STUDENT AND FACULTY PERCEPTIONS OF EFFECTS OF MIDCOURSE EVALUATION

Whitney Ransom McGowan, Russell T. Osguthorpe,
Brigham Young University

We report on faculty and student perceptions of the effects of midcourse evaluations on teaching improvement and student learning. We provided faculty with a midcourse evaluation tool, surveyed faculty and students, interviewed faculty, observed debriefing sessions, and compared midcourse with end-of-semester ratings. Of 510 mean ratings on individual learning items, 342 (67 percent) mean scores showed improvement from midcourse to the end of the semester. Faculty who read their midcourse feedback, discussed it with their students, and made pedagogical changes saw the most improvement in their ratings.

No instructor wants to be a bad teacher. Some teachers may take great joy in being considered “hard or demanding, but never bad” (Phillips, 2001, p. iv). Rather, individuals who become teachers generally want their students to have “significant learning experiences, grow, and progress” (Fink, 2003, p. 6). Nevertheless, many faculty struggle with their teaching performance. Some give up on themselves, concluding they are not effective teachers.

Faculty improvement is essential for a variety of reasons. First, faculty who improve their teaching tend to experience increased teaching satisfaction. Second, faculty who do not strive to improve in their teaching are less likely to succeed in motivating their students to achieve additional improvement (Trigwell & Prosser, 2004). Good teaching demands that faculty continue to learn and improve.

Most universities offer a variety of services to help faculty improve their teaching, but many of these services are labor intensive and reach only a small minority of those who could benefit most from the intervention.

Teaching improvement seminars and individual faculty consulting efforts require significant time from faculty development specialists and often do not reach faculty in most need of help. The online midcourse evaluation tool used in this study requires almost no investment of professional time from faculty development personnel and reaches far greater numbers of faculty. In this chapter, we investigate faculty and student perceptions of the use of an online, midcourse evaluation tool in improving teaching.

Midcourse Evaluations and Their Effect on Teaching Improvement

In many colleges and universities, students have the opportunity at the end of the course to rate the faculty's teaching. However, the feedback comes too late to benefit the current students directly. It can be used only to benefit the next class of students (Diamond, 2004). Brown (2008) suggested that evaluation reports of faculty should be obtained sooner in the course so that changes can be made before the end of the course.

Cohen (1980) reviewed seventeen studies comparing the impact on perceived quality of instruction of providing midcourse feedback with providing no feedback. He discovered a relatively small impact on teaching effectiveness (effect size = .20). Menges and Brinko (1986) updated Cohen's study six years later and reported a larger effect size when midcourse student ratings were combined with consultation (effect size = 1.10). Prince and Goldman (1981) found that midcourse evaluations led to higher student ratings at the end of a course. Brown (2008) found that 89 percent of students felt faculty should conduct midcourse evaluations because they believed the evaluations would improve instruction as well as their own (student) performance.

Although researchers have studied a variety of issues related to midcourse evaluations, no one has completed a systematic analysis of faculty and student perceptions of the effects of such evaluations on teaching and learning (see Bullock, 2003; Henderson, 2002; Bothell & Henderson, 2003; Johnson, 2003). This study aims to address the gap.

Design of the Research Study

This study used a mixed-methods design to determine faculty and student perceptions of the effects of midcourse evaluations on improving teaching and learning. We received Institutional Review Board approval for this study.

Approximately thirty-four thousand students attend Brigham Young University (BYU), and there are approximately sixteen hundred full-time faculty. BYU values teaching, research, and citizenship equally. End-of-semester online student rating scores are included in summative evaluation of teaching performance for promotion and tenure purposes.

Faculty from all twelve colleges and fifty-two departments were represented in the study. Faculty participants were male and female, included different professorial ranks and tenure status, and showed evidence of their desire to improve their teaching by volunteering to be in the study. Most or all of the faculty in several departments (organizational leadership and strategy, mathematics, the college of nursing, and the counseling and career center) decided to participate, often because their department chairs encouraged them.

Data Collection Procedures

We used a variety of methods, including scores from the midcourse evaluation and end-of-semester online ratings, faculty surveys, faculty interviews, and observations of debriefing sessions.

MIDCOURSE EVALUATION TOOL

Center for Teaching and Learning (CTL) staff created an online evaluation tool containing two survey options to help faculty conduct midcourse evaluations. The first option was a two-question open-ended survey where students indicated in narrative form what was going well in class and what could be improved. The second option consisted of four Likert-scale questions: (1) "I am learning a great deal in this course," (2) "Course materials and learning activities are effective in helping me learn," (3) "This course is helping me develop intellectual skills," and (4) "The instructor shows genuine interest in students and their learning." Response options ranged from strongly disagree (1) to strongly agree (8).

All BYU faculty received an e-mail from the CTL director in September 2008, introducing them to the midcourse evaluation tool and providing a link for them to administer the evaluation. In this e-mail, faculty were encouraged to participate in a study on midcourse evaluations and to use the four-question Likert scale as their midcourse evaluation rather than the two-question evaluation.

Once the faculty member selected the question panel, a survey link was e-mailed to all students in that course. After the survey closed, the faculty

member received an e-mail with an attached spreadsheet containing the students' feedback. The midcourse evaluation tool was available from the first week of the semester until the end-of-semester evaluations became available.

FACULTY SURVEY

A total of 105 faculty agreed to participate in the study, administered the four-question midcourse evaluation, and filled out the faculty survey (Exhibit 12.1).

Exhibit 12.1 Faculty Survey

1. What is your name?
2. What is your department?
3. Did you read the responses of your students from your midcourse evaluation? a. Yes. b. No.
4. Did you discuss the feedback you received from the midcourse evaluation with your students? a. Yes. b. No.
5. How valuable was your experience using the midcourse evaluation tool?
6. Do you think conducting this midcourse evaluation will have an impact on student learning in your course? a. Yes. b. No. c. I'm not sure.
7. Did you make any changes in your teaching because of student feedback from the midcourse evaluation, and if so, what?
8. Would you be willing to participate in a 20-minute interview about your experience using the midcourse evaluation tool? a. Yes. b. No.

DEBRIEFING SESSIONS

From the 105 participants, we randomly selected 30 faculty to be interviewed, and we attended eight debriefing sessions conducted in class by some of those 30 faculty. We observed how the faculty approached the debriefing session, documented student feedback, and recorded how the faculty planned to implement changes as a result of student feedback.

STUDENT SURVEY

At each debriefing session, we administered a survey to students about their experience with the midcourse evaluation.

FACULTY INTERVIEWS

The thirty faculty we had selected participated in interviews lasting twenty to thirty minutes each (Exhibit 12.2). Most of these interviews took place after they had received and debriefed with their students their midcourse evaluation results but before they received their end-of-semester evaluation results. The interviews were recorded and coded in NVivo. We then looked for themes and categorized the responses.

Exhibit 12.2 Faculty Interview Questions

1. Describe your experience using the midcourse evaluation tool. Was it easy to use? How long did it take to administer?
2. How did your students respond to the midcourse evaluation(s)?
3. Did you talk with your students about their feedback from the midcourse evaluation?
4. Did you make any changes in your teaching as a result of feedback from students from the midcourse evaluation?
5. Were there any suggestions from students that you did not take? If so, why not?
6. How did you decide which changes to make in your teaching?
7. Did you use the midcourse evaluation tool twice this semester?
8. Will you conduct a midcourse evaluation next semester?
9. Regardless of your student evaluations, do you feel like your teaching improved because you conducted a midcourse evaluation? What evidence do you have to support this assumption?

END-OF-SEMESTER ONLINE RATINGS

The online student ratings system allows students to confidentially rate their BYU learning experience at the end of each semester. Students can provide feedback about their courses and instructors approximately two weeks before the semester ends. The same four questions that are used to measure perceptions of student learning for the midcourse evaluations are used at the end of the semester. Once grades are submitted, faculty can view the results.

Data Analysis

Before we explain the process of data analysis, we provide definitions for the terms we use in this chapter:

- An *individual item rating* is one response (a number from 1 to 8) to one of the four questions pertaining to student learning. For example, when a student responded “strongly agree” to the eight-point Likert-scale question, “I have learned a great deal in this course,” that response was converted to the number 8. Each student provided four individual item ratings (one for each of the four learning items).
- An *item mean score* is an average of the individual item ratings for each of the four learning items. For example, if one faculty member taught a course with thirty students, we took the thirty individual item ratings from those students for each of the four learning items (120 individual item ratings). We then averaged the ratings for each of the four items to obtain four item mean scores (one item mean score for each of the four items pertaining to perceptions of student learning in each section).
- A *section mean score* is the average of the four item mean scores for a given course.
- A *composite mean score* is the average of the section mean scores for all courses evaluated in the study.

The composite mean score for the midcourse evaluation was compared to the composite mean score for the end-of-semester ratings to determine the effects of the intervention on students’ perceptions of their learning.

We conducted a factor analysis using the extraction method of maximum likelihood and varimax rotation to demonstrate that the mean scores from the four learning items could be combined into a section mean score. Cronbach’s alpha for the four items was 0.92. Based on the cumulative percentage from the extraction sums of squared loadings, the total variance that can be explained by these four factors was 81 percent. The factor loadings ranged from 0.84 to 0.92. These results show internal consistency for each of the four learning items and are good indications of students’ perceptions of their own learning.

Establishing Trustworthiness

The standards used to establish trustworthiness for the qualitative aspects of this study were credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985). To establish credibility, we used a variety of data-gathering methods, such as a comparison between the midcourse and end-of-semester scores, faculty and student surveys,

faculty interviews, and debriefing sessions. To enable transferability, we provided the CTL with a description of the context of the study, the faculty and their circumstances, and rich details from the interviews, including direct quotes. To establish dependability, we sent the faculty copies of their interview and debriefing session transcripts. We also discussed reflections from the interviews, coding structures, insights that arose while coding the data, and the decisions that were made as part of the study with the director of the CTL, faculty in the instructional psychology and technology department, and administrators. To establish confirmability, we provided copies of the recorded interviews and transcripts and made our notes available on request.

Results

Overall, 305 BYU faculty conducted 646 midcourse evaluations (some sent surveys to more than one course or section). Of these, 249 evaluations (124 faculty) used the four-question survey, to which 3,550 students responded. Of these 124 faculty, 65 said they were willing to allow us to observe a debriefing session and to participate in an interview.

Midcourse and End-of-Semester Quantitative Comparison

Overall, the composite mean midcourse score was 6.37, and the composite mean end-of-semester score was 6.71. We conducted a two-tailed, paired *t*-test comparing the midcourse with the end-of-semester composite mean scores and did the same for the mean scores for each of the four items. The mean scores for each of the four items, as well as the composite mean score, increased, showing that on average, students' end-of-semester ratings of the four learning items were significantly higher than their ratings of those same items on the midcourse evaluation (see Table 12.1).

We used Cohen's *d* to further validate the results. The effect size measure represents the standardized difference between the composite means of the midcourse and end-of-semester evaluations. Cohen (1988) defined effect sizes of 0.2 as small, 0.5 as medium, and 0.8 as large, with anything less than 0.2 considered no effect. Out of the 510 item mean scores, Cohen's *d* was 0.46, representing a medium-effect size. Of the faculty who participated in the midcourse evaluation, there was an overall medium, positive effect. Further details are shown in Table 12.2.

Table 12.1 Midcourse and End-of-Semester Student Ratings of Four Learning Items

Areas of Learning	Composite Mid-course Mean (SD)	Composite End-of-Semester Mean (SD)	Paired <i>t</i> -Test (<i>p</i> -Value)
Four areas of learning combined (<i>n</i> = 510)	6.37 (.77)	6.71 (.72)	.001
Interest in student learning (<i>n</i> = 128)	6.84 (.69)	7.08 (.81)	.011
Materials and activities (<i>n</i> = 127)	6.12 (.73)	6.44 (.73)	.001
Amount learned	6.33 (.73)	6.61 (.75)	.002
Intellectual skills (<i>n</i> = 126)	6.16 (.70)	6.64 (.66)	.001

Table 12.2 Effect Size of Midcourse to End-of-Semester Ratings Changes

	n	Mid-course Mean	End-of-Semester Mean	Effect Size			
				Large	Medium	Small	No Effect
Scores that increased	352	6.20	6.86	157 (45%)	75 (21%)	65 (18%)	55 (16%)
Scores that decreased	158	6.74	6.37	37 (23%)	23 (15%)	0 (0%)	98 (62%)

Faculty and Student Surveys

Faculty participants were asked if they read the responses of their students from their midcourse evaluation. Of 103 faculty who responded, 99 (96 percent) said yes and 4 (4 percent) said no. Faculty were asked if they discussed the feedback with their students. Of 105 faculty, 78 said yes (74 percent) and 27 said no (26 percent). Thirty-two faculty members (51 percent) said this was their first time performing a midcourse evaluation, and 31 faculty members (49 percent) said they had completed one before.

Faculty were asked in the survey, "Why did you do a midcourse evaluation?" The 104 faculty who responded to the question provided 118

different reasons. The most common reason was that they wanted to hear the students' opinions (20 responses, 17 percent), followed by feeling that feedback was helpful (17 responses, 14 percent). The third most common reason faculty said they conducted a midcourse evaluation was that they were new faculty or were teaching the course for the first time (15 faculty, 13 percent). The fourth most common reason faculty cited was to improve their teaching (and 28 of the 30 faculty interviewed felt that midcourse evaluations did improve their teaching). The fifth most common reason was to improve student learning.

Of the 126 students from six sections who filled out the survey, 78 (62 percent) had completed a midcourse evaluation before using the online midcourse evaluation tool and 48 (38 percent) had not. Of the 125 students who answered the question, "Did you fill out the midcourse evaluation? If yes, why? If no, why not?" 94 (75 percent) said yes and 31 (25 percent) said no. The most common reason students mentioned for completing the midcourse evaluation was to provide feedback to the instructor (37 students, 30 percent). The second and third reasons were that they received extra credit (25 students, 20 percent) and it was required (20 students, 16 percent). The most common reason students did not fill out the midcourse evaluation was that they forgot about it (14 students), did not receive an e-mail (9 students), were too busy (4 students), and erased it or did it late (1 person in each category). Ninety-seven (78 percent) of the students felt midcourse evaluations were somewhat important or important.

PERCEPTIONS OF STUDENT LEARNING

Of the 30 faculty who were interviewed, 27 (90 percent) felt midcourse evaluations improved student learning. Of the 105 faculty who completed the faculty survey, 62 (59 percent) felt midcourse evaluations improved student learning, and only 12 (11 percent) did not. The rest of the faculty were uncertain and wanted to see their end-of-semester ratings before deciding. From the student surveys, 88 students (71 percent) felt their learning might or would increase because their faculty conducted an evaluation.

ACTIONS TO IMPROVE STUDENT LEARNING

At the end of the midcourse in-class survey, we asked students, "What could be improved?" and "How could this course be more effective in helping you learn?" During interviews, faculty were asked, "What changes are you making or planning to make?" (At the time of the interviews, all participating faculty had had the opportunity to review feedback from their students. Some had already started implementing

Table 12.3 Top Student Changes Compared to Faculty Changes

Changes to Improve Teaching	Student Responses (<i>n</i> = 153)	Faculty Responses (<i>n</i> = 76)
Clearer expectations	60 (39%)	30 (39%)
Active learning	43 (28%)	25 (33%)
Reduce busywork	28 (18%)	6 (8%)
More review	0 (0%)	12 (16%)
No changes	22 (14%)	3 (4%)

changes based on student feedback, and others mentioned changes they planned to make.)

Because we wanted to determine how student suggestions affected teaching performance, we include only the responses from the students of the 22 faculty whose ratings improved by at least 1 point from midcourse to end-of-semester evaluations. We grouped all of the student responses (169) into 29 subthemes, and then seven overarching themes. Top suggestions for improvement from students and actions reported by faculty are in Table 12.3. Three of the four changes were the same for both faculty and students. The first two changes (clearer expectations and active learning) were the top two changes for both groups.

RELATIONSHIP BETWEEN FACULTY ACTIONS AND END-OF-SEMESTER STUDENT FEEDBACK

Of the students surveyed, 56 (45 percent) said they would rate their professor more highly at the end of the semester because he or she had conducted a midcourse evaluation. Student ratings showed improvement in proportion to the extent to which the faculty member engaged with the midcourse evaluation. Faculty who read the student feedback and did not discuss it with their students saw a 2 percent improvement in their online student rating scores. Faculty who read the feedback, discussed it with students, and did not make changes saw a 5 percent improvement. Finally, faculty who conducted the midcourse evaluation, read the feedback, discussed it with their students, and made changes saw a 9 percent improvement.

FACULTY INTERVIEWS

In general, as expressed in interviews conducted after faculty responded to the online survey, faculty thought that doing midcourse evaluations improved student learning. All of the faculty interviewed said they were going to conduct another midcourse evaluation next semester.

Discussion

The results of this study show that faculty and students who participate in midcourse evaluations perceive improvements in student learning and faculty teaching. Although the time between the midcourse evaluation and the end-of-semester student ratings is relatively short (six to eight weeks), quantitative results demonstrate that students' perceptions of their own learning increase significantly when faculty invite their suggestions and then take action to make pedagogical improvements. In brief, small changes in teaching may lead to large improvements in student perceptions of their learning.

These findings are important for those engaged in faculty development. Most centers devoted to the improvement of learning and teaching are eager to find ways to help faculty who struggle in the classroom. Although the data in this study are preliminary and more research needs to be done, the results suggest that the midcourse evaluation tool—because it is completely voluntary and confidential, and because it is easy to administer and act on—is one attractive means of assisting faculty in this area. Using this approach to teaching improvement, the faculty member does not need to be singled out by an administrator to receive help from faculty developers. Rather, the faculty member chooses to take advantage of an unobtrusive intervention that demands only a small amount of time and effort. That being said, research shows that midcourse evaluations can have even greater value when used in a supportive context (Penny & Coe, 2004).

Study Limitations and Additional Questions

Although this study showed that midcourse evaluations can have positive effects on student perceptions of their learning, it does have some limitations and raises a number of additional questions. For example, one limitation is that we did not have a comparison group of courses in which the midcourse evaluation was not administered. Also, students may have based their rating on whether they liked the course content or the instructor rather than the actual instruction.

There are several additional questions to consider for future such studies. For example, to what extent do measures of actual student performance validate student and faculty perceptions of improvements in student learning? How do students respond to pedagogical changes that faculty make as a result of using the midcourse evaluation tool? Do they recognize the changes? Do students act on those changes by making concomitant changes in their learning strategies?

If the midcourse evaluation tool is an effective intervention for assisting faculty who are struggling in the classroom, how can they be encouraged to use the tool and benefit from it? An implicit finding from this study is that personal choice plays an important role in the improvement process. Faculty were not coerced into using the midcourse evaluation tool. Neither were they forced to read or act on the results once the students had offered their suggestions. Thus, how should faculty developers encourage administrators or peers to extend the invitation to struggling faculty to use the tool in ways that will maximize the effectiveness of the intervention?

From the results of this study, it appears that students are generally eager to give feedback to faculty on the quality of their learning experience. However, if midcourse evaluations become ubiquitous and students are required to complete such surveys for every course each semester, will they come to see the evaluation as an additional burden? This concern suggests that midcourse evaluations be kept as brief as possible, requiring only a few minutes for students to complete.

Perhaps the most compelling question that remains to be answered relates to the possible effect of the midcourse tool on the faculty member's ability to read students' reactions to the course. Great teachers are constantly "gathering data" from their students about how well their students are learning (Bain, 2004; Barr & Tagg, 1995). To what extent can the midcourse evaluation tool be used as a learning tool for the faculty member? How can the tool be used to help faculty perceive more clearly how students are learning without administering the midcourse evaluation? This question should be of great interest to any institution interested in establishing an atmosphere of continuous improvement of learning and teaching.

REFERENCES

- Bain, K. (2004). *What the best college teachers do*. Cambridge, MA: Harvard University Press.
- Barr, R. B., & Tagg, J. (1995). From teaching to learning—A new paradigm for undergraduate education. *Change*, 27(6), 12–25.
- Bothell, T. W., & Henderson, T. (2003). Do online ratings of instruction make sense? In T. D. Johnson & D. L. Sorenson (Eds.), *New directions for teaching and learning*: No. 96. *Online student ratings of instruction* (pp. 69–79). San Francisco: Jossey-Bass.
- Brown, M. (2008). Student perceptions of teaching evaluations. *Journal of Instructional Psychology*, 35(2), 177–181.

- Bullock, C. D. (2003). Online collection of midterm student feedback. In T. D. Johnson & D. L. Sorenson (Eds.), *New directions for teaching and learning*: No. 96. *Online student ratings of instruction* (pp. 95–102). San Francisco: Jossey-Bass.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college teaching: A meta-analysis of findings. *Research in Higher Education*, 13, 321–341.
- Diamond, M. R. (2004). The usefulness of structured mid-term feedback as a catalyst for change in higher education classes. *Active Learning in Education*, 5(3), 217–231.
- Fink, L. D. (2003). *Creating significant learning experiences: An integrated approach to designing college courses*. San Francisco: Jossey-Bass.
- Henderson, T. (2002). Classroom assessment techniques in asynchronous learning networks. *Teaching and Learning in Higher Education*, 33, 2–4.
- Johnson, T. D. (2003). Online student ratings: Will students respond? In T. D. Johnson & D. L. Sorenson (Eds.), *New directions for teaching and learning*: No. 96. *Online student ratings of instruction* (pp. 49–59). San Francisco: Jossey-Bass.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.
- Menges, R. J., & Brinko, K. T. (1986). *Effects of student evaluation feedback: A meta-analysis of higher education research*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Penny, A. R., & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research*, 74(2), 215–253.
- Phillips, R. R. (2001). Editorial: On teaching. *Journal of Chiropractic Education*, 15(2), iv–vi.
- Prince, A. R., & Goldman, M. (1981). Improving part-time faculty instruction. *Teaching of Psychology*, 8(3), 60–62.
- Trigwell, K., & Prosser, M. (2004). Development and use of the approaches to teaching inventory. *Educational Psychology Review*, 16(4), 409–424.