College of Agriculture Course Evaluation Patterns in Overall Value of Course and Quality of Teaching

Tracy Kitchel[1]
University of Kentucky
Lexington, KY 40546

J. Shane Robinson[2]
Oklahoma State University
Stillwater, OK 74078

Charles Cordell Jenkins, III[3]
University of Kentucky
Lexington, KY 40546

Abstract

The purpose of the study was to examine course evaluations focusing on the overall course and teaching quality in light of the course characteristics and individual evaluation items in the College of Agriculture at the University of Kentucky. Students were more likely to rate the value of the course higher when a course stimulated interest in the subject for the students. Value of the course was also associated with course materials being presented in an effective manner. Student scores on the quality of teaching were influenced more by the comments students received from graded assignments, followed by the extent the course stimulated interest in the subject for the students. Instructor's ability to answer questions also related highly with students' scores on quality of teaching. When comparing the course evaluation items by course-level, hours spent studying per week, and grade expected, some differences were noted. Students in 600-level course tended to rate the course and instructor higher than any other level. Students who reported studying one hour or less per week for their course rated the course and teacher the lowest. Students believing they were to receive an "A" rated the teacher and course the highest.

Introduction

Practicing effective teaching has long been a concern for faculty within the higher education community. Higher education institutions have continually striven to produce excellence in teaching (American Association of University Professors [AAUP], 2006), along with conducting quality research and providing valuable service (Marlin, 1987). Because research and service are easily measured, and effective teaching is highly subjective, effective teaching all too often takes a backseat and is not considered as worthy in the process of providing quality scholarship. Boyer (1990) emphasized that "… excellence in the classroom all too often is undervalued" (p. 37). Furthermore, Boyer noted that in order for teaching to be considered on the same level of worthiness as research, criteria for assessing effective teaching must exist. One criteria of evaluating effective teaching is through assessing course evaluations.

Course evaluations provide a wide array of insights to instructors and administrators in higher education institutions. Thompson and Serra (2005) noted that the purpose of student evaluations is to provide feedback to instructors to improve their teaching effectiveness. Cruse (1987) noted that course evaluation results are also important for selecting instructors for salary increases, promotions, and tenure. Overall, course evaluations should be used to assist instructors in performing in the classroom, advising students, and maintaining proper informal contact with students outside of the formal classroom environment (AAUP, 2006).

A review of the literature revealed that research has been conducted on what characteristics influence course evaluations. Gramlich and Greenlee (1993) studied the impact final grades had on student evaluations on economics students at the University of Michigan and found that there is "only a slight relationship" (p. 12) when assessing students' final grades in economics courses and the evaluations of the instructor.

Wollert and West (2000) assessed how faculty characteristics at East Tennessee State University impacted course evaluations. The findings from this study revealed that faculty at the instructor rank received higher scores on course evaluations than did faculty at the full professor rank. In addition, both instructors and full professors received higher scores on course evaluations than did associate professors. The study further concluded that there was no

[1]Assistant Professor of Agricultural Education, Department of Community and Leadership Development, Tel: 859-257-4273; Email: tracy.kitchel@uky.edu
[2]Assistant Professor of Agricultural Education, Department of Agricultural Education, Leadership and Communications
[3]Graduate Research Assistant, Department of Community and Leadership Development

significant difference between the gender of the faculty member and the scores received on course evaluations.

While gender may not have a bearing on course evaluations, the level of perceived warmth does. Best and Addison (2000) studied the relationship between faculty's perceived warmth and student's evaluations of their instruction. The findings from this study suggested that the warmer students perceived the faculty member to be, the higher the scores the faculty member received.

Rodin and Rodin (1973) sought to determine if the amount students learned in class had a relationship on course evaluation scores. In assessing this study across classes, the authors found a negative correlation. The authors noted that, when comparing the mean scores of various classes, the more students learned, the lower they tended to score their professors. Krautmann and Sander (1999) assessed the impact that a students' expected grade in a course had on their evaluations of their respective instructor and concluded that grades affect student evaluations and that "faculty have the ability to 'buy' higher evaluations by lowering their grading standards" (p. 61).

The amount of work involved in a course has also been identified as a possible criterion by which to predict scores on course evaluations. Coleman and McKeachie (1980) noted that while faculty have suggested that "easy" faculty are more popular with students and receive higher course evaluation scores, their study did not yield the same results. In an experimental study on the effects that course evaluations have on students' course selection, the authors found that students selected instructors and courses which received high course evaluation scores the previous semester, regardless of the amount of work involved in the course.

While course evaluations have provided much needed data in evaluating teacher effectiveness, they are not without their share of problems. Researchers have cautioned against using course evaluations as an "end all" for evaluating teacher effectiveness. Faculty have been known to downplay student evaluations because of the lack of validity and reliability that novices (a.k.a. students) provide. In a study conducted to assess the validity of course evaluations, Onwuegbuzie et al. (2007) found a disparity between how the developers of course evaluations defined effective teaching versus how students defined effective teaching. As a result of this disparity, the authors noted that the findings of their study "… cast some serious doubt on the content-related validity (i.e., item validity, sampling validity) and construct-related validity (i.e., substantive validity, structural validity, outcome validity, generalizability)…" (p. 151) and thus caution should be exercised when making decisions on pay, promotion, and tenure of faculty based off of course evaluations completed by students. However, other researchers argue that, although course evaluations are not perfect, they are a valid and reliable instrument that should be used to assess effective teaching. Chen and Hoshower (2003) stated that: "Although methodological problems have been identified, there seems to be some support for both the reliability and validity of student ratings. Overall, the literature supports the view that properly designed students ratings can be a valuable source of information for evaluating certain aspects of faculty teaching performance" (p. 73).

Students do not always feel as though their participation is appreciated or respected. In an assessment of course evaluations at two separate institutions, Marlin (1987) reported that most students assumed that their comments and ratings of instructors would not be read or taken seriously by the instructors or the administration. Some even noted that they doubted whether administrators even cared about course evaluations at all. For students who feel as though their responses go unnoticed, they are more apt to respond as "loose cannons" and be overly critical of their instructors (Tricker et al., 2005).

Course evaluations are not always perceived well by faculty. One common faculty misperception about student evaluations is that students lacked the wisdom and expertise to properly evaluate effective teaching (Felder, 1992). Other myths of course evaluations consist of course evaluations being popularity contests (Felder, 1992) and that only easy classes are rewarded with high course evaluations (Thompson and Serra, 2005). When considering the popularity contests described by Felder (1992), relying solely on student evaluations can also pressure professors to feel as though they have to satisfy students at all costs (AAUP, 2005). However, the findings from these studies indicate that these suggestions are indeed myths and that students are capable of making "meaningful differentiations" (Thompson and Serra, 2005) p. 698, between effective and ineffective teachers.

Students should be informed that their comments and ratings of teacher effectiveness are taken into account in an effort to develop higher quality teachers. Course evaluations should be conducted with a high degree of integrity because of the emphasis placed on evaluating effective teaching. Specifically, an article by AAUP (2005) stated "…institutions, departments, and faculty members should ensure that the evaluations of teaching promote and sustain excellence of teaching and education, that faculty be primarily responsible for devising systems of evaluation and monitoring their use, and that the development and implementation of teaching evaluation methods be consistent with principles of academic freedom and shared governance (Observations section, 16)."

Course evaluations provide an avenue for students to reflect on the teaching effectiveness of the instructor as well as the overall value of the course

(Ouallal-McRiffey, 2005). Therefore, the framework for this study was rooted in reflection. York-Barr et al., (2001) stated that "reflective practices facilitate learning, renewal, and growth throughout the development of career educators" (p. 1). Schön (1983) stated that there are two forms of reflection: reflection-in-action, which refers to reflection that takes place during teaching due to an unexpected reaction, and reflection-on-action, which occurs during both the preparation stage and the post-teaching stage. This study employed course evaluations to provide a form of student feedback that assisted in the reflection-on-action stage.

Effective teaching is measured, in part, by high scores on course evaluations. At the University of Kentucky, course evaluation items of overall value of the course and quality teaching are reported and taken into consideration by college administrators for review and for promotion and tenure purposes. In addition, the results of course evaluations at the University of Kentucky are made public for students as a means of selecting appropriate courses.

While a significant amount of research exists in evaluating and assessing course evaluations in higher education, the literature is scant in colleges of agriculture. With respect to the emphasis placed on effective teaching by higher education faculty, certain questions arise. What trends exist with college of agriculture course evaluations? In particular, when focusing on such summary items as course value and quality of teaching, what aspects of teaching and the course influence those scores? Are there trends that arise when analyzing those summary items in light of characteristics of the course or students?

Purpose and Objectives

The purpose of the study was to examine course evaluations focusing on the overall course and teaching quality in light of the course characteristics and individual evaluation items in the College of Agriculture at the University of Kentucky. The following objectives were crafted to conduct this study:

1. Profile course characteristics (course level, hours per week studied, expected grade and course type – elective vs. required) of the group being analyzed for this study.
2. Determine if a relationship exists between the individual course evaluation items and the specific item "Overall Value of the Course."
3. Determine if a relationship exists between the individual course evaluation items and the specific item "Overall Quality of Teaching."
4. Compare course evaluation results items by the following categories: course level, hours per week studied, and expected grade.

Methods

The target population for this descriptive-correlational study was students completing course evaluations for a course in a college of agriculture. Because the unit of analysis was "a student completing a course evaluation for a course in a college of agriculture," students completing more than one course within the college were counted multiple times. This study utilized a time and place sample (n = 4609) of the population for the spring 2006 semester. This sample should not be considered representative as the non-respondent rate is unknown; therefore, some caution should be applied when interpreting results.

The data collection instrument was the course evaluation form at the University of Kentucky. The University Senate produced a uniform course evaluation to be implemented campus-wide. The course evaluation forms are divided into sections: course, instructor, learning outcomes and summary. These sections were comprised of eight, six, five, and two items, respectively. The summary items were "rate the overall value of this course" and "rate of the overall value of teaching by the primary instructor of

Table 1. Course Characteristics

Course Characteristics	Frequency	Percentage	Valid Percentage
Course Level			
100-level class	855	18.4	18.4
200-level class	976	21.0	21.0
300-level class	1057	22.8	22.8
400-level class	769	16.6	16.6
500-level class	582	12.5	12.5
600-level class	164	3.5	3.5
700-level class	85	1.8	1.8
800-level class	142	3.1	3.1
900-level class	12	.3	.3
Missing	0	0.0	0.0
Total	4642	100.0	100.0
Hours per Week			
1 hour or less	1492	32.2	35.1
2 hours	1224	26.4	28.9
3 hours	855	18.4	20.2
4-5 hours	419	9.0	9.9
6-7 hours	95	2.0	2.2
8 or more hours	157	3.4	3.7
Missing	400	8.6	-
Total	4242	100.0	100.0
Expected Grade			
A	2368	51.0	56.2
B	1446	31.2	34.2
C	297	6.4	7.0
D	19	.4	0.4
E/Fail	4	.1	0.1
I	1	0.0	0.0
Pass or Audit	89	1.9	2.1
Missing	418	9.0	-
Total	4224	100.0	100.0
Reason for taking course			
Required by University	216	4.7	5.1
Required by major	2964	63.9	70.0
Other	1057	22.8	24.9
Missing	405	8.6	-
Total	4237	100.0	100.0

this course." A 4-point scale was utilized for the sections: course, instructor, and learning outcomes with 1 = strongly disagree, 2 = disagree; 3 = agree, and 4 = strongly agree. A 4-point scale was utilized for the summary items with 1 = poor, 2 = fair, 3 = good, and 4 = excellent. According to the Office of Institutional Research (OIR, 2007), the course evaluation instrument was pilot tested, and validity was established. Details as to the validity process were not explained. There was no mention of reliability.

Data were collected by the University of Kentucky at the end of the spring 2006 semester courses. Instructors distribute the instruments, and then have the instruments collected and submitted to OIR by a student or teaching assistant. The OIR then scans in the results. Researchers obtained the data set from the OIR for analysis. To analyze objective one, means and standard deviations were calculated for interval or ratio scale items and frequencies and percents were reported for nominal items. For objectives two and three, relationships were calculated using the Pearson product-moment correlation. Data for objective four were reported as means and standard deviations by the various levels within each of the following categories: course level, hours per week studied, and expected grade.

Results

Objective one was to profile course characteristics (course level, hours per week, expected grade, and course type – elective vs. required) of those courses evaluated for the study. As illustrated in Table 1, the largest group, in terms of course level, was represented by students who were enrolled in 300-level classes (22.8%) followed by students enrolled in 200-level classes (21.0%). The smallest group (.3%) representative students enrolled in 900-level classes.

Regarding hours per week, the majority (32.2%) of students indicated they spent one hour or less followed by 26.4% students who indicated they spent two hours per week. The group spending six to seven hours per week represented the smallest (2%) amount of students. The majority (51%) expected to receive an "A" in the course. The students expecting to receive a "B" in the course represented 31.2%. None of the student's responses indicated they expected an "I" while only .1% of the students expected to receive an "E/Fail." The majority (63.9%) indicated the reason they took the course was "required by my major." The reason "other (e.g. elective)" was the second most popular (22.8%)

Table 2. Pearson Product Moment Correlations between Course Evaluation Items and the Summary Items

Course Evaluation Items	Rate the overall quality of teaching by the primary instructor in this course.	Ranking of magnitude of correlation for quality of teaching[1]	Rate the overall value of this course.	Ranking of magnitude of correlation for overall value of course[1]
The course simulated me to read further in the area. [L]	.58	12	.64	4
I gained an understanding of concepts and principles in this field. [L]	.66	4	.66	2
The course helped me to develop the ability to solve problems. [L]	.58	12	.62	6
The course strengthened my ability to analyze and evaluate information. [L]	.60	8	.63	5
I learned to respect viewpoints different from my own. [L]	.53	16	.55	10
The instructor encouraged student participation in class. [I]	.61	6	.55	10
The instructor stimulated my interest in the subject. [I]	.70	2	.69	1
The instructor satisfactorily answered questions raised in class. [I]	.67	3	.58	9
The instructor was available for consultation outside of class during office hours. [I]	.60	8	.52	13
The instructor had good knowledge of subject matter. [I]	.59	10	.51	15
The instructor presented course material in an effective manner. [I]	.74	1	.66	2
Graded assignments included helpful comments from the instructor. [C]	.58	12	.52	13
Graded assignments, tests, etc., were returned promptly. [C]	.51	18	.48	18
Assignments were distributed fairly throughout the semester. [C]	.53	16	.51	15
Grading in the course was fair and consistent. [C]	.62	5	.55	10
Examinations reflected what was taught in the course. [C]	.61	6	.59	8
The assignments (supplemental readings, homework, reports, etc.) helped me to understand the subject. [C]	.59	10	.60	7
The textbook(s) contributed to my understanding of the subject. [C]	.42	19	.47	19
At the beginning of the course, the instructor outlined in reasonable detail course material and grading procedures. [C]	.55	15	.51	15

Note. [C]= Course Item, [I]= Instructor Item, [L]= Learning Outcome Item, [S]=Summary Item
[1] Indicates magnitude from highest to lowest

reason for taking a course. The reason "required by university studies program" was the least popular representing only 4.7% of the group (Table 1).

Objective two was to determine if relationships existed between the individual course evaluation items (course items, instructor items, and learning outcomes) and the summary item "rate the overall value of this course." The instructor item "the instructor stimulated my interest in the subject" (r = .69) had the highest positive correlation. The learning outcome item "I gained an understanding of concepts and principles in this field" and the instructor item "the instructor presented course material in an effective manner" had the next highest positive correlations (r = .66). The course items "the textbook(s) contributed to my understanding of the subject" and "graded assignments, tests, etc., were returned promptly" were the lowest positively correlated course evaluation items with correlations of .47 and .48, respectively (Table 2).

Objective three was to determine if relationships existed between the individual course evaluation items (course items, instructor items, and learning outcomes) and the summary item "rate the overall quality of teaching by the primary instructor in this course." The course item "graded assignments included helpful comments from the instructor" had the highest positive correlation (r = .74). The instructor item "the instructor stimulated my interest in the subject" (r = .70) had the second highest positive correlation followed by the instructor item "the instructor satisfactorily answered questions raised in class" (r = .67) having the third highest positive correlation. The course items "the textbook(s) contributed to my understanding of the subject" and "graded assignments, tests, etc., were returned promptly" were the lowest positively correlated course evaluation items with correlations of .42 and .51, respectively (Table 2).

Objective four was to compare all course evaluation items by the following categories: course level, hours per week studied, expected grade, and course type – elective vs. required. As Table 3 and 4 indicate, students who filled out course evaluations rated the instructor item "the instructor had a good knowledge of the subject matter" the highest in 100-level (M = 3.55), 200-level (M = 3.54), 300-level (M = 3.56), 400-level (M = 3.55), 500-level (M = 3.66), 600-level (M = 3.77), and 800-level (M = 3.65) classes. As Table 4 indicates, students who filled out course evaluations in 700-level classes rated the instructor item "the instructor was available for consultation outside of class during office hours" the highest (M = 3.60) while students enrolled in 900-level classes rated the learning outcomes item "the course strengthened my ability to analyze and evaluate information" the highest (M = 3.50).

Students enrolled in 100 and 500-level courses rated the course item "the textbook(s) contributed to my understanding of the subject" and the learning outcomes item "the course stimulated me to read further in the area" the lowest with mean scores of 2.97 and 3.27, respectively. The course item "graded assignments included helpful comments from the instructor" was the lowest rated (M = 3.04) item for students enrolled in 200-level classes. Students enrolled in 300-level classes rated the learning outcomes item "the course stimulated me to read further in the area" the lowest (M = 3.13) while student enrolled in 700-level classes rated the course item "examinations reflected what was taught in the course" the lowest (M = 3.03). The course item "the textbook(s) contributed to my understanding of the subject" was rated the lowest by students enrolled in both 400 and 600-level classes with mean scores of 3.05 and 3.31, respectively. Students enrolled in 800 and 900-level courses rated the course item "graded assignments, tests, etc., were returned promptly" the lowest with mean scores of 2.92 and 2.33, respectively (Tables 3 and 4).

Students who reported they studied ≤ 1 hour, 2 hours, 3 hours, 4-5 hours, 6-7 hours, or ≥ 8 hours per

Table 3. Course Evaluation Items Compared by Course Level 100-500

	100			200			300			400			500		
Course Evaluation Item	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Rate the overall value of course.	837	3.14	.77	940	3.32	.80	1013	3.30	.78	753	3.25	.81	562	3.39	.76
Rate the overall quality of teaching by the primary instructor in this course.	838	3.32	.75	933	3.29	.90	1007	3.37	.80	752	3.42	.76	560	3.40	.82

Note. Scale for "summary" items: 1= Poor, 2=Fair, 3=Good, 4=Excellent

Table 4. Course Evaluation Items Compared by Course Level 600-900

	600			700			800			900		
Course Evaluation Item	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Rate the overall value of course.	162	3.59	.54	85	3.36	.80	140	3.39	.65	12	3.08	.67
Rate the overall quality of teaching by the primary instructor in this course.	161	3.61	.59	84	3.44	.75	139	3.29	.72	12	3.17	.5

Note. Scale for "summary" items: 1= Poor, 2=Fair, 3=Good, 4=Excellent

Table 5. Course Evaluation Items Compared by Hours per Week Studied

	≤1 Hour			2 Hours			3 Hours			4-5 Hours			6-7 Hours			≥8 Hours		
Course Evaluation Item	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Rate the overall value of course.	1461	3.16	.80	1209	3.34	.74	845	3.39	.76	413	3.33	.79	94	3.38	.84	152	3.50	.66
Rate the overall quality of teaching by the primary instructor in this course.	1458	3.29	.81	1204	3.41	.77	838	3.43	.82	413	3.36	.81	94	3.44	.80	150	3.42	.75

Note. Scale for "summary" items: 1= Poor, 2=Fair, 3=Good, 4=Excellent

week, rated the instructor item "the instructor had a good knowledge of the subject matter" the highest with mean scores of 3.51, 3.58, 3.64, 3.60, 3.68, and 3.65, respectively. Students who reported studying ≤1 hour, 2 hours, 4-5 hours, and 6-7 hours per week rated the learning outcomes item "the course stimulated me to read further in the area" the lowest with mean scores of 2.97, 3.15, 3.18, and 3.22, respectively. Students who reported studying for 3 hours per week rated the course item "graded assignments included helpful comments from the instructor" and the learning outcomes item "the course stimulated me to read further in the area" the lowest with a mean score of 3.26 for both items. The course item "graded assignments, tests, etc., were returned promptly" was rated the lowest by students who reported studying ≥8 hours per week (Table 5).

As indicated in Table 6, the instructor item "the instructor had a good knowledge of the subject matter" had the highest mean score for students who reported they expected to receive an "A" (M = 3.62), "B" (M = 3.53), "C" (M = 3.44), "D" (M = 3.37), or pass/audit (M = 3.73). The learning outcomes item "the course stimulated me to read further in the area" had the lowest mean scores for students who reported they expected to receive an "A" (M = 3.19), "B" (M = 3.05), or "C" (M = 2.84). Students who reported their expected grade to be a "D" rated the instructor item "the instructor stimulated my interest in the subject" the lowest (M = 2.68). Students who reported that they either "passed" or "audited" the course rated the course item "the textbook(s) contributed to my understanding of the subject" the lowest with a mean score of 3.30. There were not enough students who reported they expected an "E/Fail" (n = 4) or "I" (n = 1) that provided meaningful data.

ships with value of the course was the textbook's contribution to understanding and promptness of grades items. This implies that not all individual course evaluation items have the same relationship in regard to the overall value of the course. Therefore, when faculty reflect about their overall value score, they should take into consideration that students' scores are influenced by how much students are stimulated by the course and their effectiveness as a presenter or teacher. Juxtaposed, faculty should note that textbook selection and promptness of returned papers are less influential to the overall value of the course score. This should not indicate that less attention should be placed in these areas as very high relationships were found, but that students value them less in rating the overall value of the course.

When rating the quality of teaching, students' scores were influenced more by the feedback students received from graded assignments, followed by the amount the course stimulated interest in the subject for the students. Instructor's ability to answer questions also related highly with students' scores on quality teaching. Again, having the weakest relationships with value of the course was the textbook's contribution to understanding and promptness of graded items. This implies that some of the individual course evaluation items had stronger relationships than others to the overall value of the course. Therefore, when faculty reflect about their overall teaching quality, they should take into consideration that students' ratings are influenced by comments on graded work and the amount the course stimulated interest. Again, faculty should note that textbook selection and promptness of returned papers are less influential to the overall value of the course score. This should not indicate that less attention should be placed in these areas as very high relationships were found, but that students value them less in rating the overall value of the course.

Table 6. Course Evaluation Items Compared by Expected Grade

	A			B			C			D			P/A		
Course Evaluation Item	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Rate the overall value of course.	2329	3.37	.74	1421	3.23	.79	295	2.98	.87	19	2.74	.81	88	3.50	.70
Rate the overall quality of teaching by the primary instructor in this course.	2321	3.44	.75	1416	3.28	.83	293	3.08	.94	19	2.95	.91	88	3.67	.64

Note. Scale for "summary" items: 1= Poor, 2=Fair, 3=Good, 4=Excellent; For categories "E/Fail" and "Incomplete", not enough respondents for meaningful data

Discussion

From the findings it is concluded that the evaluations derived mostly from 300-level and 200-level classes, with 900-level representing the least represented. This would be characteristic representing the number of undergraduate versus graduate/professional students at any university. A majority of the students who completed evaluations expected an "A" in the course and were taking the course because it was a part of their major.

Students were more likely to rate the value of the course higher when a course stimulated interest in the subject for the students. Value of the course was also associated with course materials being presented in an effective manner. Having the weakest relationships with value of the course was the textbook's contribution to understanding and promptness of grades items. This implies that not all individual course evaluation items have the same relationship in regard to the overall value of the course. Therefore, when faculty reflect about their overall value score, they should take into consideration that students' scores are influenced by how much students are stimulated by the course and their effectiveness as a presenter or teacher. Juxtaposed, faculty should note that textbook selection and promptness of returned papers are less influential to the overall value of the course score. This should not indicate that less attention should be placed in these areas as very high relationships were found, but that students value them less in rating the overall value of the course.

Beyond the practical need for reflective teachers, it is also recommended that university administration should note what aspects of teaching and of the course influence the overall value of the course and quality of teaching scores from students. Because these summary items are typically reported in dossiers and review documents, administrators need a firm understanding of their meaning. This recommendation is also supported by the findings of Onwuegbuzie et al. (2007) that questioned the construct-related validity of course evaluations.

When comparing the course evaluation items by course-level, hours spent studying per week, and grade expected, some differences were noted. For

example, when comparing summary item mean scores by course level, students in 600-level course tend to rate the course and instructor higher than any other level. This raises questions as to why this level solicits such scores. The 700-level courses, also noted as graduate-level only classes, had summary mean scores that were similar to the 200- and 300-level courses. It is recommended to examine the courses taught at these levels and perhaps the student and course demographics to determine if patterns exist. Some caution should be noted, as there were only approximately 80 to 160 responses for the 600- and 700-level courses compared to the 100- to 500-level courses which ranged from 500 to 1000 responses.

Students who reported studying one hour or less per week for their course rates the course and teacher lower than students in other categories. One explanation could be that students who spend less time on the course rate instructors and the course lower because they are unhappy with the amount of work instructors expect. Another possibility is that these students rate the course and instructor lower because the course did not require more than one hour of work and the students feel they were not pushed to their potential. When reflecting on their course evaluations, instructors should consider this trend and think about the amount of work they expect in a class juxtaposed with the work ethic and attitude of their students. Other hypotheses on the subject could be developed, but this trend should be investigated further.

More students believe they will receive a grade of A than any other grade. Those students also rate the teacher and course higher than any other grade category. This supports the findings of Krautmann and Sander (1999) and implies that students' attitude about their level of performance has some bearing on how they rate their instructors and the course. When instructors reflect upon their distribution of grades, they should also keep in mind what bearing this has on their course evaluations. This is not a call for lowered standards or grade inflation. However, instructors who have reasonably high standards where the grades distribute across the scale may find lower evaluations. It is also recommended to take this aspect of the investigation further by comparing course evaluations with actual grade received.

In general, teachers should use the course evaluations as literature supports (AAUP, 2006; Cruse, 1987; Thompson and Serra, 2005). However, in using the course evaluation as a reflection tool, the tool and its use must be understood. It is recommended that faculty analyze and/or ask questions regarding their own course evaluations. The findings to this study were specific to the College of Agriculture at the University of Kentucky; findings could differ elsewhere. Being critical about a tool that is used beyond the scope of reflection (i.e. for review) should be an activity undertaken by the faculty for which the results are being affected.

Literature Cited

AAUP. 2005. Observations on the association's 1975 statement on teaching evaluation. http://www.aaup.org/AAUP/About/committees/committee+repts/teach-eval-obs.htm?PF=1 (Accessed March 13, 2007)

AAUP. 2006. Policy Documents and Reports, 10th ed. Baltimore, Maryland: Johns Hopkins University Press.

Best, J.B. and W.E. Addison. 2000. A preliminary study of perceived warmth of professor and student evaluations. Teaching of Psychology 27(1): 60-62.

Boyer, E.L. 1990. Scholarship reconsidered. New York: Jossey-Bass.

Chen, Y. and L.B. Hoshower. 2003. Student evaluation of teaching effectiveness: An assessment of student perception and motivation. Assessment and Evaluation in Higher Education 28(1): 71-88.

Coleman, J. and W.J. McKeachie. 1981. Effects of instructor/course evaluations on student course selection. Jour. of Educational Psychology 73(2): 224-226.

Cruse, D.B. 1987. Student evaluations and the university professor: Caveat professor. Higher Education 16(6): 723-737.

Felder, R. 1992. What do they know, anyway? Chemical Engineering Education 26(3): 134-135.

Gramlich, E.M. and G.A. Greenlee. 1993. Measuring teacher performance. Jour. of Economic Education 24(1): 3-13.

Krautmann, A.C. and W. Sander. 1999. Grades and student evaluations of teachers. Economics of Education Rev. 18: 59-63.

Marlin, J.W., Jr. 1987. Student perception of end-of-course evaluations. Jour. of Higher Education 58: 704-716.

Office of Intuitional Research. 2007. Teacher/course evaluations. http://www.uky.edu/IR/tce.html (Accessed March 20, 2007).

Onwuegbuzie, A.J., A.E. Witcher, K.M.T. Collins, J.D. Filer, C.D. Wiedmaier, and C.W. Moore. 2007. Student's perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. American Educational Research Jour. 44(1): 113-160.

Ouallal-McRiffey. 2005. How do you rate those instructors? The validity of student evaluations

of higher education instructors. http://idserver.utk.edu/?id=200600000000123. Electronic Thesis and Dissertations. (Accessed March 21, 2007).

Rodin, M. and B. Rodin. 1973. Student evaluations of teachers. The Jour. of Economic Education: 5-9.

Schön, D.A. 1983. The reflective practitioner. How professionals think in action. Library of Congress Cataloging in Publication Date. New York, NY: Basic Books, Inc.

Thompson, R.J., Jr., and M. Serra. 2005. Use of course evaluations to assess the contributions of curricular and pedagogical initiatives to undergraduate general education learning objectives. Education 125(4): 693-701.

Tricker, T., M. Rangecroft, and P. Long. 2005. Bridging the gap: An alternative tool for course evaluation. Open Learning 20(2): 185-192.

Wollert, M.H. and R.F. West. 2000. Differences in student ratings of instructional effectiveness based on the demographic and academic characteristics of instructors. In: Annu. Mtg. of the Mid-south Educational Research Association.

York-Barr, J., W.A. Sommers, G.S. Ghere, and J. Montie. 2001. Reflective practice to improve schools: An action guide for educators. Thousand Oaks, CA: Corwin Press.

"Advancing the scholarship of teaching and learning"