 32°C, and low temperatures from 10 to 18°C."

- Words used imprecisely may cloud meaning.

  For example, the word "natural" has a specific meaning (present in or produced by nature; not artificial or man-made). However, it is often used to describe a product to suggest that the product is better for you than if it was not natural. This might or might not be true. Consider that poisons can be natural, and medicines can be man-made.

- Important factors that could affect the conclusions are overlooked.

  For example, the statement "hungry dogs are more dangerous" overlooks differences among types of dogs and how they were raised.

- The sample size is too small or not reported, or the sample is biased.

  For example, how meaningful would it be to say that eighth grade students don't know enough material to pass benchmark biology tests if only students from one class in one school were tested? What if the conclusion was reported without mentioning how many students were tested?

- Results without a reference to what is normal or expected can be misleading.

  Suppose someone says, "The behavior of kids in this school has gotten worse. Why, there was a 100% increase in the number of students receiving detention this month!" But if the policy in that school is to emphasize detentions as punishment, an increase in detentions may say little about behavior. And if that school has 2000 students, and there were 2 detentions this month and 1 last month, the "100% increase" is not as important as it sounds.

- Correlation is not the same as cause-and-effect. It is incorrect to say that one thing causes another, when the results only show that the two things occur together.

You might show that people wearing hats are likely to be wearing gloves, but does wearing a hat cause someone to wear gloves, or is it more likely that cold weather causes them to do both?

- Observer bias, what people think before they start a study, can bias the results or change what conclusions they draw about the results.

  Suppose a researcher believes that herbal medicines are the best way to cure diseases. That person might be more likely to conclude that people in a study of an herbal medicine got better because of the herb, and might not consider other possible factors. Maybe those who got better ate a more nutritious diet than the others.

- Do the observers gain money, power, or popularity because of the conclusions?

  Could a manufacturer-funded study about the safety and usefulness of a product be biased?

- Solid conclusions cannot be based on personal stories. Isolated examples are called "anecdotal evidence" in science. Such evidence is interesting and may inspire more controlled studies, but is not to be relied upon.

  For example, "It won't hurt me to smoke — my grandmother smoked all her life and didn't get cancer."

- Study results not published and subject to review can't easily be evaluated.

- References not provided can't be checked.

  If an article says, "previous studies have shown" something, but provides no reference so you can look it up, you cannot verify the statement.

- Quotations without context can be misleading.

  For example, "A writer for *Science* magazine stated that there is 'no evidence that this product causes cancer.'" What is the whole quote? Could she have said, "There is no evidence that this product causes cancer in rats, but studies on primates did show high cancer rates"? What if the product did not cause cancer in her study, but the rest of the original article discussed limitations of her study?

- Scientific information might be inserted into a nonscientific argument to make it seem more believable.

Rough-skinned newts, *Taricha granulosa*, are poisonous. They have no predators except the occasional garter snake. When the ranges of certain garter snakes and the newts overlap, the snakes develop some immunity to their poison. Space aliens who want to rid our world of newts genetically engineer the snakes. (In case you couldn't guess, that last sentence is made up.)

Frequently asked questions

- Isn't a lot of science just unproven theories?

  The word "theory" is used differently in science than in popular culture. In everyday life, people often use the word "theory" to mean a guess, a hunch, or a personal opinion. For example, someone might say, "My theory is that the new teacher is going to be late the first day." In science, a theory is a framework that explains existing observations and can be used to make predictions. An example is the scientific theory of natural selection to explain the evolution of life on earth.

- New discoveries change scientific theories all the time. How can we ever know what's true?

  A theory may explain all observations when it is first developed and tested, but new testing methods or new discoveries may show observations that it does not explain. Then the scientists working in that field must develop a new theory that does explain all the observations, both the old and new. The idea that we can't always come up with a firm answer to scientific questions may make some people uncomfortable, but learning new things is part of the excitement and wonder of science.

- Don't scientists sometimes make up data? How can we trust scientific findings?

  Scientists are only human, and sometimes humans are wrong, are fooled, or cheat. However, part of the nature of science is for scientists to constantly try to verify, repeat, and test others' findings. You do this when you use the principles in this paper to evaluate scientific findings. Because scientific findings are subject to review and testing by others, mistakes and fraud are more easily uncovered.

Acknowledgements

Much of the information in this SETACTEACH was taken from the 1999 SETACTIP (Technical Issue Paper) entitled "Sound Science." "Sound Science" can be found on the SETAC web site, http://www.setac.org.

More on evaluating science

Barker D. 1990. Maybe yes, maybe no: a guide for young skeptics. Strassburg B, illustrator. Amherst (NY): Prometheus Books. 80 p.

An introduction to skepticism, critical thinking, and investigating claims is presented through the story of 10-year-old Andrea checking out a report of ghosts at her friends' house. Following the fictional story are some guidelines on critical thinking and investigation.

Barnard C, Gilbert F, McGregor P. 1993. Asking questions in biology: design, analysis and presentation in practical work. Essex (England): Longman Scientific and Technical. [copublished in the United States with John Wiley & Sons Inc., New York]. 157 p.

Much of this book is highly detailed, but the first chapter, "Doing Science," presents a good description of asking scientific questions, experimental design, and pitfalls in design.

Barrett S. c1997-2007. Quackwatch: Your Guide to Quackery, Health Fraud, and Intelligent Decisions [Internet]. Allentown (PA): Quackwatch; [updated 2007 April 29; cited 2007 May 31]. Available from: http://www.quackwatch.org/index.html

This web site states, "Quackwatch, Inc., which was a member of Consumer Federation of America from 1973 through 2003, is a nonprofit corporation whose purpose is to combat health-related frauds, myths, fads, fallacies, and misconduct. Its primary focus is on quackery-related information that is difficult or impossible to get elsewhere. Founded by Dr. Stephen Barrett in 1969 as the Lehigh Valley Committee Against Health Fraud, it was incorporated in 1970. In 1997, it assumed its current name and began developing a worldwide network of volunteers and expert advisors."

Brownlee C. 2006. Peer review under the microscope. Science News 170(25):392. Also available on line, for subscribers only, at: http://www.sciencenews.org.

This article discusses the history and workings of the scientific peer review process, and recent changes using internet access and preprint servers to make the process more open.

Carr JJ. 1992. The art of science: a practical guide to experiments, observations, and handling data. San Diego (CA): Hightext Publications Inc. 365 p.

This is a very readable book. Some chapters are highly detailed and mathematical, but for the purposes of understanding and judging scientific work, the introductory chapters "What is Science About?", "Thinking Scientifically," and "Theory, Hypothesis, and Law—What's the Difference?" are excellent. The Chapter "How to Win a Science Fair" uses previously discussed principles in a way students can relate to. Appendix A, "Counterfeits of Truth," explores common fallacies in scientific material.

The Committee for Skeptical Inquiry [Internet]. Amherst (NY): The Committee for Skeptical Inquiry; [updated 2007 May 20; cited 2007 May 31]. Available from: http://www.csicop.org/

Their web site states, "The Committee for Skeptical Inquiry encourages the critical investigation of paranormal and fringe-science claims from a responsible, scientific point of view and disseminates factual information about the results of such inquiries to the scientific community and the public." Many articles from their magazine, "Skeptical Inquirer," are available on line.

Committee on Science, Engineering, and Public Policy, National Academy of Sciences (US); National Academy of Engineering (US); Institute of Medicine (US). 1995. On being a scientist: responsible conduct in research. Washington (DC): National Academy Press. 27 p.

This booklet is aimed at graduate students and new researchers, but provides thoughtful material for anyone concerning the ethical issues in science, from treatment of data to scientists' obligations to society. This can also be accessed online through the National Academy of Sciences' web site, http://www.nas.edu

Fraser AB. Bad Science [Internet]. c1995-2007. University Park (PA): Pennsylvania State University; [cited 2007 May 31]. Available from: http://www.ems.psu.edu/~fraser/BadScience.html

Dr. Alistair B. Fraser, Professor of Meteorology at Pennsylvania State University, explores examples of bad science in educational facilities, articles, and texts. Links to other bad science pages (such as bad astronomy, bad chemistry, and bad meteorology) are provided. Professor Fraser's qualifications can be found at http://www.ems.psu.edu/~fraser/cv/

Gardner M. 1989. Science: good, bad, and bogus. Buffalo (NY): Prometheus Books. 412 p.

Martin Gardner is a prolific science writer, philosopher, and magician. This collection of his articles assessing claims in areas from ESP to ape intelligence is only one sample of his work. Some topics are slightly dated, but the principles illustrated still apply.

Montagu A. (ed.). 1984. Science and creationism. New York: Oxford University Press. 416 p.

Twenty essays from people such as Stephen J. Gould and Isaac Asimov discuss a variety of issues from the controversy over "creation-science," such as the age of the earth and dating methods; the fossil record; what constitutes a good scientific theory; aspects of, gaps in, and misuse of evolutionary theory; thoughtful criticism of creation theory and discussion of whether it can be considered science. Essays stand alone, but are tied together by participation in or commentary on the trial ruling the 1981 Arkansas law, "Balanced Treatments for Creation-Science and Evolution-Science Act," unconstitutional.

Plait P. 2002. Bad astronomy. New York: John Wiley & Sons, Inc. 277 p.

In a readable and humorous style, Philip Plait addresses topics ranging from the inaccuracy of the term “meteoric rise” to space-time, UFOs and astrology.

The Skeptics Society. 1992-2007. The Skeptics Society and Skeptic Magazine [Internet]. Altadena (CA): The Skeptics Society; [cited 2007 May 31]. Available from: http://www.skeptic.com/index.html

Judging the Quality of Science — Teacher's Guide

SETAC**TEACH***

* Topics for Exploration, Activities and Classroom Help

What United States National Science Standards can be met through studying evaluation of science?

Grades K-4, 5-8, and 9-12:

Content Standard A: Science as Inquiry
Content Standard E: Science and Technology
Content Standard F: Science in Personal and Social Perspectives
Content Standard G: History and Nature of Science

Source: National Research Council. 1996. *National Science Education Standards*. National Academy Press, Washington, DC. Available online at: http://www.education-world.com/standards/national/.

Suggested activities

1. Have students examine the following example extracts from articles and an advertisement, and how the material in this SETAC**TEACH** article applies to them. Then have students select a newspaper, magazine, or internet article of their own and evaluate the soundness of the science in a similar manner.

   Example 1 Article from Teen Vogue, September 2006, entitled "Sweet Nothings: Is Your Favorite Low-cal Sweetener too Good to be True?", by Eva Chen

   a) The article quotes "Kim Collier, a Sacramento, California-based sports nutritionist for AthletiCamps" and "Rebecca Appleman, R.D., a nutritionist based in New York City."

   Are the author and experts qualified? You decide. According to Eva Chen's 2007 blog http://www.teenvogue.com/beauty/blogs/beauty/all_about_eva/index.html), she was in the premed program at Johns Hopkins University for three years. According to her profile on the 2007 Athleticamps web site, http://www.athleticamps.com/cycling/about_us/staff_and_contributors, Kim Collier is an athlete with an M.A. in Holistic Health and was formerly a personal trainer. "R.D." stands for Registered Dietitian. Rebecca Appleman was not found in a cursory Internet search.

b) Ms. Chen describes three common artificial sweeteners (aspartame, saccharin, and sucralose): how sweet they are, how they work, and how they were discovered. References are not provided, but at least some of the facts can be verified easily with an Internet search. The author states, "…their origins are less than organic. All were discovered—by accident—in laboratories: aspartame by a scientist creating an ulcer medication; saccharin by a chemist working with coal tar derivatives; and sucralose by a researcher inventing an insecticide."

What does "organic" mean in this section? Do the compounds' methods of discovery have anything to do with their safety?

c) "'Girls think they can use fours spoonfuls of fake sugar and it's better than half a spoonful of real stuff,' notes Collier. 'They're wrong. These sweeteners can be a lot worse for you.' A 2004 study by Purdue University in Indiana backs up her theory, indicating that they may contribute to weight gain because the body becomes unable to gauge the number of calories consumed in other sweet-tasting foods."

Although further reference was not given, the study was located with an Internet search: Davidson TL, Swithers SE. 2004. A Pavlovian approach to the problem of obesity. International Journal of Obesity 28:933-935. Ms. Chen's statement does accurately describe one of the conclusions of the study.

d) "Some scientists caution that weight gain (and the less-than-attractive side effects of diarrhea and gas) could be the least of your concerns."

Who are "some scientists"? What was the study, or who was the expert, that provided those details on side effects?

e) "Aspartame, found in many diet sodas, was discovered in a 2005 Italian study to be associated with leukemia and lymphoma in rats that were fed the human equivalent of three cans of diet soda a day."

A reference was not given, but an article in *Science News* (Not so sweet: cancers in rats that consumed aspartame. 2006. *Science News* 169:101) indicates this is work by Dr. Morando Soffritti, who works at the European Ramazzini Foundation of Oncology and Environmental Sciences, and the work is published in the March 2006 issue of *Environmental Health Perspectives*. From what is presented in the *Science News* article, Ms. Chen has correctly presented conclusions from the study. However, according to an FDA article, the European Food Safety Authority ruled that the conclusions in the study were not supported by the data, and the U.S. Food and Drug Administration (USFDA) is reviewing the data. (http://www.fda.gov/fdac/features/2006/406_sweeteners.html USFDA's *FDA Consumer* magazine, July-August 2006, article entitled "Artificial sweeteners: no calories…sweet!").

f) "Anecdotal evidence also has it associated with migraines, seizures, and Alzheimer's disease. 'I have a college-aged client who drank her first diet soda before class, had her second afterward, and throughout the day, had two six-packs,' says Collier. 'She began displaying symptoms of multiple sclerosis. It may not be proven in studies, but I know that when she cut out diet pop, she got better.'"

Anecdotal evidence may suggest avenues of research, but is not the result of controlled studies. For example, the National Multiple Sclerosis Society does not support claims that aspartame causes MS, since they are undocumented claims, unsupported by evidence or peer-reviewed research. They also state that symptoms of MS may come and go randomly. Read their article at: http://www.nationalmssociety.org/site/PageServer?pagename=HOM_ABOUT_headlines_aspartame

g) "Saccharin, one of the oldest artificial sweeteners, also has a checkered past. In light of studies from the 1970s in which rats developed bladder cancer after being fed high quantities, the FDA once sought to ban the additive. It's since been declared safe by the World Health Organization, but the National Cancer Institute and FDA stated (about three decades ago) that there was "suggestive evidence" that people who were heavy users had some indication of an increased risk of cancer. "

The meaning of the word "safe" here is not clear. Similarly, "suggestive evidence" is a vague term. However, articles (e.g. http://www.fda.gov/fdac/features/1999/699sugar.html USFDA's *FDA Consumer* magazine, article entitled "Sugar substitutes: Americans opt for sweetness and lite," by John Henkel) that refer to this study generally use "suggestive evidence" in quotes, so that may be the phrase used in the original article. Since the reference is not given in this article, it is difficult to examine what the original article may have meant by the phrase.

h) Would you use artificial sweeteners? Why or why not?

Example 2 Paragraph from "Awake!", September 2006 issue, article entitled, "Does it Matter What You Believe?", no author listed.

"Do you think that life has a purpose? If evolution were true, then the statement quoted in the journal *Scientific American* would be valid: 'Our modern understanding of evolution implies…that ultimate meaning in life is nonexistent.'"

a) The author is anonymous, so we cannot judge his or her qualifications.

b) There is not enough information provided with the quotation from the *Scientific American* article to verify the statement and the context. It is impossible to tell from the information given where the quote came from. There is no date, issue number, article title or author given. The quotation is described as "the statement quoted in the journal," so what is presented in the "Awake!" article is a quote of a quote. Finally, we cannot evaluate what may have been contained within the ellipsis points (the group of three periods that show something was omitted from the quotation).

c) The author believes that acceptance of evolution means that life has no meaning. The incomplete, out-of-context quotation from *Scientific American* is an example of putting a reference to science in a nonscientific article to bolster an argument. Note

that we are not discussing acceptance or rejection of evolutionary theory, but only looking at the way the argument is presented. A more straightforward statement would simply state the author's opinion: "Do you think that life has a purpose? If evolution were true, then life would have no meaning."

Example 3 Advertisement for Hylexin®, *Parade* magazine, Sept. 10, 2006. A similar advertisement can be found on their web site at http://www.Hylexin.com (Bremenn Research Labs®.)

This product claims to "dramatically reduce the appearance of serious dark circles." At the top of the ad is a photograph of a young woman with dark circles under her eyes, but they are obviously colored onto her face, since they are unnaturally dark, rectangular, and do not follow the shape of the eye. The text of the ad is written by Tiffany Stobel, although we are not told who she is or what her qualifications are.

The ad states that the kind of dark circles the product works on are caused by capillaries around the eye leaking blood, which then oxidizes through hemoglobin degradation and creates bluish-red pigmentation under the eyes. The ad claims,

> "In scientific studies, Hylexin® has been shown to help reduce hemoglobin degradation byproducts by optimizing enzymatic activity…reducing the appearance of deep discoloration and red-blue pigmentation. Studies also confirm that Hylexin® actually helps strengthen the capillary matrix to help stop the 'leaking'…so your orbital eye area is protected against further damage. Science is soooo cool!"

The next paragraph continues,

> "In a double-blind clinical trail more than 72% of women who had serious dark circles and used Hylexin's® key compound saw an obvious, tangible, visible reduction in the color and severity of their dark circles. These results were confirmed with high-speed laboratory photography that clearly showed a significant reduction in the appearance of the blue and red color that makes dark eye circles so noticeable and so unattractive."

No references to the studies mentioned are provided, so we cannot verify the methods or results, or evaluate the conclusions. It also may be that the company selling this product funded and conducted the studies, and that company certainly gains when conclusions state that their product works. Sample sizes are not reported for any of the studies.

The Hylexin® web site does not provide any details about the studies either. The only scientist or data review mentioned is a short article from *Instyle* (August 2006) reproduced on the web site that states "N.Y.C. dermatologist Brad Katchen (who reviewed the Hylexin studies) says he is 'wary of the claims' but is

interested in seeing more independent clinical research." An Internet search reveals that Dr. Katchen is the founder of SkinCareLab (http://skincarerx.com/).

Would you buy this product based on the information presented in the ad? The cost is $95 for one 0.78-oz. tube.

2. Have students select a newspaper, magazine, or Internet article, find a technical study (or summary of it in a scientific publication) discussed in the popular press article, and evaluate whether the conclusions were the same.

   For example, carefully compare the content provided in the *Science News* article:

   Not so sweet: cancers in rats that consumed aspartame. 2006. *Science News* 169:101. Also available on line at http://www.sciencenews.org for subscribers.

   with the section in the popular article from *Teen Vogue* discussed in Exercise 1, Example 1 above.