ISSN 1430-4171
http://journals.springer-ny.com/chedr
10.1007/s00897980163a

Of Special Interest

Report On "Action Research: Overcoming the Sports Mentality Approach to Assessment/Evaluation"

NANCY S. GOROFF
Department of Chemistry
State University of New York at Stony Brook
Stony Brook, NY 11794
ngoroff@notes.cc.sunysb.edu

Don't confuse precision with accuracy, and don't assume that quantitative assessment methods will be more accurate in helping you make decisions about instruction.

In this workshop, George Bodner contrasted evaluation with assessment, and presented *Action Research* as a method for evaluating curriculum changes. He emphasized that any major change in teaching will have some effect; if test scores or other analyses do not reflect this effect, then one must use other tools to uncover it. Bodner defined Action Research as qualitative information-gathering and analysis from all available sources, including reflective journals, comments from students, and personal interviews. In the past few decades, such qualitative research has given way to more quantitative studies, such as analysis of test scores. Bodner argued in his presentation that these quantitative

"Action Research: Overcoming the Sports Mentality Approach to Assessment/Evaluation" by George Bodner was presented at the "Day 2 to 40" workshop symposium held May 10–11, 1997. The two-day event was held in the Willard H. Dow Chemical Sciences laboratory building on the central campus of The University of Michigan in Ann Arbor, Michigan. Each of the articles that comprise this issue was written by one of the group of reporters whom I asked to attend each session to take field notes and then follow up with the session leader and participants afterwards.

—Brian P. Coppola, Proceedings Editor

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

measures are more precise, but less accurate, than Action Research's subjective, informal methods of evaluation.

Descriptive Outline

This workshop consisted of a lecture-style presentation by Bodner with occasional discussion or questions from participants. Bodner used transparencies and wrote some additional information on the board. At the end, participants received a handout of the information in the transparencies.

Report

Don't confuse precision with accuracy, and don't assume that quantitative assessment methods will be more accurate in helping you make decisions about instruction. Those were the major points of George Bodner's presentation about how to evaluate changes in a curriculum. Bodner stressed that *assessment*, "the process by which an instructor (or researcher) measures the performance of individual students or groups of students," is only one technique available for *evaluation*, "the process by which information is collected in order to make decisions on how instruction can or should be improved." Assessment generally comes in the form of quantitative exams, designed for purposes other than reviewing the curriculum (i.e., for assigning grades). Instead of relying on typical assessment methods when deciding on curriculum issues, Bodner recommended *Action Research*, a qualitative and iterative approach to evaluating instruction and making changes to it.

Quantitative assessment techniques are popular in education research because they are easily available and precise. Nonetheless, Bodner cautioned workshop participants to stay away from the "sports-mentality trap" of relying too heavily on scores: At the end of the game, no matter the margin of victory, "there is no doubt about who wins when 'Da Bulls' play the Pistons. The same cannot be said about the results of studies of student achievement." Specifically, Bodner warned against three mistakes that can result from the "sports-mentality approach" to improving instruction.

1. Researchers assume that *assessment* and *evaluation* are synonyms, and therefore "that decisions on the efficacy of a program can be based solely on student achievement;"

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

2. the assessment data fail to address the issues of concern to the researcher; and

3. preliminary experiments lead to incorrect conclusions because student bodies vary from semester to semester.

Comparing students' results on standardized exams, the conventional method for evaluating the success of new changes, gives precise, and even statistically significant, results, but they may not have any value, according to Bodner. As an example, he cited two mythical sets of data representing test scores of students, manufactured so that the mean of the second set was significantly higher than the first; however, 20 out of 24 students in the second group had lower scores than in the first, while only four students had improved. Using the mean score as a simple metric of student success removes any value judgment from the process, and this is to the detriment of the research, Bodner said.

The key issues Bodner would like to consider in evaluating a course are: What is learned? What is retained? and Which knowledge we value is gained? (vs. Which knowledge is easily tested?). Change should be targeted, not random, and therefore evaluation of the change needs to take into consideration the goals that prompted the change in the first place. In particular, Bodner suggested that educators focus on retention of material as a key, an often ignored aspect of success in the classroom. To address the issue of retention, researchers need to include some post-class assessment in their evaluation.

Bodner also argued, crediting Steve Kemmis for the idea, that educators should not make a change in curriculum and then "see if it has an effect," because *all* changes will have both positive and negative effects. "If the evaluation shows no effect, then you have done the wrong evaluation." Bodner suggested that change should be iterative: "Once you've done something, find the people who benefit. Find those who don't." Make improvements by maximizing the positive effects and minimizing the negatives, through a repeated process of change and evaluation. This iterative effort is Bodner's Action Research. Still, Bodner cautioned, "[the methodology of] Action Research doesn't tell you what is good;" to determine that, the instructor and students need to work together and make a subjective evaluation.

Bodner defined Action Research as "an informal, qualitative, formative, subjective, interpretive, reflective, and experiential model of inquiry in which all individuals

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

involved in the study are knowing and contributing participants." Breaking that hefty definition down into parts, he explained that the research is *informal* in that there is no second observer, external to the classroom. Because the research is *qualitative*, there are "no data which are not worth collecting." It is *formative*, occurring while the course is in progress, and involving everyone; Bodner eschewed the practice of changing the curriculum for some students but not all, to establish a "control" group. Action Research is *subjective* and *interpretive*, relying on personal evaluation rather than statistical tests. It is *reflective* and *experiential*, and everyone involved with the class is asked to participate. Including students explicitly and openly in the process increases the amount of information gathered.

Bodner laid out the four stages of the Action Research cycle as follows:

- Reconnaissance and general plan: recognize the problem and get an idea of how to solve it.
- The action: carry out your idea.
- Monitoring the implementation: collect observations in various forms, including reflective journals, comments from students, and personal interviews.
- The revised plan or maintaining the action: look at the data; develop a new strategy; and repeat the process "until you have achieved 90% of optimum." At that point, Bodner explained, you have achieved as much as you can.

Bodner finished off his presentation with a few words on assessment, citing several papers from *Assessment in the Service of Instruction* [1].

Adoptive Participation

Bodner urged workshop participants to use Action Research as a tool for evaluating any changes they may try in their own curricula. He presented the example of one professor at Purdue, Mark Loudon, who changed his style of teaching and evaluated the change using the recommended methods. Loudon wanted to increase his own enjoyment in the classroom and increase retention of material among his students. To accomplish these goals, he decided to have students work together in groups of three, both outside the classroom and in class. He also allowed them to discuss the hour

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

exams in groups of three at the beginning of the test period before writing down their answers to the exam questions.

Bodner read a long list of questions considered by Loudon and his students during their Action Research investigation. He also used this example to point out the dangers of depending too much on early results; Loudon had very different responses from students in the first two terms in which he made the changes to his course.

- How do students interact when they work together?
- What do the students choose to talk about?
- What do the students discuss when they preview exams?
- What are the students' perceptions of their instructor and each other?
- What are the instructor's and the students' beliefs about how learning occurs?
- What do the students believe is the role of the instructor in organic chemistry?
- What is the nature of the dissatisfaction that led the instructor to change the instructional situation?
- What factors make it difficult to change the instructional situation?
- What factors interfere with the ease with which this technique can be used by other instructors, or transported to different institutions?
- What effect does this mode of instruction have on the instructor's attitude toward teaching? Is this approach to organic chemistry more fun to teach?
- What effect does this approach have on students' perception of the difficulty of organic chemistry?
- What do students understand about organic chemistry?
- Does this change when we alter the approach taken to instruction?

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

- Does the new instructional mode produce students who think more like an organic chemist?
- Does it produce students who understand what organic chemists do?

Feedback

Expectations

"As with this one, many workshops turn out to be mini-plenaries, with the 'sage on the stage' sharing his profound insights with a passive audience. His insights were wonderful, and he was entertaining."

"This was not really a workshop. Except for a few responses to questions, it consisted of a lecture by George."

"The session was not a workshop but primarily a lecture in which the class could ask questions. Fortunately, I found Bodner to be a fount of information."

"It was more of a presentation than a workshop. George tried with a couple questions to open it to the group, but to me a 'workshop' is interactive where everyone gets a chance to participate in a hands-on or at least oral way. The discussion was interesting, but primarily between only a few audience members."

"In a workshop I imagined that the participants would accomplish some sort of directed task, and then analyze the results with the help of the workshop leader. In reality, most of the sessions I attended were not really workshops, but were more like informal lectures. People in the audience were more prone to speak freely because of the 'workshop' title, but I didn't feel that I participated in 'workshops.' This is not bad. I like lectures because of the tremendous amount of information which can be communicated. Whether intentional or not, I think the workshop title was a very effective way to bring about informal, participant-active lectures, which I prefer to actual workshops in most cases."

"I don't think this was a workshop. It was basically a long presentation by George with the opportunity to ask questions, discuss with him and other attendees the points he made, etc.… On the other hand, I came away from this 'workshop' with a generally positive opinion and would give it a B grade."

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

“I expected the role of the workshop leader to be more of a facilitator, moderating a discussion with the other participants after setting some type of framework in the introduction.”

“It seems strange to talk about making change work when the format for discussing change is the traditional lecture format.”

“Personally, I was looking for a seminar format and that is what was given. I wanted to know the professional opinion of the speaker on the topic of assessment/evaluation and I ‘got my money’s worth.’”

Things to Take Home:
“I was intrigued by his admonition that, if you have made a substantial change in your teaching, but you see no change in the student learning, then change your metric. I had blithely assumed that many changes lead to ‘null’ results because they are not causative.”

“The idea of action research was useful. It is similar to some of the evaluation tools we use..., but the idea of making students more part of it was particularly helpful.”

“The most useful [thing] I learned was that classes of students can vary considerably; groups of students can respond and react quite differently from one term to another. (I needed to hear that since in the last two times I taught a particular course the student reaction and my evaluation were very different for the two terms.)”

“I’ll definitely carry back George’s story of the two students who changed sections to get out of the group study, then begged their new instructor to incorporate it. Part of my job is ‘selling’ faculty and students on collaborative learning techniques, and there is always the problem of how it will benefit the academic overachievers. This was a nice story that should help that problem.”

“One key point that was made is that we usually fail to measure retention of material by our students. I have been aware of the learn/write exam/forget mentality of students and give a comprehensive final, but I hadn’t considered actually building later tests of retention into the curriculum. I will have to see how receptive my colleagues will be to periodic retesting of students.”

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

"One that I have already tried is the pre-exam cooperative discussion of the exam by groups of 3 students.… We tried it for the Final Exam in Instrumental Analysis (12 Students, 4 groups of three, 20 minutes). The students found that their exam anxiety was reduced greatly or even eliminated. One student noted that her exam anxiety was gone and that, whatever her grade, she couldn't blame her performance on anxiety but would accept whatever she earned as a reflection of what she knew. My general observation is that student performance did not change much (the sample is too small, probably, to do any statistical analysis)."

"Perhaps the most meaningful take-home lesson that I gained was that I can convert many of the 'informal' methods of assessment (that I am using for my own purposes) into formal 'assessment instruments' to provide as 'cold hard data' for the 'bean counters.' Also, from the workshop (and discussions with several participants afterward) I was able to better understand why it is not possible to present 'cold hard data' for the intangibles which are the real changes that result from 'reform.'"

"The overall ideas on assessment and evaluation will be useful for project on women in science that I am working on with other members of the ... faculty. We need to look closely at what we are trying to accomplish and make sure we design assessment tools to answer the questions we feel are important to the success of the project."

"The idea of having students preview exams is something that I am thinking of doing in my biochemistry course. I already encourage the students to work together on homework which emphasizes applying what they have learned and integrating the material. This may be an effective way to continue to have the students apply these skills. It would allow me to ask more of these kinds of questions. I do worry about time constraints and am considering having a separate section that asks more straight-forward questions that the students will do at a different time."

Improvements:
"Because people learn and retain more when they have a chance to reflect and then rephrase the concepts in their own words and from their own experiential base, the session would have benefited from some pair-wise or small-group discussions."

"He could solicit more involvement by the attendees. He could bring more handouts. I did not receive a packet of the lecture notes (Bodner's transparencies), which was

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

unfortunate since when he stated that we would all receive copies I stopped taking extensive notes."

"Use overheads that are easier to read...bolder letters and lighter background. Provide some time for small discussion groups to react to the presentation. And don't overlook the importance of a break!"

"a. Provide sufficient hand-outs. b. Make it a workshop."

"Selecting a specific aspect/example of the difference between assessment and evaluation, then inviting the participants to discuss/brainstorm the distinction and its impact on the larger picture of chemical education and reform efforts would have dramatically improved the workshop."

"I think handing out the handout at the beginning and then spending more time following the topics of interest to the attendees would help. That way, rather than try to track a lecture, everyone could be more engaged in actual dialogue."

"I would have liked to have more examples of how to provide 'meaningful assessment' strategies for my classes...and an opportunity to share my experiences with others in a smaller group."

"More concrete examples of how others have applied his evaluation methods."

Workshop Participants

George Bodner (Leader; gmbodner@vm.cc.purdue.edu), Nancy Goroff (reporter; nancy.goroff@sunysb.edu), Richard Bertrand (rdbertrand@uccs.edu), Lisa Bell-Loncella (kintner+@pitt.edu), Bob Blake (rblake@post.cis.smu.edu), Charles Cornman (charles_cornman@ncsu.edu), Seyhan Ege (snege@umich.edu), Martha Gach (mgach@ umich.edu), Louise Liao (lwliao@facstaff.wisc.edu), Nadia Marano (nmar@music. stlawu.edu), Malinda Matney (mmatney@umich.edu), Deanna Mitchell (djmitchell@ nwu.edu), Harvey Nikkel (nikkelh@gvsu.edu), Jim Pennington (jdpennin@ umich.edu), Tracy Schoolcraft (tascho@ark.ship.edu), Ron Schroeder (rrs@chem.wayne.edu), Bill Sweeney (wsweeney@shiva.hunter.cuny.edu), Joel Tellinghausen (tellinjb@vuctrvax.vanderbilt.edu), Clif Thompson (clifton.thompson

ISSN 1430-4171
http://journals.springer-ny.com/chedr
S 1430-4171 (98) 01163-9

@ cmich.edu), Sarah Thompson (sarah.e.thompson@cmich.edu), Donald Wink (dwink@uic.edu).

REFERENCES

1. *Assessment in the Service of Instruction*; Champagne, A. B., Lovitts, B. E., Calinger, B. J., Eds.; American Association for the Advancement of Science: Washington, DC, 1992.