Peer Review by School Science Students: Its Role in Scientific Inquiry

William S. Carlsen[1]
Penn State University

Christine M. Cunningham
Nancy M. Trautmann
Cornell University

Paper presented at the annual meeting of the
National Association for Research in Science Teaching

St. Louis, MO
March 25-28, 2001

This research has been supported by grants from the National Science Foundation (9454428, 9618142, and 9979516) and the Dwight D. Eisenhower Title IIA program (0132-00-0008). The views expressed are those of the authors.

---

[1] To whom correspondence should be addressed: Dept. of Curriculum & Instruction, 173A Chambers Building, Penn State University, University Park, PA 16802. Email wcarlsen@psu.edu. This paper includes some unpublished material from Carlsen, Cunningham, and Trautmann (March, 2001).

Introduction

Sociologists of science have described how communities, not individuals, are the agents of scientific knowledge—for a "fact" to be produced, it must be judged and accepted by a larger collective of scientists. This is accomplished through a system of formal and informal peer review. However, this central feature of science is absent in virtually all science classes and curricula. We are interested in how aspects of the peer review system impact students' abilities to construct and defend scientific arguments.

Some general questions driving this work:

- Does the process of assessing other students' projects and responding to others' critiques foster student understandings of (a) what constitutes valid scientific research questions, evidence, and claims; and (b) the role of peer review in creating scientific knowledge?
- How do students shape and reshape the presentation and justification of their findings based on the feedback they receive from peers?
- How does face-to-face peer review compare with anonymous web-based peer review in terms of development of student understandings, constructiveness of feedback, and student comfort with giving and receiving reviews?
- Is peer review by student researchers reliable and valid?
- Can peer interactions cultivate high-quality original student research?

This paper, from a program (*Environmental Inquiry: Learning Science as Science is Practiced*) focusing on sociologically authentic science education, offers a theoretical framework for thinking about peer review in relation to science curriculum and instruction. We review the functions of peer review as reported in scientific fields, as well as perspectives of science and technology studies (S&TS), including sociology, history, and philosophy of science. We identify normative, economic, and epistemological functions of peer review, variously emphasized in the different S&TS fields. We also outline ethical and psychological considerations that interact with "authenticity" in guiding instructional decisions concerning the use of students as peer reviewers.

We then examine the use of peer review by secondary students in successive annual Research Congresses, attending to issues of reliability, validity, and insights gained by students. We also briefly describe other forms of peer review with which we are experimenting, including web-based anonymous review.

The data used for this paper include questionnaires probing students' understandings of peer review, administered after students completed extended toxicology research projects. Group interviews with 31 students (from the second Congress) further probe students' perspectives on peer review and the activities in which they participated. Reliability of students' reviews are measured statistically, and validity is assessed through comparison of students' reviews by reviews done independently by 13 graduate students enrolled in environmental science fields.

Theoretical Framework

State and national science education standards call for school science to promote public understanding about the practices and processes of science. Science educators concerned with depicting science in more authentic ways have described and researched attributes of science that deserve attention in school science, including the use of inscriptions (Roth & McGinn, 1998), group work (Bianchini, 1997), and argumentation (Kelly, Druker, & Chen, 1998). However, one of the pivotal steps in the creation of scientific knowledge—formal peer review—has been largely overlooked.

Scientific peer review: Perspectives from science. Most of what has been written about peer review by mainstream scientists has functioned to analyze, critique, and refine the review process. For example, in 1990, *JAMA: The Journal of the American Medical Association* published a special issue focusing on the subject of peer review; articles reported the effects of blinding (in a naturalistic study, blinding increased the quality of the review as gauged by independent editorial judgment: McNutt, Evans, Fletcher, & Fletcher, 1990), authors' evaluations of the quality of editorial review (assessments were comparable between authors whose manuscripts were accepted and those whose were rejected: Garfunkel, Lawson, Hamrick, & Ulshen, 1990), and disciplinary differences in peer review (they were great: Hargens, 1990). Eight years later, another special issue in the same journal reported that in experimental trials, blinding had no significant effect on reviewers' decisions (Godlee, Gale, & Martyn, 1998; van Rooyen, Godlee, Evans, Smith, & Black, 1998). However, other research has emphasized the importance of reviewer qualifications over other predictors of review outcome, such as individual reviewer rejection rates (Callaham, Baxt, Waeckerle, & Wears, 1998). In fields like medical science, where the consequences of publication decisions may be economically profound, we should probably expect continued interest by scientists in its implementation. In other fields, discussions about peer review are much less likely to be carried out publicly.

Scientists recognize the importance of peer review, even when its mechanisms are not routinely discussed. For example, conference papers delivered without subsequent publication in the referred literature are sometimes denigrated as "gray literature," not as trustworthy as their peer-reviewed counterparts (Lacanilao, 1997).[2]

Peer review concerning publication of scientific results is different than peer review concerning the funding of scientific research, and the difference has probably grown as funding agencies like the National Science Foundation have increased the participation of representatives from industry, education, and other organizations in reviewing proposals, partly in response to public pressure from outside the scientific community (Dickson, 1988) and partly in response to revelations about scientific misconduct and revelations about economic conflicts of interest

---

[2] However, it should be noted that in some fields and particularly in non-Western nations, the gray literature is an important resource for scientists and policymakers (Boon & Brock, 1994; Deorani & Dabral, 1997).

(Chubin, 1990). Peer review of published research is much more likely to insulated from direct participation by non-experts.

Most of the time, peer review within the scientific community is a background phenomenon. Its important role may be openly asserted at times of crisis, such as during the cold fusion controversy of the late 1980s: After a paroxysm of high-profile publicity that bypassed scientists and appealed directly to the mainstream media, the physics community successfully reasserted the importance of independent experimentation, replication, and most significantly, peer review to debunk cold fusion's claims (Gieryn, 1999).

Finally, it should be noted that scientists themselves are cognizant of the potential for peer review to function conservatively; that is, for it to discourage truly imaginative and creative methodologies and claims, because they may be difficult to publish (Horrobin, 1990; Ziman, 1994).

Scientific peer review: Perspectives from science studies. Science and technology studies (S&TS) scholars have described the importance of scientific communities in fact production; scientists must convince their colleagues to accept their claims (Hull, 1988; Latour & Woolgar, 1986; Longino, 1990). In peer review prior to publication, methods and results are carefully scrutinized and critiqued, often anonymously. Peer review is only one step in the process of "fact-making" (more than half of all scientific papers are not cited within five years of publication (Campanario, 1996); clearly, there is a difference between a published claim and an accepted fact), but it is a crucial one.

The origins of peer review are usually traced back to the 1700s, but "antecedent" peer review practices have been identified by historians in the 1600s (Kronick, 1990). Modern editorial peer reviewing developed haphazardly in a variety of forms beginning in the mid-19th century, but became a general and institutionalized practice after World War II (Burnham, 1990).

The mechanism of peer review can be easily related to the Mertonian norms that influenced the growth of sociology of science (Merton, 1973), particularly the norms of communism and organized skepticism. According to the first of these norms, scientific knowledge is common property and scientists' claims to knowledge are restricted to getting credit for new discoveries; according to the second norm, scientists are enjoined to withhold judgment on new claims until empirical facts are assembled and scrutinized. Contemporary sociologists have identified counternorms that reject Merton's taxonomy but remain consistent with the institutional practices of peer review; for example, the counternorm of secrecy is both prevalent and to a large extent institutionalized in review practice. Although journals release new knowledge into the public domain (communism), discretion by reviewers concerning what they review (secrecy) helps to prevent priority disputes (Mulkay, 1975; Mulkay, 1991). In an age when knowledge is increasingly a valuable commodity, such disputes can have serious implications.

Historian of science David Hull's fascinating account of the process of peer review in one community of biologists demonstrates many facets of the process that are usually invisible. For example, by sending a manuscript to a large number of reviewers, an editor can generate enough conflicting advice that the decision to "publish if reviewers' concerns are addressed" is functionally a decision to reject. Manipulating the editorial system in this fashion may suppress the emergence of unfavored ideas or methodologies; however it may stimulate dissenters to break away from the community and create a rival community with its own forms of publication (Hull, 1988).

A recent book about the "Science Wars," by sociologist of science Thomas Gieryn, argues that scientists routinely engage in what he calls "boundary work," defining and redefining the edges of science to suit their needs. At times like the cold fusion controversy, those needs may best be served by insistence on the primacy of peer review (Gieryn, 1999); at other times, the community may acquiesce to the establishment of truth claims through stage tricks (Gieryn & Figert, 1990). Institutionally, rigorous peer review provides a criterion of demarcation between science and non-science.

Scientific peer review: Educational perspectives. Given the importance of peer review in the scientific process, its absence from the school science and science education literatures is surprising. A handful of professional publications for science educators have printed peer-reviewing “how to” articles (e.g., Gratz, 1990; Pechenik & Tashiro, 1991); another handful of research papers—mostly at the college teaching level—report effects of peer reviewing in the production of student research papers or multimedia reports. Student and faculty concerns about peer reviewing—typically reported anecdotally—often resemble those raised in the scientific and science studies literatures. For example, Gratz (1990) noted that students sometimes express concerns that their ideas will be stolen or that reviewers may not provide competent feedback on how to improve their writing.

New technologies have provided interesting tools for engaging students in peer review (e.g., Rushton, Ramsey, & Rada, 1993). MacLeod (1999) reported on the use of two types of technology for engaging business students in technology-mediated peer review: asynchronous public newsgroups (not anonymous) and anonymous real-time review. Although there was some preference among students for reviews that were not anonymous, half of the students reported that anonymity helped them to be more honest in their feedback, and a majority indicated that they incorporated some peer reviewer comments into the final drafts of their papers (which were evaluated by the instructor, not by peers).

The usual emphasis of both high-tech and low-tech studies has been the role of peer review in improving students' writing (Koprowski, 1997; Towns et al., 1999) and presentations (Bos, Krajcik, & Soloway, 1997) or, occasionally, on modeling peer review to help students "appreciate" it (Lightfoot, 1998). The use of peer review in teaching science may motivate students to produce better work--to impress their peers--and it may signal a more collaborative relationship between teacher and students (Billington, 1997). However, teachers may report difficulty getting students to write critically or substantively about each others' work, and peer reviews may focus on the surface level features of reports. This phenomenon may be attributable to students' lack of in-depth content knowledge, their unwillingness to criticize classmates, or the logistics of how peer review assignments are scheduled (Bos et al., 1997).

To date, little emphasis has been placed on the syntactical function of peer review: the filtering of facts-in-the-making from analytical noise (but see Cunningham & Helms, 1998; Kelly, Carlsen, & Cunningham, 1993). Content is typically a relatively unimportant part of the peer review process, representing about one-fifth of overall evaluation weighting in one representative approach (Billington, 1997). Furthermore, almost nothing has been published on the reliability and validity of peer assessments at the precollege level, a situation that is only slightly better for college-level studies (Topping, 1998). The mechanisms and functions of peer review in precollege school science are largely unexplored.

Bringing together science, science studies, and science education. A principal goal of our project is to integrate the diverse perspectives of the worlds of science, science studies, and education to craft approaches to teaching science that represent science authentically, while also engaging all students' interests. We believe that the sociological nature of scientific practice is important not because it challenges science's status (although it sometimes does), but because it invites participation by a wider range of citizens and better prepares them to participate fully in an information-based society in which science and technology are important persuasive currency. Consequently, it is important to us to carefully observe who engages and who succeeds in what we call "sociologically authentic school science."

Part of this project is to identify norms of teaching practice that potentially conflict with "authenticity" to mitigate against peer review as an instructional practice. These norms are curricular ("How can we ensure that important content is taught?"), cognitive ("What if students reject true claims or accept false ones?"), affective ("Will critical or inappropriate feedback reduce students' interest or confidence?"), and accountability-related ("Will students plagiarize each other?"). In working with teachers, scientists, and students, what we have found is that for each norm, parallel questions have been observed in "real" scientific communities. For example, scientific peer review sometimes results in the initial rejection of ultimately important claims; these initial rejections may help improve research reporting (Campanario, 1996) or may motivate scientists to work extraordinarily hard to disprove their detractors (Hull, 1988). Of course, the existence of functional similarities between science and science education should not lead to an "anything 'authentic' goes" attitude: educators have different responsibilities than journal editors. Nevertheless, considerations that at first appear exclusively educational usually have analogs in scientific communities.

This convergence between educational concerns and those of science practice cohere fortuitously with the emergence of a view of learning as social participation, both in and out of science (Wenger, 1998). Scientific literacy represents much more than the individualistic mastery of facts or the ability to respond appropriately to questions from an interrogator; scientific literacy involves participation in a complex community of practice, where collaboration and competition coexist in discipline-specific ways.

Peer Review in Practice

This section of the paper describes the engagement of teachers and students in peer review, in the context of an innovative student research program. In our program, *Environmental Inquiry: Learning Science as Science is Practiced*, teachers and eventually students conduct original research projects on environmental topics. Teacher work involves residential summer collaboration between selected science teachers, scientists, and university science educators. These summer efforts focus on identifying and adapting promising research methods in five targeted *Environmental Inquiry* curriculum areas: (1) environmental toxicology, (2) bioremediation and waste management, (3) ecology of invasive species, (4) watershed dynamics, and (5) urban ecosystem modeling. These curricular areas were selected in part because they represent areas of scientific strength and opportunity for Cornell, and in part because they lend themselves readily to local student-driven scientific investigations. Our project completed three successive summers of this developmental cycle and participants created and pilot-tested investigations in a number of research areas. The first of four teacher texts on the project includes the toxicology experiments utilized by students in this paper (Trautmann, Carlsen, Cunningham, & Krasny, in press).

The model of student inquiry developed in *Environmental Inquiry* (EI) rejects the "naïve inductivist" view of science that portrays science as a straightforward, rational movement from naturalistic observation to inductive generalization.[3] EI's framework begins with clearly defined well-established <u>protocols</u> for investigative methods, with the perverse twist that they are almost immediately applied to novel problems. For example, to determine the toxicity of a particular chemical, students are given very specific instructions about how to prepare a serial dilution of the chemical and how to conduct a bioassay using a particular plant or animal. However, neither the students nor the teacher is ever given a “right answer;” consequently, the resolution of the chemical’s toxicity is locally established. The protocol is provided; the answer is not. At this first stage of inquiry, methods are *adopted*, not created, and conclusions are *created*, not confirmed. Compared to conventional cookbook science “investigations” in which results are pre-established, we believe that protocol labs better resemble the mechanism through which novices are brought into scientific communities of practice.

The second stage of student inquiry in the EI model is interactive research. In interactive research projects, students work in groups to plan and conduct well-designed and controlled experiments to test their own novel hypotheses, and then communicate their findings to others. The communication phase, which is interactive, is intended to help students further develop their understandings and to improve their science skills through a process of social argumentation, persuasion, and peer review. To date, we have developed and explored three types of interactive research: (I) Research Congresses, (II) Technology-Mediated Peer Review, and (III) Engineering Design Challenges.

For peer review to be successfully integrated into school science, we believe that it is necessary to develop written guidelines and other training materials for students new to the peer review process. Fortunately, good models for the beginning scientist are available (e.g., Jaeger & Toft, 1998) and provide a starting point. *EI* has developed materials of this kind for the

---

[3] For a discussion of “naïve inductivism,” see Chalmers (1982) and Millar (1989).

secondary science student. They are available in traditional print format[4] and in an HTML tutorial (http://ei.cornell.edu/Bioassays/PeerReview/). In fact, most of the materials used in this project are currently online at our web site, as are detailed instructions for engaging students in environmental toxicology research using bioassays (http://ei.cornell.edu).

Peer review on environmental toxicology research projects: Preliminary work. One topic of EI student research in the 1998-99 and 1999-2000 school years was the assessment of the toxicity of road deicers and other locally applied chemicals through the use of bioassays (lettuce seed germination and growth, aquatic plant asexual reproduction and growth, and aquatic invertebrate behavior). Student research projects involved participation in a distributed "research community," which used either face-to-face meetings at a Research Congress or Internet message boards and other resources to link geographically distributed schools and classrooms. For example, when students began their research projects, they logged on to a Cornell-based file server, which created a user account, then gave them access to tutorials and background information on bioassays. Working in small groups, students developed research hypotheses, entered descriptions of their hypotheses in a public database, and were then given access to online conferences that provide technical assistance for these school-based projects. In our pilot year (1998-99), students from three schools completed the experimental process by preparing public presentations, which were evaluated by peers at a Student Research Congress that brought student experimenters face-to-face for the first time.

Over a three-week period in July 1999, ten teachers from eight schools worked with program staff and a computer programmer to expand online resources for the support of peer

---

[4] For a paper or PDF copy of these materials, please contact Nancy Trautmann, Cornell Center for the Environment, Rice Hall, Cornell University, Ithaca, NY 14853, nmt2@cornell.edu.

review, and this system was evaluated during 1999-2000 and 2000-2001 in bioassay research projects at nine schools.

We conducted a second student Research Congress in April 2000. At this one-day event, 46 students from a wide range of grade 7-12 classes convened on the Cornell campus with their teachers to present posters summarizing their toxicology research. Students completed a total of 112 peer reviews of each others' research projects. The evaluation instrument used included two rubrics: a six-item "poster presentation and content" subscale and a five-item "experimental design" subscale. Subsequent analysis showed strong subscale alpha reliabilities of 0.83 and 0.70 respectively. Because different posters were rated by different students, analysis of variance was used to evaluate within-poster agreement across student raters. In spite of a wide range of student raters (in addition to variation in age, classes represented ability tracks ranging from 9th grade remedial to A.P. Environmental Science), reliability by reviewed-poster was good, and between-poster variance was much greater than within poster variance ($p<.001$). It should be noted, however, although a range of student ages and abilities were represented at the Congress, participation at this Saturday event was limited to volunteers, typically 5-7 students brought by a teacher. In addition to this selection effect, the evaluation rubric was identical to that used locally; consequently, most students were familiar with the criteria for evaluation in advance.

Although the peer review process had a number of general elements that were common across the eight study schools (e.g., access to a computer tutorial, online conferences, and the online peer review system), individual teachers implemented their toxicology research projects in different ways, providing us with natural variation for assessing the strengths and weaknesses of different teacher-selected approaches. For example, some teachers have used students' peer reviews (what they wrote as well as what they received) in calculating students' grades; others have not.

Reliability, Validity, and Learning Outcomes, 2000-2001

The preliminary work described above left several questions unanswered:

- How consistent are students as peer reviewers?
- How reliable is the evaluation instrument that we directed the students to use?
- How do students' judgments compare with subject-matter experts?
- Do students' experiences with peer review vary with grade level, subject, gender, or other demographics?
- What learning outcomes result, with respect to understandings about the nature of science?
- What other educational outcomes result?
- How are curriculum materials being used by teachers?

To address these questions, we undertook another cycle of student experimentation and peer review, culminating in a Research Congress at Cornell in December, 2000. The process varied in several ways from prior Congresses. First, we collected background information from each participant and assigned each an ID code that went on their actual reviews. (In previous review sessions, all reviews were completely anonymous; consequently, we were limited in the analyses we could undertake). Second, we utilized our first cohort of Fellows participating in the CEIRP project[5] as subject-matter experts for the purpose of assessing the validity of high school students' reviews. Although the high school student reviewers reviewed 2-3 projects each, each CEIRP Fellow reviewed 8-9 projects. Approximately half of the Fellows' reviews were done during the actual Research Congress (Fellows circulated among the posters and talked with

---

[5] The Cornell Environmental Inquiry Research Partnership (CEIRP) is an NSF-funded GK12 project, described in Cunningham, Meyer, and Avery (March, 2001). CEIRP Fellows range from senior undergraduates to doctoral students, and all have substantial prior research experience in the environmental sciences. As part of the GK12 project, the Fellows have a weekly seminar and extensive school-based work, including assisting teachers implementing EI activities.

student researchers) and half were done in the week after the Congress (student researchers left their posters at Cornell). Third, we conducted interviews with all of the participating students after the Congress, in their original school groups. These interviews were done by an EI staff member assisted by a CEIRP Fellow. These interviews explored students' perceptions of the process of peer review. Finally, we administered a post-test to both student researcher and Fellows, to explore and contrast their perspectives on the function of peer review in science.

Results

At the December 2000 Congress, 32 student researchers represented 19 groups that had prepared and brought posters and other artifacts describing their original toxicology research projects. The projects ranged from basic protocols taken directly from the EI curriculum materials to ambitious and original projects, such as an evaluation of the effectiveness of toxin uptake by hydroponically grown plants. Students ranged from 9th to 12th graders, and the courses represented ranged from very basic non-college preparatory to AP Environmental Science. A Saturday morning was devoted to orienting students to the evaluation process (all were familiar with the evaluation rubric, which had been distributed to them at the beginning of their research) and rotations by student reviewers among posters. Students were randomly assigned posters to review and were encouraged to discuss the research projects with the authors. A rotation was established so that each poster was constantly manned but that all students circulated and reviewed others' work.

Reliability of student reviews and the evaluation instrument. The scoring rubric used by reviewers is reproduced in Appendix A. It contained eight items corresponding to different aspects of the research and reporting process. A four-point scale was used, and a descriptor assigned to each point (e.g., 1="Very Clear"). Seven of the items referred explicitly to the research poster and ancillary materials, and an eighth concerned presenters' ability to respond to questions asked by reviewers. The overall score calculated by the students was a simple arithmetic sum of the eight categories. We utilized an unweighted arithmetic mean for later analysis, which is mathematically equivalent.[6]

Table 1 displays baseline data for the 8 review subscores, the overall review score (RScore), and the overall review score for the seven items that did not involve oral questioning (PScore), for the 94 student reviews. Means and standard deviations were similar for the first seven items.

Table 1. Peer review instrument baselines, Student reviews only
(the form is reproduced in Appendix A)

Item	N	Minimum	Maximum	Mean	Std. Deviation
R1 Layout	94	1	4	1.78	.764
R2 Question	94	1	4	1.66	.824
R3 Design	94	1	4	1.71	.728
R4 Procedures	94	1	4	1.93	.895
R5 Data display	94	1	4	1.66	.874
R6 Results	94	1	4	1.86	.850
R7 Conclusions	94	1	4	1.90	.704
R8 Oral response	93	1	4	2.04	1.132
RSCORE Review Score total	94	1.000	3.250	1.81459	.462095
PSCORE Review Poster only	94	1.000	3.286	1.78267	.460285
Valid N (listwise)	93				

Inter-item correlations among the 8 items ranged from -.0432 to .5199. Cronbach's alpha was computed for the entire scale and for the scale with each item removed. Alphas for the eight alternatives ranged from .6616 to .7413. The standardized scale alpha of .7427 with all eight items was judged satisfactory, and provides good evidence for the internal reliability of the instrument as used by secondary students. (Reliability for the Fellows was higher, with a

---

[6] This solved a problem during validity assessment: because approximately half of the CEIRP fellows' reviews were done after the students had left, we calculated the mean for those reviews using the first seven categories.

standardized alpha of .8615, but also had higher intercorrelation among items, ranging from .1707 to .7481).

Reliability of students' reviewing skills was gauged through analysis of variance, contrasting between-group variance and within-group variance.[7] Between-group variance was significantly greater than within-group variance, at p<.05, lending support to the reliability of students' reviewing skills (see Table 2).

Table 2. ANOVA to Assess the Reliability of Student Reviews

RScore Variance	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.065	18	.337	1.832	.036
Within Groups	13.794	75	.184		
Total	19.858	93			

We also assessed the reliability of Fellow reviews, which was stronger (F=5.361, p<.001).

Validity of students' reviews. To assess the validity of students' reviews, we compared students' reviews with reviews done by Fellows. As a proxy for a true score for each poster, we used the means of independent review scores by 4-5 Fellows. The correlation between students' scores and Fellows' mean scores was 0.386, which is significant (p<.01). The regression line, 95% confidence intervals, and individual data points are plotted in Figure 1. To reveal overlapping data, data points have been randomly jittered in two dimensions.

One thing to note in Figure 1 is that students' reviews are consistently lower than Fellows' reviews. In independent ANOVAs, we found no significant effects of student gender, grade, or school on review scores.

---

[7] "Within-group" variance measures disagreements among randomly-assigned reviewers poster-by-poster. "Between-group" variance measures differences between posters.

The peer review process as implemented in the EI project calls for multiple reviews of students' work, just as scientists' work is reviewed by more than one person. When the means of

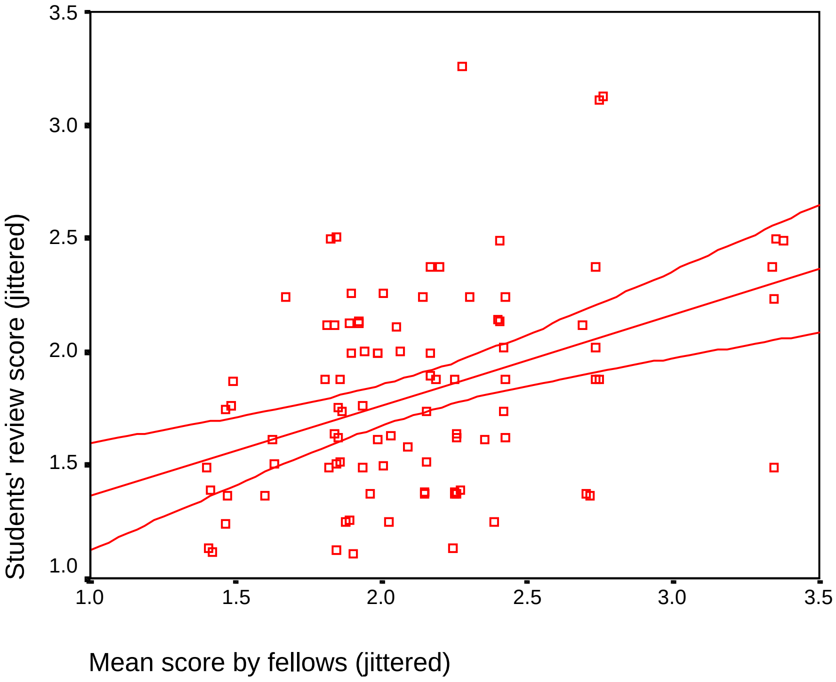

Figure 1. Correlation between Fellows' average review scores and individual students' scores.

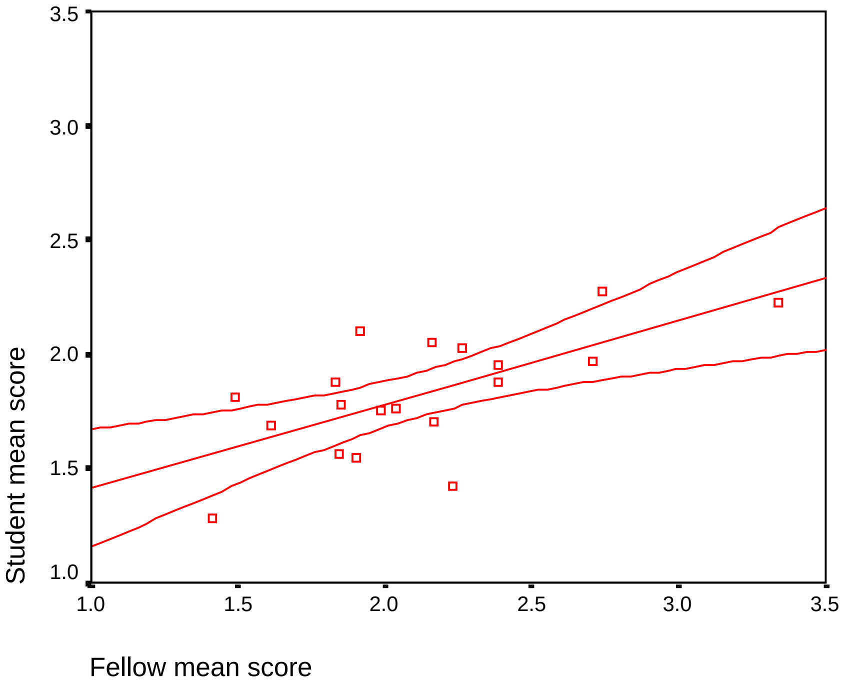

Figure 2. Correlation between Fellow review scores and student scores (both means).

student reviews (by poster) are compared to mean Fellows' review scores, the correlation jumps to 0.659 across the 19 posters, which we believe provides good evidence of the validity of student reviews ($p<.01$; see Figure 2). However, it should again be noted that students' review scores were consistently lower than Fellows' scores.

<u>Students' perceptions of peer review in action</u>. Appendix C summarizes the results of the group interviews conducted with students at the conclusion of the Research Congress. It includes descriptions of students' projects, by school and school subject. Students from approximately half of the schools reported that teachers assigned their research questions. All of the groups utilized computers in some fashion, but at the time of the Congress, groups from only one school had posted reports online. Students from two of the six schools had completed in-school face-to-face peer reviews and another had done a more informal review. Experience with peer review prior to the Cornell Congress had no detectible effect on students' scoring or on the scores they received.[8]

Students were consistently very positive in their reviews of the bioassay projects and generally positive (with a commonly expressed caveat about nervousness) about their experiences giving and receiving peer review (questions 6 and following). Students were mixed in their assessment of whether they would consider it fair to have peer assessments considered in determining their project grades. We are exploring–initially at the college level–mechanisms for reducing that concern (e.g., by having a routine appeals process); however we should emphasize that in our experience, an important part of what makes the peer review process work for precollege students is ensuring that all student reviewers have similar experiences with experimental protocols.

---

[8] These hypotheses were tested with analyses of variance.

Students' and Fellows' perceptions of the role of peer review in science. We administered a short questionnaire ("posttest") to participating students and to the CEIRP Fellows, in order to gauge and contrast their perspectives on peer review in science. The questionnaire is reproduced (with variable names) in Appendix B. The first item described a scenario and asked subjects to identify what they believe would be the most and least convincing reasons to believe an experimental claim. The reasons and subjects' responses are summarized in Tables 3 and 4. We recommend referring to the actual questionnaire as well, because the choices are written so that none of them is completely implausible.

Table 3. Most convincing reason to accept experimental claim

			Total
	CEIRP Fellows	HS Students	
1 a. Affiliation	0	8	8
2 b. Peer reviewed & published	6	7	13
3 c. Multiple experiments, personal	1	7	8
4 d. Other scientists similar observations	2	3	5
Total	**9**	**25**	**34**

Table 4. Least convincing reason to accept experimental claim

			Total
	CEIRP Fellows	HS Students	
1 a. Affiliation	2	6	8
2 b. Peer reviewed & published	0	2	2
3 c. Multiple experiments, personal	2	3	5
4 d. Other scientists similar observations	1	3	4
5 e. Boss reviewed and announced	4	11	15
Total	**9**	**25**	**34**

Although we didn't get responses from all Fellows and students, the partial results reveal interesting differences. For example, a third of the high school students felt that the fact that "Dr. Shales works at a major research hospital known for important discoveries in heart disease" was the most convincing reason given (what we call the "affiliation" argument); none of the Fellows did. Two student respondents even reported that the fact that "Dr. Shales's written report was reviewed by three other scientists and published in a medical journal" (the "peer review" argument) was the least convincing reason given. In general, however, students and Fellows showed greater agreement on what isn't a convincing reason that on what is.

In response to our queries concerning when peer review is typically undertaken, students were twice as likely as Fellows to believe that peer review is a common element of "developing an experimental question" (28% of students, 10% of Fellows) or "when they are developing methods or techniques for the research" (52% S, 20% F). These differences are significant (t-test p<.05, with critical values adjusted to correct for multiples), but there were no differences concerning beliefs about whether peer review occurs "when they apply for a grant to fund their research" (40% S, 60% F), "when they are analyzing their data" (32% S, 30% F), or "after they write up their results" (88% S, 80% F). These differences may be attributable in part to our web site, which has some peer collaboration tools that can be used in the early stages of student research. Although none of these subjects used that part of the system, some may have seen it.

Conclusions

We believe that this project provides evidence of the feasibility and importance of engaging students in peer review practices as part of their school science work. Although peer review is an indisputably important component of scientific practice, it is virtually nonexistent in the science curriculum. Properly conducted, experiments that utilize peer review can yield reliable, valid assessments of the products of student research. Although such data might prove to be a useful component of routine school assessment practices, their sociological role transcends the evaluation function of schooling. By making–and accepting–judgments on other students' work, student scientists make facts. They also learn that experiments are only one part of the process of science, which is fundamentally social.

Appendix A.
Rubric for Peer Review of Students' Research Projects *{Variable names added for reference}*

Your ID # _____________
Poster #_____________

Key:
1 – Very Clear
2 – Mostly Clear
3 – Somewhat Clear
4 – Largely Unclear

	(+) (-)	*Question*
Does the poster include: Title, Research Question, Hypothesis, Procedure, Results, Conclusions, & Acknowledgments?	1 2 3 4	*R1*
Is there a clear statement of the research question and hypothesis?	1 2 3 4	*R2*
Does the experiment appear to be designed appropriately to address the research question?	1 2 3 4	*R3*
Are the procedures described in enough detail for the experiment to be copied by someone else?	1 2 3 4	*R4*
Are the data presented clearly?	1 2 3 4	*R5*
Is there a clear explanation of the results?	1 2 3 4	*R6*
Do the conclusions seem well supported by the data?	1 2 3 4	*R7*
Presenters were able to answer questions clearly	1 2 3 4	*R8*

TOTAL SCORE ______

Comments:

What was a particular strength of this experimental design?

What suggestions do you have for improvement in this experimental design?

Appendix B. Peer Review Posttest *{Variable names added for reference}*

A Short Science Questionnaire

The following questions do not have absolute right answers. We are interested in your perspective, not in how many questions you get "right."

1. Dr. Amanda Shales, a scientist studying a new blood pressure medicine ("Cornellix") notes that some patients who take the drug develop heart valve problems. Knowing that, here are some reasons that a person might conclude that Cornellix causes heart valve damage:

   a. Dr. Shales works at a major research hospital known for important discoveries in heart disease.
   b. Dr. Shales's written report was reviewed by three other scientists and published in a medical journal.
   c. Dr. Shales conducted several experiments that convinced her of the link.
   d. Scientists studying other blood pressure medicines have reported similar side effects in some patients.
   e. Dr. Shales's boss—a very well-respected doctor—reviewed her experiment and reported her discovery to the public.

   To the scientific community, which reason do you think would be most convincing?
   Letter ___ *{Q1most}*

   To the scientific community, which reason do you think would be least convincing?
   Letter___ *{Q1least}*

2. When do scientists typically get anonymous feedback from other scientists? Check all that apply:

   ☐ When they are developing an experimental question *{Q2a}*
   ☐ When they apply for a grant to fund their research *{Q2b}*
   ☐ When they are developing methods or techniques for the research *{Q2c}*
   ☐ When they are analyzing their data *{Q2d}*
   ☐ After they write up their results *{Q2e}*

3. In describing an experiment at a research conference, a scientist concludes that a new genetically engineered soybean variety is toxic to earthworms. Which of the following do you think are likely to occur? Check all that apply:

   ☐ Scientists at the conference will ask him questions about his experimental design *{Q3a}*
   ☐ The soybean variety will be taken off the market *{Q3b}*
   ☐ Other scientists will conduct similar experiments to check his results *{Q3c}*
   ☐ The company that developed the soybean will discourage further research of this type *{Q3d}*
   ☐ The scientist will report his results to scientists who did not attend the conference *{Q3e}*

Appendix C. Summary of Post-Review Discussion Groups

Summary of Discussion Groups at Bioassay Congress – 12/00

The following table summarizes discussions held during the Student Bioassay Congress at Cornell University on 12/9/00. At the congress, 31 high school students from 6 schools presented a total of 19 posters representing the results of bioassay research. EI faculty and staff, in conjunction with Cornell graduate and undergraduate students, conducted the focus group discussions. The groups were organized by school, with groups of students collectively answering questions posed by each discussion leader.

Question	School and Course					
	1 – AP Environmental Science	**2 – Independent Study**	**3 – Chemistry in the Community**	**4 – Applied Science**	**5 – Applied Science**	**6 – AP Environmental Science**
1. Briefly describe the project that you did.	Bioassay on effect of Cl on Daphnia and lettuce seeds.	Compare toxicity of “environmentally safe” cleaner to regular cleaner.	Students who are volunteer firefighters decided to assess the toxicity of foam used to fight fires. 1st time doing bioassays, last minute project.	Were inspired by Erin Brockovich movie to look at chromium toxicity.	Bioassays using brine shrimp, lettuce seeds, duckweed to test metallic elements, WD40 and other lubricants, and storm drain water.	Class bioassay/practice project/put on EI website. Come up with individual ideas.
2. How did you decide what question you were going to research?	Assigned by teacher.	Students came up with their own question.	Fought a fire near a lake, and doing the bioassays in class sparked an idea. Wanted to be different than others.	Students helped pick overall topic, and teacher assigned specific topics.	Teacher specified organisms. Student groups generated list of possible chemicals.	Talk about septic tanks led to one project. Another idea stemmed from librarian, “All around where we live is agriculture… it’s where we live.”
3a. Did you use a computer during this project? How?	For graphing, data analysis	Excel, online research (couldn’t find good info on web)	Type up presentation. Internet research on toxicity of the foam (tried to find MSDS info).	Learned how to use animation in Powerpoint. “This was the 1st time I used a computer, and I learned it in 45 minutes.”	For posters – digital camera, tables, website research	Looked at EI website peer review for better ideas. Used Internet searches, satellite images.

3b. Did you do peer review online?	No	Not yet	No	Posted reports	No	No
4. Did you review your classmate' bioassay research at your school?	No	No	Yes	Went around reviewing posters in classroom – what's wrong with this one? With that one?	No	Yes
4a. If yes, how was that similar or different from the experience today?			More refined than what is going on today. Looked over other students' results, read over students' research to try to replicate results, class discussions, brainstorming.	In class, used hard copy peer review form. Later we'll put it on the website. We tried EI website but couldn't save it.	Will do next week with other students' posters.	
5. In doing the bioassay project did your group ever disagree? About what? How did you work it out?	Agreed by consensus		Sometimes disagreed about predicting toxic concentrations. Didn't disagree much. "Had lots of fun doing this."	Helping each other out, personality differences we had to deal with.	Disagreed on conclusions. People told others what to do, how they would do the tests. Projects were better due to disagreements. Some groups disagreed, others didn't.	"Not that I know of." "We brainstormed together." "Not much to disagree on when there are only 3 of us (in the class).

6. How was the bioassay project different from other projects you’ve done in science class?	Groups designed own experiments. Students knew process, but developed a better idea of what they meant. “We made our own conclusions, no one told us any of it.” “It was applied.” “We saw the bigger picture.”	So much research to look at; open-ended.	“It was more fun.” Discussed methods Seemed more “college-like” “Don’t just have teacher telling us what to do.” “There was less hand-holding.” “We did the entire project.” “We made our own connections.” “There was a reason behind this.”	This is hands-on, not paper work, not note-taking. Don’t like being lectured at. I can’t take it all in. Hate note-taking. Need demonstration.	More freedom, harder, more by yourself, on your own, more complicated, used computers more, got to pick own things, could choose more things to test this year, got to know people better.	“More detailed.” “More creative.” “It isn’t ‘Here’s a lab, do it.’ It’s your individual project.”
7. We assume that doing the bioassay took some time. Was this a good use of your time?	“Didn’t take that long.” “Poster was most intense part.” Lot of after-school work.	Loved it, more interested in it now. “Involved my whole life.”	Yes! For the first time, what I was doing outside has to do with chemistry. “Let’s do this project to see if this question we have had about the environmental problems of this foam is legitimate.” “Maybe we can have an impact as young students.” “This let us really test what we’ve been talking about for years.”	Showed us how to put things together and how to present. Learned communicating, talking with group members. Had to rush on it.	Yes.	Yeah. It was interesting, we learned a lot – it was fun too.

8. What parts of the bioassay project were disappointing?	Would have liked to do comparison of results to another substance. Small data set. Needed longer period of time for study.	The results. Disappointed in companies’ ethics with animal testing.	More time would have been helpful. Disappointing that they have been told that this foam is fine and might not be in the long run. Don’t want to stir up trouble in the community, but at least want to make them aware. Might be better options. “We wished we could have done more. We would like to take it further.”	Rough spots – writing the conclusion was hard. We don’t like having to write up everything we did: that’s hard. We also realized that we had messed up.	Needed more time. Frustrating putting it together. Many materials.	“My original lab didn’t come out right but I talked to these two and figured out errors.”
9. How did you feel about reviewing other people’s posters?	A lot more harsh evaluations than expected. Easier to evaluate similar projects, but unfair to evaluate those closest to your own because you knew more than what was on the poster. “Who was I to judge his clarity?”	“We were a little intimidated that ours wouldn’t look as complete and professional as others.”	Showed a lot of things. “We can help them and they can help us.” You could tell the differences in the materials available to the schools. Showed us how we can help each other, and changes we can make in our own.	Shows us how to correct things. I don’t like doing it because I don’t like disappointing people.	Got other ideas for other projects. Liked how students can review the work.	“I tried to be nice about it.” “Held back a little.”
9a. Could you understand what the group had done?	A huge obstacle to understanding – technical explanation of procedure (e.g. Used molecular formula, not name.) Some groups had unclear intentions. Presenter was not always there to answer questions.		Were able to ask questions and get answers.	Perspective on flow might be different. Some were mind boggling in terms of flow – e.g. conclusion 1st, then hypothesis. Use arrows?	Yes, when the person explained it. Others in their school had done bioassays last year.	“I think some of the stuff we learned help us understand theirs.”

9b. Did you feel qualified to assess other students' work?	No, in that we didn't actually do any of the work. Yes, in the sense you can get a basic understanding just by reading.	"I didn't like being so critical… they know what they did."	Felt as though they had some right to give input. Were able to see differences in teachers – compared Congress to teacher conference. "We can talk about our rationale and learn from each other like teachers do at the big conference."	Yes, give their opinion. Benefit.	Yes, definitely. Scared, shy, not comfortable asking. Felt like didn't know what to evaluate.	"Yeah, I feel like we learned a lot about it." "We are all on an equal level."
9c. Do you feel that you could assess their work fairly?	No. "I felt bad giving bad grades." "That's just mean." "It was fair in that a standard was set for our own." "Not clear understanding." "No scale."		Yes – they were all great and format was similar, so it made reviewing easier.	Fair, yes.	Yes. It was fair since these people didn't know them.	
10. How did you feel when other students were reviewing your poster?	Nervous. Why? "It was my work out there!" Different if prizes? "No!"		Kind of nervous. Open to questions to enable a good assessment. Why were you able to work better with other students than with adults? – "More comfortable because they were going through the same process."	Nervous. I wasn't nervous because I saw theirs' and ours were as good.	Nervous. Ready.	"I didn't really like it." "I don't take criticism too well."
10a. Were you able to answer the questions that people asked?	Adults asked, not a lot of kids. "Some adults were nicely inquisitive." "I'd have asked more if I was forced to do so."		Easy to explain because they had a good amount of background.	Nervous.	Yes. Some trouble with hypothesis.	Weren't around our poster because we were peer reviewing others.

11. Do you think it would be fair to use the assessments of your peers in determining a grade for you on this project? Why/why not?	No. Each person has their own view of what's good. You need an objective point of view. "I might be biased." Is the teacher objective? → "I hope so!"	Should use both peers and grownups. "Adults and people with experience know where they're coming from."	In some cases – might have biases or might not be doing it completely.	Yes, we can grade each other.	Yes, they trust that people will fairly judge them, but not if there were different types of projects.	No. If people were more experienced – some people didn't get into it.
12. Did you learn anything today that would help you to create a better bioassay research experiment in the future? What?	Yes. We should vary individual projects (have different students use different toxins). But, compiling data was good.	"Frustrating – so precise… our experiment led to so many questions. I got stressed out, mad at people." "Hard to figure out how to put it together." Important to "stick to what your specific question is."	Presentation: a lot more graphs, colors. Experimental design: use other animals, other concentrations, look at other variables. Presentation and explanation is important.	Don't use a particular shampoo. Let's try Daphnia instead of brine shrimp. (science content)	Change it, test different chemicals next year (idea from a professor).	"I would have repeated the experiment." "It would have been nice to use additional species rather than just using lettuce seeds." Time to look up more chemicals – do more research.
13. During this project, did you learn about aspects of doing science that you did not know previously?	"Absolutely." "I learned a lot about bioassays." "You really learned to use the scientific method."	Teacher: "Process is more important than outcome. The girls' process was just beautiful." Students: "A lot more space to learn because it was not as structured." "In this, you feel like you're actually doing something" (unlike regular chem. labs) "Outcome was totally unexpected."	"Science can be fun." Learned organizational patterns of science. Everything has to be able to be replicated.	It takes a lot of different things (methods) to find out something interesting. Cool to find out new stuff. I like doing it (science) because it's fun.	Yes, e.g. process of elimination on evaluation form, give an overall grade (more about peer review).	"I learned you kinda look at people who have done projects before." "You build on other people." "The creative part is cool." "You have your own choice."

Other Comments	“Did anyone else feel bad about killing Daphnia?” “We really wanted to do more with the Daphnia.” “More dilutions would have allowed for my surviving Daphnia.” “What I learned – posters!” What would be useful to leave behind for next year? Diagrams of serial dilution. Another student disagreed, “They should develop it on their own.”		Community involved: fire chief, relatives, teachers, other fire officials. Want to take the project further – don’t want to stop. Want to use other organisms for bioassays and compare with other fire companies throughout the area.. Age – people in the community paid attention even though they were teens. Talked to friends and family about project and they thought this was interesting and were able to learn a lot about the foam’s history through personal interviews. Took a good amount of work and thought. Really took “Chemistry in the Community” to heart. Proud. One question leads to another, builds on previous questions (excited). Might make web site describing research.	“I had never done something like this in my whole life.”	They imagine that scientists always have some disagreements. May help you do science better. “What would you do differently next time?” - change the elements - work harder If scientists had research results, what would they do with them? - show them to other scientists - show them to someone “higher” - redo the experiment - tell on news, Internet, shouting, magazines, handing in a report to a “higher “ person Interviewer told them about journals. They did not know that.	“Make sure you have ample amount of time to get it done.” “I’m looking into biotechnology” – future plans. “I was thinking about conservation.”

Bibliography

Bianchini, J. A. (1997). Where knowledge construction, equity, and context intersect: Student learning of science in small groups. Journal of Research in Science Teaching, 34, 1039-1065.

Billington, H. L. (1997). Poster presentations and peer assessment: Novel forms of evaluation and assessment. Journal of Biological Education, 31(3), 218-220.

Boon, P. I., & Brock, M. A. (1994). Plants and processes in wetlands: A Background. Australian Journal of Marine and Freshwater Research, 45(8), 1369-1374.

Bos, N., Krajcik, J., & Soloway, E. (1997). Student publishing in a WWW digital library--goals and instructional support. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Burnham, J. C. (1990). The evolution of editorial peer review. Jama, 263(10), 1323-1329.

Callaham, M. L., Baxt, W. G., Waeckerle, J. F., & Wears, R. L. (1998). Reliability of editors' subjective quality ratings of peer reviews of manuscripts. Jama, 280(3), 229-231.

Campanario, J. M. (1996). Have referees rejected some of the most-cited articles of all times? Journal of the American Society for Information Science, 47, 302-310.

Chalmers, A. F. (1982). What is this thing called science? (2nd ed.). Milton Keynes, England: Open University Press.

Chubin, D. E. (1990). Scientific practice and the contemporary politics of knowledge. In S. E. Cozzens & T. F. Gieryn (Eds.), Theories of science in society (pp. 144-163). Bloomington, IN: Indiana University Press.

Cunningham, C. M., & Helms, J. V. (1998). Sociology of science as a means to a more authentic, inclusive science education. Journal of Research in Science Teaching, 35(5), 483-500.

Cunningham, C. M., Meyer, D. Z., & Avery, L. (March, 2001). Promoting sociologically authentic inquiry in school science communities. Paper presented at the annual meeting of the National Association for Research in Science Teaching, St. Louis, MO.

Deorani, S. C., & Dabral, S. L. (1997). Forestry grey literature in India and its management. Indian Forester, 123(7), 607-614.

Dickson, D. (1988). The new politics of science. Chicago: University of Chicago Press.

Garfunkel, J. M., Lawson, E. E., Hamrick, H. J., & Ulshen, M. H. (1990). Effect of acceptance or rejection on the author's evaluation of peer review of medical manuscripts. Jama, 263(10), 1376-1378.

Gieryn, T. F. (1999). Cultural boundaries of science: Credibility on the line. Chicago: University of Chicago Press.

Gieryn, T. F., & Figert, S. E. (1990). Ingredients for a theory of science in society: O-rings, ice water, C-clamp, Richard Feynman, and the press. In S. E. Cozzens & T. F. Gieryn (Eds.), Theories of science in society (pp. 67-97). Bloomington, IN: Indiana University Press.

Godlee, F., Gale, C. R., & Martyn, C. N. (1998). Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. Jama, 280(3), 237-240.

Gratz, R. K. (1990). Improving lab report quality by model analysis, peer review, and revision. Journal of College Science Teaching, 19, 292-295.

Hargens, L. L. (1990). Variation in journal peer review systems. Possible causes and consequences. Jama, 263(10), 1348-1352.

Horrobin, D. F. (1990). The philosophical basis of peer review and the suppression of innovation. Jama, 263(10), 1438-1441.

Hull, D. L. (1988). Science as a process: An evolutionary account of the social and conceptual development of science. Chicago: University of Chicago Press.

Jaeger, R. G., & Toft, C. A. (1998). Writing for scientific journals II: The review process. Herpetologica, 54(Suppl), S54-S63.

Kelly, G. J., Carlsen, W. S., & Cunningham, C. M. (1993). Science education in sociocultural context: Perspectives from the sociology of science. Science Education, 77, 207-220.

Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. International Journal of Science Education, 20, 849-871.

Koprowski, J. (1997). Sharpening the craft of scientific writing. Journal of College Science Teaching, 27(2), 133-135.

Kronick, D. A. (1990). Peer review in 18th-century scientific journalism. Jama, 263(10), 1321-1322.

Lacanilao, F. (1997). Continuing problems with gray literature. Environmental Biology of Fishes, 49(1), 1-5.

Latour, B., & Woolgar, S. (1986). Laboratory life: The construction of scientific facts. Princeton, NJ: Princeton University Press.

Lightfoot, J. T. (1998). A different method of teaching peer review systems. Advances in Physiology Education, 19(1), S57-S61.

Longino, H. E. (1990). Science as social knowledge: Values and objectivity in scientific inquiry. Princeton, NJ: Princeton University Press.

MacLeod, L. (1999). Computer-aided peer review of writing. Business Communication Quarterly, 62(3), 87-94.

McNutt, R. A., Evans, A. T., Fletcher, R. H., & Fletcher, S. W. (1990). The effects of blinding on the quality of peer review. A randomized trial. Jama, 263(10), 1371-1376.

Merton, R. K. (1973). The sociology of science: Theoretical and empirical investigations. Chicago: University of Chicago Press.

Millar, R. (1989). Bending the evidence: The relationship between theory and experiment in science education, Doing science: Images of science in science education (pp. 38-61). London: Falmer Press.

Mulkay, M. (1975). Norms and ideology in science. Social Science Information, 4-5, 637-656.

Mulkay, M. (1991). Sociology of science: A sociological pilgrimage. Bloomington, IN: Indiana University Press.

Pechenik, J. A., & Tashiro, J. S. (1991). Instant animals & conceptual loops: Teaching experimental design, data analysis & scientific writing. The American Biology Teacher, 53, 220-228.

Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. Review of Educational Research, 68, 35-59.

Rushton, C., Ramsey, P., & Rada, R. (1993). Peer assessment in a collaborative hypermedia environment: A case study. Journal of Computer-Based Instruction, 20(3), 75-80.

Topping, K. (1998). Peer assessment between students in colleges and universities. Review of Educational Research, 68(3), 249-276.

Towns, M. H., Marden, K. L., Sauder, D., Stout, R., Long, G., Vaksman, M. A., Kahlow, M., & Zielinski, T. J. (1999). Interinstitutional peer review on the Internet: Generating a professional-like atmosphere for writing in the discipline (Unpublished manuscript). Muncie, IN: Ball State University.

Trautmann, N. M., Carlsen, W. S., Cunningham, C. M., & Krasny, M. E. (in press). How toxic is it? Measuring the effects of chemicals on living things. Arlington, VA: National Science Teachers Association.

van Rooyen, S., Godlee, F., Evans, S., Smith, R., & Black, N. (1998). Effect of blinding and unmasking on the quality of peer review: a randomized trial. Jama, 280(3), 234-237.

Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. Cambridge: Cambridge University Press.

Ziman, J. (1994). Prometheus bound: Science in a dynamic steady state. Cambridge: Cambridge University Press.