Class Meeting Schedules in Relation to Students' Grades and Evaluations of Teaching

Robert C. Reardon
Stephen J. Leierer
Donghyuck Lee

A six-year retrospective study of a university career course evaluated the effect of four different class schedule formats on students' earned grades, expected grades and evaluations of teaching. Some formats exhibited significant differences in earned and expected grades, but significant differences were not observed in student evaluations of instruction. Career services providers, including curriculum designers, administrators and instructors, will find the results of this study helpful in the delivery of services, especially with high-risk freshman students.

Keywords: career, teaching, course, instruction, evaluation, grades

While individual counseling has been shown to be effective in helping students develop career decision-making skills (Brown & Ryan Krane, 2000; Reese & Miller, 2006; Whiston & Oliver, 2005; Whiston, Sexton, & Lasoff, 1998), undergraduate career courses also can be effective interventions (Folsom & Reardon, 2003; Reardon, Folsom, Lee, & Clark, 2011; Whiston et al., 1998).

Although college career courses have been shown to offer substantial benefits (Brown & Ryan Krane, 2000; Osborn, Howard, & Leierer, 2007; Reed, Reardon, Lenz, & Leierer, 2001; Reese & Miller, 2006; Whiston & Oliver, 2005; Whiston et al., 1998), the content and format of such courses vary greatly (Folsom & Reardon, 2003). The present study sought to focus on one aspect of such career course variability: alternative class schedule formats.

Effective career classes can be characterized by these features: (a) structured course approaches appear to be more effective than unstructured approaches (Smith, 1981); (b) individual career exploration should be a cornerstone of the course (Blustein, 1989); and (c) five components (written exercises, individualized interpretations and feedback, in-session occupational exploration, modeling, and building support for choices within one's social network) (Brown & Ryan Krane, 2000; Brown et al., 2003).

What is the effect that class schedule might have on course effectiveness? Only one study (Vernick, Reardon, & Sampson, 2004) has examined this issue, and the results showed that such courses should be designed to meet more than once a week and avoid over-exposure to materials and activities so as not to overwhelm the student. Extending this concept, we hypothesized that certain course schedule formats (weekly meeting frequency and term length) could make a difference in student learning and evaluation of teaching.

Alternative Career Class Schedules

This study focused on a course based on cognitive information processing theory incorporated into the course textbook, *Career Planning and Development: A Comprehensive Approach* (Reardon, Lenz, Sampson, & Peterson, 2000). All sections of the course followed a prescribed curriculum comprising a mixture of lectures, panel presentations, small and large group instructional activities, personal research, and field work; however, the classes differed in terms of the class meeting schedule (class duration, number of weekly meetings, and number of weeks a class met during an academic term).

Robert C. Reardon, NCC, NCCC, is Professor Emeritus at the Career Center at Florida State University. Stephen J. Leierer is an Associate Professor at East Carolina University. Donghyuck Lee is an Assistant Professor at Konkuk University in Seoul, Korea. Correspondence can be addressed to Robert C. Reardon, Career Center, Florida State University, 100 S. Woodward St., Tallahassee, FL 32306-4162, rreardon@admin.fsu.edu.

We examined 57 course sections that met over a six-year period and were team-taught by lead instructors and co-instructors with an instructor/student ratio of about 1:8. Lead instructors included both professional staff and faculty who supervised the co-instructors. During the time of this study, four class schedule formats were used. In the case of a 16-week semester, the class met once per week for 3 hours; twice per week for 1.5 hours; or three times weekly for 1 hour. A fourth schedule option was for a 6-week term with the class meeting four times weekly for about 8 hours per week. In the 16-week semester, the class met once per week for 3 hours on Wednesdays (W); twice per week for 1.5 hours on either Monday/Wednesday or Tuesday/Thursday (MW/TuTh); or three times weekly for 1 hour on Monday, Wednesday, and Friday (MWF). A fourth schedule option was for a 6-week term where the class met four times weekly for about 8 hours per week on Monday, Tuesday, Wednesday, and Thursday (MTuWTh). In summary, we sought to evaluate the influence of these four class schedule formats upon the educational experience of the students as measured by expected grades, earned grades, and student evaluations of teaching.

Course Measures

The following section gives details about the three measures of student learning and perceptions of teaching used in this study.

Earned Grade (EG)

Although a student's grade point average has limitations as a measure of academic achievement, class grades are nevertheless a widely accepted method of quantifying students' level of educational achievement and future success in graduate school or employment (Plant, Ericsson, Hill, & Asberg, 2005). Specific to career development, Reardon, Leierer, and Lee (2007) showed that grades might be useful measures of career course interventions, "especially if the treatment variables are carefully described and the grading procedures are fully explained and replicable by other researchers" (p. 495). For this study, we assumed that a student's final EG would accurately reflect learning in the course.

Expected Grade (XG)

Grade expectations are a complex phenomenon that combines realistic data-driven grade expectations with unjustified optimism or wishful thinking (Svanum & Bigatti, 2006). The XG reflects the student's assessment of course demands and optimism about successfully meeting those demands. This grade prediction may be informed or uninformed; however, after completing multiple assignments over the course of the semester, Svanum and Bigatti (2006) noted that students lower the value of their XG such that it will be only moderately inflated and will reliably predict their final EG. Because students in our course had the course grading scale in the syllabus, a signed performance contract, and predicted their grades during the last week of the semester when 85% of their grade had already been accounted for, we hypothesized that in aggregate their predictions would be only moderately inflated and thus a reliable predictor of their earned grades and success in the course. We felt this grade variable was important as a measure of students' confidence in their mastery of the career development subject matter and the problem-solving skills taught in the course, and therefore a valid measure of the relative effectiveness of different class schedule formats.

In addition, comparing EG and XG informs us about students' self-evaluation of learning and their actual performance in the course. When there is not a significant difference between the two scores, we might suppose that students have a fairly accurate understanding of their performance on completed assignments and those still to be graded. By contrast, a significant difference between XG and EG indicates a discrepancy between students' self-evaluations of graded and as-yet-ungraded assignments and the official final grades. If XG is significantly higher than EG in a section, one may conclude that the academic work has been undervalued by the instructor or overvalued by the students. Conversely, if XG is significantly lower than EG, one might conclude that students' estimates were conservative or instructors recognized a level of performance not seen by the students.

Student Evaluation of Teaching (SET)

Student evaluation of classes and teaching effectiveness is standard practice at most postsecondary institutions. There is substantial anecdotal and experimental evidence supporting the usefulness of SETs (Centra, 1993; Marsh & Dunkin, 1992; Marsh & Roche, 1997). Certain student ratings forms provide important feedback that can be used to improve teaching performance (Greenwald & Gillmore, 1997; Marsh & Roche, 1997; McKeachie, 1997), and when asked most faculty members support the use of SETs as a tool for teaching improvement (Baxter, 1991; Griffin, 1999; Schmelkin, Spencer, &

Gellman, 1997). Although SET is not without its critics, it appears to be a pragmatic way to access and compare student perceptions of teachers' effectiveness and therefore a potential measure of the relative efficacy of different class schedules.

In an effort to better evaluate students' course experiences, the influence of EG (Goldman, 1985) and XG (Greenwald & Gillmore, 1997) on SET is receiving considerable attention in the literature. The present study provided an opportunity to examine the relationship of SET to both EG and XG relative to four different class schedule formats.

Research Questions

In seeking to discover if particular class schedules were more effective in a team-taught career course, we evaluated grades and participant feedback from undergraduate students. The goal was to determine if any of the four differing class schedules produced significant differences in the course evaluation measures EG, XG, and SET. Although we were examining these measures from the students' perspective and such measures are typically scored at the individual student level, we chose to examine class section level scores because XG and SET data were only available to us in this way.

The first group of research questions examined differences between mean evaluative measures, aggregated by class format and averaged for classes that met one (W), two (MW/TuTh), or three times per week (MWF) for 16 weeks, or four times per week (MTuWTh) for 6 weeks.

> Research Question 1: Were there any significant differences in the career course evaluation measures among the four class formats?
> > RQ 1.1: Are there differences in mean EG between formats?
> > RQ 1.2: Are there differences in mean XG between formats?
> > RQ 1.3: Are there differences in mean SET between formats?

The second group of research questions explored the differences between the evaluation measures (EG, XG, and SET) within the sections.

> Research Question 2: Within any given format, are there significant differences between the mean of the aggregated class evaluation measures?
> > RQ 2.1: Is the mean XG significantly different than the mean EG?
> > RQ 2.2: Is the mean XG significantly different than the mean SET?
> > RQ 2.3: Is the mean EG significantly different than the mean SET?

Method

Participants

Over a 6-year period, 1,479 students were enrolled in 57 sections of a career course to fulfill elective requirements for the baccalaureate degree. The class met in a standard classroom in academic buildings on the campus. Although the class was offered for variable credit, over 95% of the students took it for 3 credit hours. The number of students per section ranged from 19–34 with a mean of 26.5.

Ethnic diversity was generally proportional to the general student population of the university: Caucasian, 74%; African American, 12%; Hispanic American, 7%; Other, 4%; Asian, 3%; and American Indian, .4%. The course typically enrolled about 60% females and 40% males, including freshmen (15%), sophomores (45%), juniors (20%), and seniors (20%). Depending on the semester, between 15% and 25% of the course was composed of students with officially undeclared majors, and the large percentage of sophomores was the result of academic advisors referring these undeclared students to the class. While almost 40% of the members in a typical class reported satisfaction with their present career situation, about 60% were unsure, dissatisfied, or undecided.

Course Grading Procedures

Student grades were computed using scores earned on assignments contained in the performance contract. This contract was comprised of 28 different graded activities spread across the three units of the course. Given the use of the performance contract, students in this course should have had a very good idea of what their final grade would be when

they filled out the SET and estimated their grade, because only two of the 28 activities accounting for 125 of 653 total points were still ungraded at that point.

Student Evaluation of Teaching Ratings

We used a standardized instrument for SETs, the Student Instructional Rating System (SIRS; Arreola, 1973), a student course on form developed at Michigan State University (Davis, 1969) and adapted for use at our university. SIRS provided an opportunity for instructors to obtain reactions to their instructional effectiveness and course organization and to compare these results to those of similar courses offered within the university.

The SIRS consisted of 32 items and 25 of these items enabled students to express their degree of satisfaction with the quality of instruction provided in the course by using a 5-point Likert scale. For example, *the course was well organized* could be marked *strongly agree, agree, neutral, disagree,* or *strongly disagree.* One item on the SIRS was of special interest in this study: *What grade do you expect to receive in this course? A, B, C, D, or F.*

We also employed a second instructional rating instrument, the State University System Student Assessment of Instruction (SUSSAI) which had been used at the university for five years prior to this study. This instrument consisted of eight items focused on class and instructor evaluation. One item was of special interest in this study: *Overall assessment of instructor: Excellent=4, Very Good=3, Good=2, Fair=1, Poor=0.*

Data Collection

After obtaining permission from the university institutional review board, we received the archived career course grade data for a six-year period. We aggregated the grades of these 1,479 students by class schedule and averaged the results to achieve a mean EG for each class schedule format.

The data relating to students' perceptions of what they had achieved and the quality of instruction they had received was collected as follows: On the last week of class, while filling out their teacher evaluations, all students in a section were asked to indicate the grade they expected to receive and the results were tallied and averaged to determine a class mean XG. These class averages of 57 sections were forwarded to the researchers, and the results were tallied and averaged to find the mean XG for each class schedule format. In addition, we retrieved overall class ratings of instructors for an ad hoc sample of career classes over the 6-year period. These data enabled us to examine the relationships between mean EG and XG, EG and SET, and XG and SET.

Procedures

In this team-taught course where all instructors were involved in making large- and small-group presentations, each co-instructor had primary responsibility for evaluating the progress of students in his or her small group and assigning a grade, while the lead instructor of the team had overall responsibility for course presentations and management. In completing the SIRS and SUSSAI items for the SET, students were asked to provide a composite rating of the instructional team for their section. SETs were completed anonymously during the final two class meetings while instructors were out of the room and then returned by a student proctor to the university's office of evaluation services.

Data Analysis

We examined how different class formats influenced mean EG, XG, and SET. The independent variable of class schedule format had four levels. The first three levels met over the course of a 16-week fall or spring semester for either 3 hours once a week (W), 1.5 hours twice a week (MW/TuTh), or 1 hour three times a week (MWF). The final level met for 2 hours four times a week over the course of a 6-week semester (MTuWTh). Because the assumptions related to independence for the three evaluative measures could not be met (i.e., the evaluations for each class section were correlated), we analyzed the data using a split-plot design.

Results

As is the case for other ANOVA and MANOVA tests, the dependent variables were assumed to be normally distributed. We tested the dependent variables to determine if they were normally distributed by computing skewness and kurtosis of

each of the dependent variables to see if they fell between −1.0 and +1.0. Both the SET and EG scores did not violate the assumptions of normality as measured by skewness and kurtosis. However, while the skewness of XG did fall within the appropriate range, the kurtosis score was 1.04. Although this score is above 1.00, we believe this minor violation does not seriously affect the results and their interpretation.

Research Question 1

Using the split-plot MANOVA, we found a significant interaction of the three evaluative measures across the four class formats $F(6, 106) = 4.47$, $p < .0005$, $\eta^2 = .20$. Specifically, there was a significant difference in EG between the four course formats, $F(3, 53) = 19.15$, $p < .0005$, partial $\eta^2 = .52$. The EG for schedule MTuWTh ($M = 3.50$) was significantly higher ($p < .005$) than that of formats W, MW/TuTh, and MWF ($M = 3.25$, 3.32, and 3.31, respectively). Next, there was a significant difference in XG between the four course formats, $F(3, 53) = 3.62$, $p = .019$, $\eta^2 = .02$. The means for XG for the W, MW/TuTh, MWF, and MTuWTh were 3.71, 3.57, 3.34, and 3.64, respectively. There was not a significant difference for XG between formats W, MW/TuTh, and MTuWTh. However, there was a significant difference between format MWF and format MTuWTh ($p = .036$), and format MWF was trending lower when compared with format W ($p = .097$) and format MW/TuTh ($p = .051$). Finally, there was not a significant difference on SET scores across the four formats, $F(3, 53) = 1.36$, $p = ns$. The mean SET scores for formats W, MW/TuTh, MWF, and MTuWTh were 2.88, 3.15, 3.31, and 3.11, respectively.

Research Question 2

When we compared evaluation measures within each format, we found significant differences with each one, $F(2, 52) = 23.61$, $p < .0005$, $\eta^2 = .47$. We found XG significantly greater than EG within schedule format W (.46, $p = .002$) and format MW/TuTh (.35, $p < .0005$). By contrast, the difference between XG and EG was smaller and not statistically significant within format MWF (.13, $p = ns$) and format MTuWTh (.13, $p = ns$). This lack of a significant difference between EG and XG indicates that these students earned grades very similar to the grades they expected to receive. It is apparent that the students and instructors used similar evaluation and grading methods. Stated another way, this finding suggests that students in classes meeting more frequently per week have a slightly more accurate perception of how they are doing in the class.

We also found that mean XG was significantly greater than mean SET for format W (.83, $p = .003$), format MW/TuTh (.42, $p < .0005$), and format MTuWTh (.53, $p < .0005$). However, there was not a significant difference between XG and SET for format MWF (.13, $p = ns$). Finally, in comparing the difference between mean EG and mean SET within each of the four formats, we found a significantly higher EG only for format MTuWTh (.40, $p < .0005$). No significant differences were observed for formats W, MW/TuTh, and MWF, which had differences of .37, .07, and .13, respectively.

In summary, we found significant differences in the evaluation measures of XG, EG, and SET across the four different career course formats. Class sections which met four times a week for 6 weeks had a significantly higher EG than classes meeting one, two, or three times a week for a 16-week semester. Interestingly, formats W, MW/TuTh, and MTuWTh all had mean XG scores over 3.55, while format MWF's XG was not only lower than the other formats, but significantly lower than that of format MTuWTh. Finally, mean SET scores were not significantly different from one another. Notably, they were all well above the rating of "good" (good = 2.0), with a mean of 3.15 on a 4-point scale. Means for the sections ranged between 2.88 and 3.31; thus we concluded that students found the instruction to be very good or excellent.

Discussion

Career course interventions have been developed to help students improve their academic and career decision-making skills. Comprehensive career courses offered for academic credit represent a cost-effective intervention that could be described as a "mega-dose" of career services (Reardon et al., 2011). While the benefits of college career courses are clear, it is unclear what contributions specific class formats (differing by length of class period, number of classes per week, length of course in weeks) might make to their effectiveness. Thus, the purpose of our study was to analyze the influence of different schedule formats on earned and expected grades and students' evaluation of their instructors.

Previous studies on career development classes have described various limitations (see Gold, Kivlighan, Kerr, & Kramer, 1993; Reese & Miller, 2010), and we attempted to address these in the following ways. First, although we did not

directly address random selection and random assignment issues, we aggregated class section scores instead of individual student scores, thus reducing the effect of individual outliers. By using the aggregate mean for each career planning section, individual students' evaluation of the teacher remained anonymous yet the evaluation of the course section remained intact. The second limitation described by other researchers is the small number of participants in the career class analyzed. Over a six-year period we were able to collect data from almost 1,500 students from 57 sections of the course. The third limitation we attempted to address was the lack of equal representation of different ethnic groups. While we did not have equal percentages of students from different ethnicities, the demographic composition of our sample closely matched the composition of our university.

Perhaps the greatest strength of this study's design was the replication of the intervention. That is, because the course structure and specific assignments were very similar for all sections, in effect the replication of the career course occurred across all 57 of the course sections analyzed. In each section, the course content and procedures were clearly specified and grades were based on the successful execution of a performance contract by the student.

Earned and Expected Grades
We examined how schedule influenced mean earned grade (EG) and expected grade (XG) scores. Like Vernick et al. (2004), we found that sections meeting only once per week over 16 weeks (format W) had the lowest EG, though not significantly lower than formats MW/TuTh and MWF. By contrast, schedule MTuWTh had a significantly higher EG than all the other formats, suggesting that a 6-week semester of 2-hour class meetings four times a week was more conducive to learning than a 16-week semester of classes meeting one, two, or three times per week for 3 hours, 1.5 hours, or 1 hour, respectively; that is, the "mega-dose" of career development interventions given in the course were intensified with MTuWTh.

Further analysis of the difference between mean section EG and XG scores enables us to compare the students' view of their performance in the course with their actual performance. Ideally, we would prefer that there not be a significant difference between XG and EG in order to increase students' confidence about the fairness of the grading and their sense of having mastered the material in the course. Expanding on these points, when the section mean XG was significantly higher than the mean EG, students could have left the course with a sense of failure and disappointment. Interestingly, in this study schedules W and MW/TuTh had significantly higher mean XG than mean EG, indicating an incongruity between the expected and earned grades. By contrast, for both schedules MWF and MTWF, the difference between mean XG and mean EG was not significant. One might conclude that fewer course meetings per week increased the difference between XG and EG scores.

Student Evaluation of Teaching
With regard to SET, there were no significant differences between the four class schedule formats, although we had suspected this might be the case. Perhaps a significant difference between section means for SET and XG would describe an incongruity between the students' estimate of instruction quality and their evaluation of their own performance in the course. If XG were significantly higher than SET, this finding might indicate that students in these sections believed their performance was more related to their abilities and efforts rather than course instruction. By contrast, sections with significantly lower XG than SET scores may have rated instructors' presentation of material higher than their own performance in the course. Interestingly, for schedules W, MW/TuTh, and MTuWTh, XG was significantly higher than SET, suggesting that students evaluated themselves more favorably than they did their instructors. We found it curious that for schedule MWF alone, XG was not significantly higher than SET.

Finally, EG is assigned to the student by the instructor, while SET is assigned to the instructor by the student. By comparing mean EG with SET, we can examine the relationship between an instructor's evaluation of his or her students with students' evaluation of the instructor. When EG is greater than SET, this means that instructors evaluated their students more favorably than they themselves were evaluated; conversely, when SET is greater than EG, students evaluated instructors more favorably than they themselves were evaluated. For schedules W, MW/TuTh, and MWF, there were no significant differences between mean EG and SET scores. However, for MTuWTh, in which students achieved a significantly higher mean EG than the other formats, the EG also was significantly higher than the SET, suggesting that this high-performing group had higher expectations for their instructors than they felt the instructors met.

Limitations

Because this study is field research, there are a few limitations to discuss. First, participants were undergraduates taking a career planning course from one university. The advantage to using this approach was consistency of teaching content, training and quality control of teaching personnel, administration of tests, and assignments, thus reducing the possibility that course differences were responsible for random error variance. But, because these results come from only one university's career course, caution should be exercised when generalizing them to other courses.

Second, participants were not randomly selected. In fact, random assignment was impossible given the students' autonomy in selecting this course. Random selection is seldom an option in field research at an educational institution, but this fact does restrict the robustness and generalizability of results to other populations (Babbie, 2001).

Third, participants in the study may have been experiencing more career-related difficulties than other students who did not elect to take the course. It is to be expected that participants perceived a career course as more important to their progress than nonparticipants, which limits generalizability of these findings (Smith & Glass, 1987).

Fourth, because the data were collected over a six-year period, it is difficult to determine the effect of historical events on the behavior and attitudes of participants (Smith & Glass, 1987; Van Dalen, 1979). For example, students from the initial semester of the study took the class at the height of the tech bubble, while others took the class in the shadows of the 9-11 tragedy. Although we were not able to control for these events, we acknowledge that researchers and practitioners must be aware of the influence of external events upon any college course.

Implications

There are several implications regarding the findings of this study. The significant differences found between schedule formats in the outcomes of EG and XG serve to remind instructors, those who supervise them, and those managing career courses about the potential impact of this variable. For example, these findings indicate that classes meeting one time per week for three hours are not characterized by higher earned grades, and by implication this means student learning. Additional studies should isolate and evaluate format variables such as length of the entire course, number of classes per week, and length of individual classes so that those evaluating teachers might consider this in their evaluations. At the same time, the absence of any differences in student evaluations of teaching across the four schedule formats is reassuring for those teaching and supervising instructors, at least in a course that was as highly structured and standardized as the one in this study.

Career services providers, curriculum designers, administrators, and instructors may wish to consider these findings when making decisions about the design and delivery of career courses, especially for high-risk freshmen (Osborn et al., 2007). Students meeting for four classes a week over a 6-week semester earned and expected significantly higher grades overall than students meeting over a 16-week semester. Taking the 6-week intensive course during the summer term before beginning the freshman year could both increase students' chances of academic success and their confidence in navigating the college experience.

References

Arreola, R. A. (1973). A cross-institutional factor structure replication of the Michigan State University SIRS faculty evaluation model. *College Student Journal, 7*, 38–42.

Babbie, E. (2001). *The practice of social research*. Belmont, CA: Wadsworth/Thomson Learning.

Baxter, E. P. (1991). The TEVAL experience, 1983–88: The impact of a student evaluation of teaching scheme on university teachers. *Studies in Higher Education, 16,* 151–179.

Blustein, D. L. (1989). The role of career exploration in the career decision making of college students. *Journal of College Student Development, 30,* 111–117.

Brown, S. D., & Ryan Krane, N. E. (2000). Four (or five) sessions and a cloud of dust: Old assumptions and new observations about career counseling. In S. D. Brown & R. W. Lent (Eds.), *Handbook of counseling psychology*

(3rd ed., pp. 740–766). New York, NY: John Wiley & Sons.

Brown, S. D., Ryan Krane, N. E., Brecheisen, J., Castelino, P., Budisin, I., Miller, M., & Edens, L. (2003). Critical ingredients of career choice interventions: More analyses and new hypotheses. *Journal of Vocational Behavior, 62,* 411–428.

Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness.* San Francisco, CA: Jossey-Bass.

Davis, R. H. (1969). *Student Instructional Rating System (SIRS) Technical Bulletin.* East Lansing, MI: Michigan State University, Office of Evaluation Services.

Folsom, B., & Reardon, R. (2003). College career courses: Design and accountability. *Journal of Career Assessment, 11,* 421–450.

Gold, P. B., Kivlighan, D. M. Jr., Kerr, A. E., & Kramer, L. A. (1993).The structure of students' perceptions of impactful, helpful events in career exploration classes. *Journal of Career Assessment, 1*, 145–161.

Goldman, L. (1985). The betrayal of gatekeepers: Grade inflation. *Journal of General Education, 37*, 97–121.

Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist, 52,* 1209–1217.

Griffin, B. W. (1999). *Results of the faculty survey on student ratings of instruction: Preliminary report.* Statesboro, GA: Georgia Southern University, Student Ratings Committee.

Marsh, H. W., & Dunkin, M. (1992). Students' evaluations of university teaching: A multidimensional perspective. In J. C. Smart (Ed.), *Higher education: Handbook on theory and research* (Vol. 8, pp. 143–234). New York, NY: Agathon Press.

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52*, 1187–1197.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52,* 1218–1225.

Osborn, D. S., Howard, D. K., & Leierer, S. J. (2007).The effect of a career development course on the dysfunctional career thoughts of racially and ethnically diverse college freshmen. *Career Development Quarterly, 55*, 365–377.

Plant, E. A., Ericsson, K. A., Hill, L., & Asberg, K. (2005). Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance. *Contemporary Educational Psychology 30*, 96–116. doi:10.1016/j.cedpsych.2004.06.001

Reardon, R. C., Folsom, B., Lee, D., & Clark, J. (2011). *The effects of college career courses on learner outputs & outcomes: Technical report No. 531.* Tallahassee, FL: Center for the Study of Technology in Counseling and Career Development, Florida State University. Retrieved from http://career.fsu.edu/techcenter/whatsnew/TechRept53.pdf.

Reardon, R. C., Leierer, S. J., & Lee, D. (2007). Charting grades over 26 years to evaluate a career course. *Journal of Career Assessment, 15*, 483–498. doi:10.1177/1069072707305767.

Reardon, R. C., Lenz, J. G., Sampson, J. P., Jr., & Peterson, G. W. (2000). *Career development and planning: A comprehensive approach*. Pacific Grove, CA: Wadsworth-Brooks/Cole.

Reed, C., Reardon, R., Lenz, J., & Leierer, S. (2001). Reducing negative career thoughts with a career course. *Career Development Quarterly, 50*, 158–167.

Reese, R. J., & Miller, C. D. (2006). Effects of a university career development course on career decision-making self-efficacy. *Journal of Career Assessment, 14*, 252–26.

Reese, R. J., & Miller, C. D. (2010). Using outcome to improve a career development course: Closing the scientist-practitioner gap. *Journal of Career Assessment, 18*, 207–219.

Schmelkin, L. P., Spencer, K. J., & Gellman, E. S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education, 38,* 575–592.

Smith, G. E. (1981). The effectiveness of a career guidance class: An organizational comparison. *Journal of College Student Personnel, 22,* 120–124.

Smith, M. L., & Glass, G. (1987). *Research and evaluation in education and the social sciences.* Englewood Cliffs, NJ: Prentice-Hall.

Svanum, S., & Bigatti, S. (2006). Grade expectations: Informed or uninformed optimism, or both? *Teaching of Psychology, 33*, 14–18.

Van Dalen, D. (1979). *Understanding educational research: An introduction.* New York, NY: McGraw-Hill.

Vernick, S. H., Reardon, R. C., & Sampson, J. P. Jr. (2004). Process evaluation of a career course: A replication and extension. *Journal of Career Development, 30*, 201–213.

Whiston, S. C., Sexton, T. L., & Lasoff, D. L. (1998). Career-intervention outcome: A replication and extension of Oliver and Spokane. *Journal of Counseling Psychology*, *45*, 150–165.

Whiston, S. C., & Oliver, L. W. (2005). Career counseling process and outcome. In W. B. Walsh & M. Savickas (Eds.), *Handbook of vocational psychology* (3rd ed., pp. 155–194). Hillsdale, NJ: Erlbaum.

nbcc