What's Wrong with Giving Students Feedback?

Judith E. Sims-Knight, Richard L. Upchurch
University of Massachusetts Dartmouth

Abstract
This paper reviewed the extensive evidence on the effectiveness of feedback on learning. The research supported five claims about feedback. First, informational feedback is effective in domains with clear right or wrong answers when tested immediately after training. Second, when the same maximal feedback conditions are tested for retention or transfer, they are less effective than conditions with less feedback. Third, feedback can draw attention away from the learning task. Fourth, feedback apparently plays a minor role in actual classroom situations. Fifth, teaching students to provide their own feedback and explanation is an effective alternative. These findings suggest that instructors may be more effective if they put less effort into grading and commenting on students' products and more effort into structuring their courses to help students learn how to assess and reflect on their state of learning themselves. Two specific pedagogical strategies are suggested. First, giving students more assignments than the instructor could grade or comment on will provide more of the kinds of practice they need to develop expertise. Second, helping students to learn how to assess and reflect on their state of learning will help them learn how to provide their own feedback and thus help them to become independent life-long learners.

I. Introduction

Conventional wisdom dictates that giving students feedback promotes learning. Instructors in all disciplines spend hours laboriously correcting and explaining errors on quizzes and tests, reports and papers. Their belief in the importance and effectiveness of feedback is supported by research in a variety of contexts. For example, Azevedo's meta-analysis of research on the effectiveness of feedback in computerized instruction found strong and consistent superiority of feedback conditions compared to nonfeedback[1]. In addition, a number of studies have found that elaborated feedback, in which students are helped to find the right path, are more effective than situations in which they are simply told whether they are right or wrong.[2,3,4,5]

Although there are strong and consistent findings that feedback improves immediate performance under some circumstances, it is also clear that in some situations feedback is irrelevant and sometimes even harmful. In a meta-analysis of research in educational, organizational, and laboratory settings, Kluger and DeLisi[3] found that in one-third of the comparisons the feedback condition had worse performance than the group who was given no feedback.

Because so much of instructors' time is spent giving students feedback by commenting, correcting, and grading student work, it is important to know if this labor fosters learning. Understanding the factors that produce the inconsistency in feedback research should guide

educational practice in two ways. It can indicate whether typical evaluation practices promote learning. If they do not, it can suggest ways in which educational practices can change.

The word feedback is used in many ways. In this paper we will use the following definitions. In informational feedback people are told whether their answers are right or wrong. Such informational feedback may be a simple right-wrong response (minimal feedback) or it may include an elaboration. Typically, the elaboration involves an explanation of errors or guidance toward the correct response, but it need not include a reward component. A reward is something of value to the individual, such as a grade or money, that follows a response. Praise is a particular kind of reward that is widely used in educational settings and has variable effects. A reward may or may not include informational feedback.

This paper will focus on the effects of feedback on performance, and, when possible, on performance in educational settings. It must be noted, however, that feedback has other effects. Feedback also affects intrinsic motivation, which is important in its own right. Intrinsic motivation encourages people to choose to put in the long, hard hours of work required to develop expertise. In addition, feedback elicits both emotion and arousal, which can affect learning.

Our review of the research has led us to make five claims about the relation between feedback and learning performance, which we will review in turn:

1. Informational feedback is effective in domains with clear right or wrong answers when tested immediately after training.
2. When the same maximal feedback conditions are tested for retention or transfer, they are less effective than conditions with less feedback.
3. Feedback can draw attention away from the learning task.
4. In studies of the role feedback plays in actual classroom situations, feedback is not a major variable.
5. Teaching students to provide their own feedback and explanation is an effective alternative.

II. Research Review

Claim 1. Informational feedback is effective in domains with clear right or wrong answers when tested immediately after training. A recent meta-analysis of 22 studies found strong support for the effectiveness of feedback in computer-based learning, although the size of the feedback effect varied with a number of factors.[1] It may be that computer delivery of feedback is especially effective. In another meta-analysis of studies that compared feedback to no-feedback conditions, Kluger and DeLisi[3] reported larger effect sizes for computer feedback than for noncomputer feedback.

Claim 2. When the same maximal feedback conditions are tested for retention or transfer, they are less effective than conditions with less feedback. When students are asked to recall what they have learned at a later time (retention) or are asked to use what they have learned to solve new problems (transfer), informational feedback after every response is not so effective. Students often retain and apply what they have learned more effectively with less feedback.[5-6]

Schooler and Anderson,[5] who compared delayed vs. immediate feedback in their LISP computer tutor, suggested two factors to explain this effect. First, students in their delayed feedback

conditions engaged in more self-correcting than did students in the immediate feedback conditions. To solve problems beyond the tutorial, students need to be able to self-correct. Second, the protocol analyses suggested that feedback may take up cognitive space and reduce problem-solving ability (see also Kanfer & Ackerman[7]).

An analogous finding was reported by Charney, Reder, and Kusbit[8] in a study of learning to use a computer spreadsheet. They compared a condition in which students had to figure out how to solve a problem and then were given feedback to a second condition in which the subjects were given the correct commands and had to copy them. Both were given the correct answer as feedback, but in the first condition it was delayed until the students had made a response. Thus these conditions are roughly analogous to Schooler and Anderson's delayed vs. immediate feedback.[5] The delayed feedback condition required longer training times, but resulted in better performance at test than the tutorial condition.

Schmidt and Bjork[6] reviewed evidence that compared getting feedback after every response (continuous feedback) to getting intermittent or summary feedback after a number of trials. In these studies continuous feedback resulted in superior immediate performance, but poorer retention and transfer than summary feedback.

Several studies have explored the efficacy of different kinds of feedback, usually minimal vs. some sort of explanation or guidance. Elaborated feedback often results in superior performance on immediate tests, but no differences on retention and transfer. This pattern of results was found by Anderson et al.[2] in studies of the effectiveness of their LISP intelligent computer tutor and by McKendree[9] with a comparable computer tutor for geometry. McKendree,[4] however, in a second study, found significant improvement in solving transfer problems by a simple change in the feedback condition. In the original geometry tutor[9] the computer allowed students to follow possibly correct, but nonoptimal proof paths, that is, they were allowed to make correct inferences even when such inferences did not lead directly to the proof. This, she surmised, allowed students to retain misconceptions. In the revision of the tutor students could not choose nonoptimal proof paths, which prevented them from choosing inferences for the wrong reasons. The improved elaborated feedback forced optimal reasoning and was more effective in both the immediate test and the transfer test than minimal feedback. It was also more effective than a condition that explained students' errors.

This body of research suggests several things. First, what is most effective in the short run is not necessarily the most effective in the long run. Applied to the classroom, this suggests that what promotes the best quiz performance may not promote the best final exam performance. Second, feedback can prevent students from engaging in the active processing that results in generalizable understanding. Delayed feedback and intermittent feedback force students to do this. Third, explanations that help guide the student toward correct reasoning are effective in helping students learn generalizable skills.

Claim 3. Feedback can draw attention away from the learning task. Students sometimes pay more attention to the feedback than to the task. Kluger and DeLisi[3] concluded from their meta-analysis that feedback becomes ineffective or deleterious if that feedback directs students' attention (and emotion) away from learning and toward ego issues, such as how smart they are, whether they can do better than others, whether they will succeed. The meta-analysis indicated that both praise and feedback that was designed to discourage eliminated the positive effect of

feedback. They argue that these seemingly opposite manipulations both pull recipients' attention away from the task and to ego needs.

Mueller and Dweck[10] have discovered that the kind of praise matters. They compared the effects of praising for intelligence to praising for effort. They found that praising students for intelligence resulted in reduced perseverance, lower performance, less task enjoyment, and more low-ability attributions than praising them for their effort. It is noteworthy that in both conditions students were not getting informational feedback, yet there are substantial effects on both performance and subsequent effort. Grading has been shown to have effects comparable to praising students for intelligence. Butler[11] found that grading students' performance led the students to focus on ego issues rather than on task characteristics, and to perform more poorly than students who received comments, but no grades.

The research described in this section suggests that grading and other kinds of evaluative feedback have a counterproductive aspect to them in that they lead students to focus on ego issues rather than on the task to be learned and because they reduce perseverance. One would expect that this disruption is more severe with difficult material[7] and when recovering from errors.

Claim 4. In studies of the role feedback plays in actual classroom situations, feedback is not a major variable. Three quite different meta-analyses have reviewed the effectiveness of feedback in classroom settings. Although the effect of feedback varied, all three meta-analyses found that other factors have a stronger influence on student performance. Harris and Rosenthal[12] reviewed 135 studies that investigated the effect of teachers' expectancies and other behaviors on the performance of their students. Whereas the amount of feedback students received was not related to performance, other variables, such as classroom climate, yielded strong relationships with performance.

Feldman,[13] in a meta-analysis on the effectiveness of course evaluations in large-section classes with common exams, found of 25 dimensions of teaching effectiveness, student's assessment of the nature, quality, and frequency of faculty feedback from their instructor was related to student achievement ($\underline{r}$ = .23), but it was relatively unimportant. Only 4 of the 24 dimensions studied had smaller correlations. Similar to Harris and Rosenthal's findings, the two strongest predictors were (i) instructor's preparation and organization of the course and (ii) instructor's clarity and understandableness.

Hillocks[14] in a meta-analysis of studies on the effectiveness of writing instruction found that the size and direction of feedback effects varied widely in different studies—from -.27 to +.82. He was able to identify some factors that influence feedback. First, when the instructors' comments had clear objectives (usually one or two), feedback was effective in improving writing (average effect size = .74); when they had no clear objectives, feedback was ineffective. When feedback was positive, students' performance at the end of the course was better than at the beginning (average effect size = .43); when comments were negative, post performance was lower (average effect size = -.20).

These three meta-analyses indicate that instructor feedback is not the potent instructional tool it is often supposed to be. In only two specific circumstances in writing instruction has it been shown to be effective—when the comments are focused on one or two objectives and when they

are positive. The latter might be akin to the finding of the effectiveness of elaborated feedback when it leads students to right answers.

Claim 5. Teaching students to provide their own feedback and explanation is an effective alternative. The research just described not only disabuses us of the belief that the more feedback the better, but it suggests why. When students can rely on an external authority for determining whether they are right or wrong, they don't need to develop such skills on their own. Luckily, research exists that demonstrates the power of students developing such skills, which are often called metacognitive. A series of studies on self-explanations nicely exemplifies this phenomenon. Chi et al.[15] first discovered that, when studying physics, good students (i. e., those who solved more physics problems) were more likely than poor students to explain to themselves what they understood and what they didn't. It has subsequently been demonstrated experimentally that students can be induced to self-explain and that so doing improves their learning.[16 17] Further, Van Lehn and colleagues[18] demonstrated that most of what students do when they self-explain is to fill in gaps in their understanding. Thus, fundamental to the self-explanation process is figuring out what one understands (has right), what one does not understand, and what one misunderstands.

A similar phenomenon has been demonstrated by requiring students to focus on their problem-solving process.[19] Students in the metacognitive processing groups were asked questions that focused their attention on their own learning strategies and the consequences of those learning processes (e. g., "What are you going to do next?" "What do you think will happen?" "Why do you think that happened?") while they solved problems like the Tower of Hanoi problem. Control groups were asked to think aloud or work silently, but were not asked specific questions. Students in the metacognitive processes group were better able to solve transfer problems than the control groups. Furthermore, negative self-evaluative statements (as captured in thinking-aloud protocols) decreased markedly for the metacognitive processes group but not for the control groups.

In both these examples it was found that many college students do not spontaneously use these metacognitive strategies,[15, 19] but that they can easily be taught to do so. Furthermore, developing skill at self-explanation and metacognitive processes seems a crucial component of life-long learning.

III. Pedagogical Implications

What does this research suggest for the classroom? In a sentence, it suggests that less emphasis on grading and providing individual feedback on reports, exams, etc. should improve instructional effectiveness. Reducing grading has four advantages. First, it frees up instructor time to do something more productive. Second, it reduces instructors' focus on deciding who are the good students and who are not. Third, it reduces students' focus on whether they are good (or the best) students. Fourth, it reduces the adversarial relationship between instructor, the giver of the reward (or punishment) of grades, and students, the receivers.

If faculty and students are less focused on evaluation, they can be more focused on learning. With the time freed from grading, instructors can institute practices that will more effectively help students learn. Here are two suggestions. First, faculty can give more of the kind of assignments that are difficult to grade. Difficult-to-grade assignments are often those that are

more authentic, such as real designs and written reports. Once instructors decide not to grade or comment on every assignment, they can determine the number and form of assignments based on what the students need or can do in the time available. Students need extensive practice to acquire authentic skills. For example, in a class in which students were required to write two reports per week, and the instructor gave no individual comments for most of the assignments, every student commented that one of the best things about the course was that they learned to write better.

When it is suggested to faculty that they give a higher proportion of ungraded assignments, they typically worry that students will object. It has been our experience[20] that students accept the responsibility for their own learning and like it. They do experience a "culture shock," at the beginning, and want feedback on everything, but they come to see the value of the approach as the semester progresses.

A second way faculty can institute pedagogical strategies that help students learn is to structure learning environments to help students learn to provide their own feedback. The process was exemplified in the data structures course described in Upchurch & Sims-Knight.[21] A goal of the course was to help students learn how to "read" programming code and detect errors, a skill that does not come automatically with the ability to write code.[22-24] The students were organized into teams. One student wrote the program and the other students reviewed it and organized a review meeting. Then the author had to respond in writing to the reviewers. This was repeated three times during the semester. Thus, students practiced providing feedback about a computer program from reading it. In addition, this procedure was embedded in a continuous improvement model to provide feedback on the students' review processes. The students kept track of how many errors they found and how many the authors corrected. After the debugging process, the errors the team missed would be found. Then, as a class, they would reflect upon the reasons the errors were missed and what they could do to improve their review behaviors the next time. Thus, students were providing their own feedback in a context in which they could assess how well they were accomplishing their goals and could figure out how to improve the process by which they wrote programs.

IV. Conclusions

The extensive research on the effectiveness of feedback reveals that providing feedback to learners is often very effective in promoting learning, but that in a significant proportion of instances it is ineffective or even deleterious. In addition, it has side effects that prevent students from focusing on learning in the most effective manner. Finally, when the role of feedback in classroom situations has been assessed, it has been found to be less important than issues of how instructors structure their course. Thus, it appears warranted for instructors to spend less time giving assignments to be graded or commented upon and more time structuring courses in other ways. Two ways seem particularly germane. First, giving students assignments that the instructor neither grades nor provides individual feedback will provide more of the kinds of practice they need to develop expertise. Second, helping students to learn how to assess and reflect on their state of learning will allow them to become independent life-long learners.

Bibliography

1. Azevedo, R., & Bernard, R. M. A meta-analysis of the effects of feedback in computer-based instruction. *Journal of Educational Computing Research*, 13(2): 111-127 (1995).

2. Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4(2), 167-207 (1989).

3. Kluger, A. N., & DeLisi, A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284 (1996).

4. McKendree, J. Effective feedback content for tutoring complex skills. *Human-Computer Interaction*, 5, 381-413 (1990).

5. Schooler, L. & Anderson, J. R. The disruptive potential of immediate feedback. *Proceedings of the Cognitive Science Society* (1990).

6. Schmidt, R.A. & Bjork, R.A. New Conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3, (4)*, 1992, 207-217.

7. Kanfer, R. & Ackerman, P. L. Motivation and cognitive abilities: An intergrative/aptitude-treatment interaction approach to skill acquisition, *Journal of Applied Psychology*, 74(4), 657-690 (1989).

8. Charney, D. Reder, L. & Kusbit, G. R. Goal setting and procedure selection in acquiring computer skills: A comparsion of tutorials, problem solving, and learning exploration, *Cognition and Instruction*, 7(4), 323-342 (1990).

9. McKendree, J. E. *Impact of feedback content during complex skill acquisition.* Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh, PA (1986).

10. Mueller,C. M, & Dweck,C. S. Praise for intelligence can undermine children's motivation and performance. Journal of Personality and Social Psychology, 75(1), 1998, 33-52

11. Butler, R. Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance, *Journal of Educational Psychology*, 79, 474-482 (1987).

12. Harris, M. J., & Rosenthal, R. Mediation of interpersonl expectancies effects: 31 meta-analyses. Psychological Bulletin, 1985, 97, 363-386.

13. Feldman, K. A. The association between student ratings of specific instrucitonal dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. Research in Higher Education, 1989, 30(6), 583-645.

14. Hillocks, G., Jr. (1986). *Research on written composition: New directions for teaching*. Urbana, Ill.: (ERIC Clearinghouse reading and Communication Skills.

15. Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems. *Cognitive Science*, 13, 145-182 (1989).

16. Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. Eliciting Self-Explanations Improves Understanding. *Cognitive Science*, 18, 439-477 (1994).

17. Bielaczyc, K., Pirolli, P. L., Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. Cognition and Instruction, 13(2), 221-252.

18. VanLehn, K. Cognitive Skill Acquisition. *Annual Review of Psychology*, 47, 513-539 (1996).

19. Berardi-Coletta, B., Buyer, L. S., Dominowski, R. L., & Rellinger, E. R. Metacognition and Problem Solving: A Process-Oriented Approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1), 205-223 (1995).

20. Upchurch, R., & Sims-Knight, J. "Reflective Essays in Software Engineering," Frontiers in Education Conference, San Juan, Puerto Rico, November 10-13 (1999).

21. Upchurch, R., & Sims-Knight, J. E. "In Support of Student Process Improvement," Proceedings of CSEE&T'98, Atlanta, Georgia, February 22-25, 1998.

22. Boehm-Davis, D. A. (1988) Software Comprehension. In M. Helander (ed.), Handbook of Human-Computer Interaction. Elsevier Science Publishers. p. 107-121.

23. Corbi, T. A. (1989) Program Understanding: Challenge for the 1990's. IBM Systems J. 28, 2, 294-306.

24. Deimel, L. E., and D. V. Moffat. (1982) A More Analytical Approach to Teaching the Introductory Programming Course. Proc. Natl. Educational Computing Conf. 1982. Columbia, Mo.: University of Missouri, 114-118.

JUDITH E. SIMS-KNIGHT
Judith E. Sims-Knight is Chancellor Professor of Psychology at University of Massachusetts Dartmouth. She also serves as Faculty Assessment Coordinator for the College of Engineering at UMASS Dartmouth, under the auspices of the NSF-sponsored Foundation Coalition, a consortium of six engineering schools dedicated to reforming engineering education. She has collaborated with Professor Upchurch on several NSF-supported projects exploring innovative ways of teaching software design.

RICHARD L. UPCHURCH
Richard L. Upchurch is a Professor of Computer and Information Science at the University of Massachusetts Dartmouth. He is currently working with the assessment team of the College of Engineering, under the auspices of the NSF-sponsored Foundation Coalition, in developing software support for assessment and reporting. He and Dr. Sims-Knight have collaborated on many occasions over the past fifteen years.