Students’ Perception of a Discussion-Driven Classroom Environment in an Upper-Level Ruminant Nutrition Course with Small Enrollment

M. A. Wattiaux*1 and P. Crump†
*Department of Dairy Science, and
†Department of Computing and Biometry, University of Wisconsin, Madison 53706

ABSTRACT
This 3-yr classroom research project studied students' perception of the learning environment when class time was reserved primarily for discussion of pre-assigned reading material as an alternative to a lecture in ruminant nutrition. A student-assessment-of-learning-gain instrument with 46 items was administered anonymously as a survey 3-wk into the semester and as a course evaluation at the end of the semester. Scores, collected on a scale of 1 to 10, were analyzed to determine differences between survey and evaluation, the variation due to student cohort (years 2003, 2004, and 2005), and student standing (undergraduate or graduate) for 5 selected items: item 1 = My level of interest/curiosity for nutrition is; item 17 = The reading assignments help me learn; item 18 = The in-class discussions help me learn; item 24 = I would prefer the professor lecture; item 38 = I am learning a lot in this class.

Although students indicated a preference to using class time for discussion rather than lecture, the degree of consensus was low, as 10% of scores indicated a strong desire for lecture, whereas another 10% indicated strong opposition to lecture. Reading assignments and in-class discussions contributed positively to the perception of learning and both were correlated positively with level of interest in the topic. In the evaluation, the desire for lecture was correlated negatively with level of interest in the topic and the self-assessed level of learning. Students reported a greater interest in the topic at the end of the semester than at wk 3; however, the self-reported learning gains from discussion-related activities varied with student cohort. Graduate students reported a higher interest, more learning from discussion-related activities, and less desire for lecture than undergraduates. Although graduate students were more impervious to the discussion format, undergraduates perceived more learning if the discussion was student-centered (i.e., focused on their questions and answers), rather than subject-centered (i.e., focused on the paper content).

Key words: teaching method, classroom discussion, scholarship of teaching, student learning

INTRODUCTION
Lecture remains by far the major form of professor-student interaction in higher education classrooms (Diamond, 1998). The teacher-centered, 50-min lecture presentation with PowerPoint slides may be an efficient mode of content delivery, but may be less effective in promoting active engagement of students with either the instructor or the topic. Students have almost no opportunity to become actively engaged, and typically, little class time is spent on higher-level thinking skills when lecture is used as the primary form of interaction in the classroom (López and Whittington, 2001). Instructors have been encouraged to broaden their teaching methods to accommodate for the diversity of students' learning styles (Hartel, 1995). Although institutional efforts have facilitated the transition away from a teacher-centered classroom to a more student-friendly learning environment (Whittington, 1999; McMahon, 2000), what happens within a classroom remains essentially the prerogative of the instructor (Kreber, 2002).

In contrast to lecture, classroom discussion could contribute to each of the 7 principles of good practice in undergraduate education (Chickering and Gamson, 1991). Depending on its design and implementation, discussion provides for more instructor-student interaction, and reciprocity and cooperation among students. It is an active teaching technique that emphasizes time on task, allows for prompt feedback, helps set high expectations, and respects diversity and ways of learning. Classroom discussion may take many forms with the instructor being a leader, a facilitator, or at times, simply an observer (Palmer, 1998). Recent reports have documented success in using discussion as the only mode of instruction (Glazer, 2000; Bradley et al., 2002; Cameron, 2003). However, abandoning lecture entirely may not be advisable in all circumstances and the implementation of discussion as a teaching method is not always a guaranteed success (Palmer, 1998). Students’
or instructors’ inertia, lack of experience, and the need to cover material are only a few of the impediments of incorporating discussion in the classroom (Brookfield and Preskill, 1999). Thus, our objective was to describe attempts to develop an effective discussion-based classroom environment, and to study students’ attitude and perception of the learning environment when class time was reserved primarily for discussion of preassigned reading material.

**MATERIALS AND METHODS**

**The Class: Ruminant Nutrition 414**

Ruminant Nutrition 414 is a 2-credit “lecture” course, which is the third and last component of the nutrition curriculum for dairy science majors at the University of Wisconsin-Madison. As indicated by the syllabus, course objectives are for students to learn nutritional concepts and their implications in feeding ruminants and to develop analytical and critical thinking skills as related to ruminant nutrition. The first author has taught this course since 1999, initially with a heavy reliance on PowerPoint presentations. However, from 2003 to 2005, lecture presentations were essentially abandoned. Class time was structured instead with a variety of activities meant to engage students with the course material including discussions of preassigned readings and case studies, computer laboratories, and student presentations of a class project, which consisted of collecting and analyzing dairy cattle rations from commercial farms. The relative weight of each activity varied somewhat from year to year in response to students’ input, but the topics remained relatively constant. Typically, the first half of the semester focused on principles of ruminant nutrition, including current systems of feed evaluation and use by dairy cattle. The second half of the semester was geared toward real-life application of feeding practices with an emphasis on understanding and solving problems associated with metabolic disorders such as milk fever, displaced abomasum, rumen acidosis, and ketosis, and issues related to calf and heifer nutrition (a complete course schedule can be found at http://www.dairynutrient.wisc.edu/414). Typical enrollment, which included a mix of undergraduate and graduate students, varied from 15 to 24 in the last 3 yr (Table 1).

**Use of Class Time, 2003.** In the first year of the study, PowerPoint presentations were discontinued and replaced with discussions of preassigned reading material. Papers (5 to 10 pages long) were selected carefully from proceedings of nutritional conferences (e.g., TriState Nutrition Conference, 4-State Nutrition and Management Conference), the NRC (2001), and the *Journal of Dairy Science*. An effort was made to select papers of appropriate difficulty with a degree of overlap and connection. A hard copy of the readings was distributed at least 1 wk before the scheduled discussion. To help students prepare for meaningful discussion, they were requested to read the paper and e-mail the instructor at least 2 multiple-choice and 2 mini-essay questions (with or without answers) focusing on a new concept, principle, or application presented in the reading. After grouping the questions into themes, a handout (without answers) was printed, photocopied, and used for in-class discussion. After class, the material discussed was posted on the course Web site.

**Use of Class Time, 2004.** Use of class time in 2004 resembled that in 2003 with a few important exceptions. Students were responsible for downloading readings from the course Web site. In addition, a set of 15 to 20 true/false and multiple-choice questions were created and posted online for each reading. Some of the questions were from the 2003 student cohort, but others were added to guide the students’ attention to the main points of the paper. Everyone in the class had to take an online quiz, print it, and bring it to class as a proof of preparation. Completion of the quizzes contributed 10% of the final grade. Furthermore, students were placed in teams of 2 or 3 early in the semester, and each team took their turn in leading the discussion. Team members met with the instructor a few days before the scheduled discussion to identify the parts of the paper worthy of in-depth attention and to discuss ways to facilitate the in-class discussion with the rest of the class. Most teams chose to prepare a handout with key points and questions as a way to structure the discussion. After every team had its first turn, an informal survey indicated that students were uncomfortable leading discussions and needed more guidance to understand the content of papers and the connection among them. As a result, the instructor and a teaching assistant led the discussions for the second half of the semester with a handout of thought-provoking questions drawn from the expanded online quiz.

**Use of Class Time, 2005.** In 2005, attempts were made to improve student preparation and quality of discussions, and to help students make connections across readings. Compared with 2004, the major modifications to the learning environment were as follows. First, the number of readings was decreased with some papers being discussed over multiple periods. Second, at the beginning of each discussion period, a quiz of 10 randomly generated multiple-choice questions from the online quiz was administered. Third, projection equipment was used as needed for discussion of data presented in the tables and figures of the paper. Fourth, 5- to 10-min “mini-lectures” were introduced periodically to either “fill the gap” between papers or help...
students integrate a series of papers. Fifth, the students and instructor did not sit in a circle (as in 2004) or around a conference room table (as in 2003), but the physical layout of the classroom was more traditional, with all students sitting in rows facing the instructor. The latter modification was to assert the instructor’s role as a leader of class discussion.

### Experimental Approach

**The Instrument.** A series of Likert-type items can be used to evaluate instructional effectiveness along distinct mental dimensions or constructs in a single instrument (Brown, 1983; Abrami and d’Apollonia, 1990). In this study, the Student Assessment of Learning Gains (SALG; Seymour, 1997) was modified to explore the learning environment in the class by assessing students’ perception of their interest in the topic, their sense of the relevance and value of course materials, their learning gains as a result of various class activities, the adequacy of the grading system and the overall impression of the class. For each of these aspects, a set of 6 to 12 statements (referred to herein as items) was formulated resulting in a total of 46 items in the instrument. Scores for each item were collected on a response-scale of 1 to 10. Five descriptive categories were used to assist students in understanding the response-scale (Table 2). The cover page of the instrument included a short overview of the project and requested student descriptors including current standing (undergraduate or graduate) and reason for taking the course (required or elective). In addition, students were requested to identify the instrument with the first 2 letters of the first names of their mother and father as a way to identify the instrument anonymously. A consent form was prepared to provide students with additional details about the project and their right to be excluded from the study. Before initiating the study, both the instrument and the consent form were approved by the UW-Madison Institutional Review Board for research with human subjects.

**Data Collection.** The instrument was administered as a class survey 3 wk into the semester and again as a course evaluation conducted the last week of the semester. Students were allowed to complete the survey outside class, but the evaluation was administered with ample time for completion during the last day of class. The modes of administration were selected as a compromise to avoid time pressure on the student while maintaining the instructor’s ability to collect the completed instruments. Survey and evaluation instruments matched unequivocally by the anonymous identifier were included in the study. Items pertaining to the objective of the study were selected for detailed analyses (Table 2). Item 1 was directed at student’s motivation. The next 2 items focused on learning gains from the preclass reading assignments (item 17) and the in-class discussion (item 18). Item 24 measured student’s desire for lecture and item 38 addressed student’s perception of overall learning in the class.

### Statistical Analyses

Scores were analyzed according to a randomized complete block design. Each student was considered a block,

---

### Table 1. Class enrollment and study enrollment

<table>
<thead>
<tr>
<th>Year</th>
<th>Undergraduate</th>
<th>Graduate</th>
<th>Total</th>
<th>Undergraduate</th>
<th>Graduate</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>12</td>
<td>3</td>
<td>15</td>
<td>6</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>2004</td>
<td>16</td>
<td>2</td>
<td>18</td>
<td>8</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>2005</td>
<td>18</td>
<td>6</td>
<td>24</td>
<td>16</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td>11</td>
<td>57</td>
<td>30</td>
<td>11</td>
<td>41</td>
</tr>
</tbody>
</table>

1Number of survey and evaluation instruments for which student identification was matched unequivocally.
2Juniors and seniors (except 1 sophomore in 2005).

---

### Table 2. Items and response alternatives used to collect students’ perception

<table>
<thead>
<tr>
<th>Item</th>
<th>Statement</th>
<th>Not at all</th>
<th>A little</th>
<th>Somewhat</th>
<th>A lot</th>
<th>A great deal</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My level of interest/curiosity for nutrition is:</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>17</td>
<td>The reading assignments help me learn</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>18</td>
<td>The in-class discussions help me learn</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>24</td>
<td>I would prefer the professor lecture</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>38</td>
<td>I am learning a lot in this class</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

1Item number reflects the order in which it appeared in the instrument.
and administrations of the instrument (wk-3 survey and wk-15 evaluation) were considered treatment effects. Characteristics of data distribution were explored with the UNIVARIATE and FREQ procedures of SAS (SAS Institute, 1999). The percentage of observations in each of the 10 possible scores were plotted to inspect the distribution function of each item. Both parametric and nonparametric techniques (e.g., Mann-Whitney or Wilcoxon tests) have been used to analyze individual Likert-type items (Clason and Dormody, 1994). The relative efficiency and adequacy of these statistics depend upon data distribution and assumptions regarding the distribution of the error term. Parametric methods require that residual errors be normally distributed. In contrast, nonparametric methods, which are typically performed on ranked scores rather than actual scores, remain valid regardless of the population distribution function (Conover, 1999). In this study, analyses were performed on ranked values as an equivalent to the Wilcoxon test (SAS Institute, 1999). However, ANOVA of actual scores and ranked scores yielded the same results. Thus for simplicity, results of the former analyses will be presented here.

Differences among items were studied separately for survey and evaluation scores with the GLM procedure (SAS Institute, 1999) and the following model (model 1):

\[ \text{Score}_{ik} = S_{ti} + I_{tk} + e_{ik}; \]

where \( \text{Score}_{ik} \) (\( n = 205 \)) was the dependent variable, \( S_{ti} \) identified student \( i \) (\( i = 1 \) to 41), \( I_{tk} \) identified item \( k \) (\( k = 1 \) to 5), and \( e_{ik} \) was the error term. The independence of the residuals in the error term was confirmed after using the MIXED procedure with various covariance structures for the repeated measurements on the same student. After a significant \( F \)-test (\( P < 0.05 \)), item mean scores were separated with Fisher’s LSD (\( P < 0.05 \)).

The association among items was evaluated with the Spearman correlation. In addition, the degree with which each item was tapping into a common mental construct was measured with the corrected item-total correlation (CITC) and the Cronbach alpha coefficient (CAC; Cronbach, 1951) using the CORR procedure of SAS (1999). The CITC is the correlation between an item score and the sum score across all items, excluding the studied item. In contrast, the CAC ranges from 0 to 1.0 and measures internal consistency for a composite score (Nunnally and Bernstein, 1994). The higher the CAC is, the higher the indication that taken collectively, the items in the data set represent a particular mental construct reliably. Low CAC reflects either an insufficient number of items or the fact that items have very little in common. In addition, the degree with which an item contributed to internal consistency was estimated by comparing CAC obtained with all items in the data set to the CAC obtained after deleting an item. If the deletion caused an increase in CAC, then the remaining items were more internally consistent, indicating that the deleted item contributed less than the others taken together to internal consistency and the associated mental construct (SAS Institute, 1999).

The model used to analyze the difference in students’ response between the survey and the evaluation for each item was as follows (model 2):

\[ \text{Score}_{ij} = S_{ti} + \text{SuEv}_{j} + e_{ij}; \]

where \( \text{Score}_{ij} \) (\( n = 82, 81, 64, 62, \) and \( 81 \) for items 1, 17, 18, 24, and 38, respectively) was the dependent variable, \( S_{ti} \) identified student \( i \) (\( i = 1 \) to 41), \( \text{SuEv}_{j} \) identified whether the score was from the survey or the evaluation (\( j = 1 \) to 2), and \( e_{ij} \) was the error term. Single degree of freedom contrasts were used to separate the means of student cohort (years 2003, 2004, and 2005), and student standing (undergraduate vs. graduate). Significance was declared for \( P < 0.05 \) and tendencies for \( 0.05 \leq P < 0.10 \). Furthermore, model 2 was modified to explore whether the change in scores between survey and evaluation varied among student cohorts or between student standings. These interactions were added to the model after the variation due to student (40 df) was partitioned into a year effect (2 df), a student standing effect (1 df), and a remaining student blocking factor (37 df). Interaction terms with \( P \leq 0.12 \) were explored further with separate analyses by yearly student cohort or by student standing with model 2.

RESULTS AND DISCUSSION

Enrollment

Forty-one surveys and evaluations were unequivocally paired with data coming from 30 undergraduate students and 11 graduate students in the 3 yr of the study. Study enrollment averaged 72%, but was 60, 56, and 92% of class enrollment for 2003, 2004, and 2005, respectively (Table 1). One student declined participation in 2004. Moreover, reliance on students’ goodwill without an established enforcement strategy for collection of anonymous data contributed in part to lower study enrollment in the first 2 yr of the study. There was no evidence for nonrespondent bias, but simply procrastination in completing the survey by some students in 2003 and 2004. In turn, the greater success in collecting paired-data in 2005 was associated with additional time spent explaining the intent and protocol of the study, and the involvement of a teaching assistant who checked students from a list when the survey
was turned in, verified and enforced instructions for anonymous identification, and reminded procrastinators of the importance of their collaboration.

**Data Distribution and Alternatives in the Response Scale**

The frequency of scores in the 10-response alternatives for the sum of the 5 items increased gradually with score’s ordinal value, but plateaued at approximately 16% for score 7 and greater (Figure 1A). The distribution characteristics of the data varied considerably among items (Figure 1B to 1F), among students, and to a lesser degree between survey and evaluation (data not shown). In this study, it was our intent to capture the widest possible range in attitude and to minimize the number of ties by using 10 response alternatives. Scores for Likert-type items are collected most often with 7-response alternatives using qualifiers such as “strongly” or “slightly” to differentiate adjacent choices (Clason and Dormody, 1994). In this study, the subdivision of 5 descriptive categories into 10 numerical values may have allowed for meaningful differentiation of each response alternative and possibly for more constant intervals along the mental scale of the respondent.

**Student Effects**

Some students are expected to score consistently high and others consistently low. To identify possible outliers (i.e., a student who might have intentionally scored low or high regardless of item), the average score of each student was calculated for all items without a single missing observation (n = 23 items). Average (survey and evaluation) scores ranged from 4.9 to 9.7. Inspection of the data sorted in descending order indicated that no student departed from an overall linear trend (Figure 2). In addition, there was no difference between evaluation and survey scores (0.06 ± 1.17; mean ± SD). However, this difference ranged from −2.26 to +2.65 among students. The sign of the difference appeared relatively evenly distributed across students’ average score. As expected, larger differences seemed to occur as scores
declined; however, the largest differences were not observed in the few students near the lowest score (Figure 2).

Comparison and Relationships Among Items

Level of Agreement. A mean score provides a numerical measure of students’ level of agreement with an item statement. Results suggested a preference for using class time for discussion rather than for lecture (7.6 vs. 5.6 and 7.8 vs. 6.2 for item 18 vs. item 24 in survey and evaluation, respectively; Table 3). In the survey, the self-reported learning gain from reading assignments (item 17) was numerically lower than for in-class discussion (item 18); however, in the evaluation, the difference was significant. These mean scores should be interpreted carefully because they may not represent “the voice of a majority,” especially if the data are not normally distributed and because they fail to account for possible diverging opinions among students.

Degree of Consensus. The degree of consensus refers to the degree with which students were united or polarized in answering an item. It can be measured with the interquartile range (IQR), which contains the middle 50% of the responses and excludes scores on either end (Kreber, 2003). A larger IQR reflects greater data dispersion and lower degree of consensus. In the survey, IQR was 3 for most items, but 5 for item 24, indicating that the desire for lecture was the most polarizing item (Table 3). The degree of consensus was somewhat higher for items 1 and 24, but lower for items 17, 18, and 38 for the evaluation compared with the survey. The divergence of opinion among students for item 24 was apparent in Figure 1e, which showed that 10% of the combined survey and evaluation scores pointed to a strong opposition to using class time for lecture (score of 1), but another 10% pointed to a strong preference for lecture (score of 10), whereas no single category received more than 15% of total scores.

Correlation Among Items. In both the survey and the evaluation, students’ perception of overall learning (item 38) was associated positively with their level of curiosity and interest (item 1), the reading assignments (item 17), and the in-class discussions (item 18, Table 3). In addition, there were negative correlations in the evaluation data. Students who indicated higher preference for lecture scored lower for level of interest and curiosity in the topic (r = −0.36, Table 3). These students were also among those who indicated the least benefits from reading assignments (r = −0.38), and the least overall learning in the class (r = −0.40). These results
Table 3. Mean scores (± standard deviation), Spearman correlations, and Cronbach alpha coefficient (CAC) and corrected item-total correlation (CITC) for survey and evaluation scores of items measured on a scale of 1 (not at all) to 10 (a great deal)

<table>
<thead>
<tr>
<th>Item</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
<th>IGR</th>
<th>Score</th>
<th>Spearman correlation</th>
<th>Item 1</th>
<th>Item 17</th>
<th>Item 18</th>
<th>Item 24</th>
<th>CAC</th>
<th>CITC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey (wk 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>41</td>
<td>7.3a</td>
<td>2.29</td>
<td>3</td>
<td>—</td>
<td>0.54</td>
<td>0.34†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>40</td>
<td>6.7a</td>
<td>2.29</td>
<td>3.5</td>
<td>0.43**</td>
<td>0.43</td>
<td>0.54**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>32</td>
<td>7.6a</td>
<td>2.01</td>
<td>3</td>
<td>0.47**</td>
<td>0.46**</td>
<td>0.50</td>
<td>0.42*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>31</td>
<td>5.6b</td>
<td>2.93</td>
<td>5</td>
<td>−0.28</td>
<td>−0.02</td>
<td>0.79</td>
<td>−0.18</td>
<td>0.32</td>
<td>0.70**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>41</td>
<td>7.1a</td>
<td>2.17</td>
<td>3</td>
<td>0.57**</td>
<td>0.66**</td>
<td>0.69**</td>
<td>−0.12</td>
<td>—</td>
<td>0.70**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evaluation (wk 15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>41</td>
<td>7.9a</td>
<td>1.71</td>
<td>2</td>
<td>—</td>
<td>0.38</td>
<td>0.40*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>41</td>
<td>6.0b</td>
<td>2.49</td>
<td>4</td>
<td>0.43**</td>
<td>0.39</td>
<td>0.40</td>
<td>0.36*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>32</td>
<td>7.8a</td>
<td>2.05</td>
<td>4</td>
<td>0.39*</td>
<td>0.35*</td>
<td>0.29</td>
<td>0.54**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>31</td>
<td>6.2b</td>
<td>2.40</td>
<td>4</td>
<td>−0.36*</td>
<td>−0.38*</td>
<td>−0.15</td>
<td>0.80</td>
<td>−0.40*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>40</td>
<td>7.6a</td>
<td>2.06</td>
<td>4</td>
<td>0.67**</td>
<td>0.69**</td>
<td>0.70**</td>
<td>−0.40*</td>
<td>0.11</td>
<td>0.75**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Item scores within survey and evaluation without a common superscript differ (P < 0.05).

Item 1 = My level of interest/curiosity for nutrition is; item 17 = The reading assignments help me learn; item 18 = The in-class discussions help me learn; item 24 = I would prefer the professor lecture; item 38 = I am learning a lot in this class.

1IQR = Interquartile range.

CAC with no variable deleted was 0.61 and 0.52 for survey and evaluation, respectively.

†P < 0.10; *P < 0.05; **P < 0.01.

corroborated the findings of Seevers et al. (1999) indicating that the instructors’ ability to stimulate student’s interest influenced both overall course evaluation and student achievement.

**CAC and CITC.** For the survey data, the deletion of item 24 increased the CAC from 0.61 to 0.79, indicating that item 24 was contributing to a different construct than the others taken collectively. The different nature of item 24 was reflected in its low CITC (r = −0.18) compared with the other items. In contrast, item 38 had the highest CITC and was a unifying item. Most of the variance that item 38 shared with its construct was unique and when removed from the data set, the collection of the remaining items no longer covered the construct adequately (CAC = 0.32, Table 3). Items 1, 17, and 18 contributed somewhat to internal consistency. For each of these items, CITC was positive and their deletion caused a decrease in CAC, albeit to a lesser extent than when item 38 was deleted. Thus, interest in the topic and class discussion activities contributed positively, but only partially to students’ perception of overall learning gain. As pointed out by Nunnally and Bernstein (1994), there is not an absolute threshold to declare a set of items reliable descriptors of a mental construct; that depends on the objective of the study. However, identification and addition of items that reflected other student attributes or aspects of the course that contributed to learning would increase the overall CAC to a value closer to 1.0.

**Factors Influencing Individual Items**

**Item 1: Student’s Interest in Nutrition.** Students’ interest and curiosity for the topic of the course was higher in the evaluations than in the surveys (7.9 vs. 7.3, Table 4). This result suggested that, overall, the discussion format had a positive effect on students’ interest and curiosity for the topic. In addition, there were differences among student cohorts. These differences were associated in part with the proportion of graduate students in the class for whom the self-assessed level of interest and curiosity was 2.4 units higher than for the undergraduate students (Table 4). The difference in interest for the topic between undergraduate and graduate students was presumably associated with the reason to take the course. In this study, most, but not all, undergraduates took the course to fulfill a requirement, and most, but not all, graduate students took the course as an elective.

**Item 17: Learning with Reading Assignments.** Students’ perception regarding the usefulness of reading assignments to help their learning was higher in graduate students than in undergraduate students and higher in surveys than in evaluations. However, the latter difference varied among student cohorts (P = 0.04, Table 5). In 2003, the year that students were responsible for generating the in-class discussion questions as a mode of preparation, scores increased numerically over the course of the semester. However, in 2004 and 2005, the years in which students’ preparation for
Table 4. Change in scores associated with time of instrument administration within the semester [survey (Su) vs. evaluation (Ev)], student standing [undergraduate (Un) vs. graduate (Gr)], and yearly student cohort (year 2003, 2004, and 2005) for items measured on a scale of 1 (not at all) to 10 (a great deal)

<table>
<thead>
<tr>
<th>Item</th>
<th>SuEv</th>
<th>UnGr</th>
<th>Year</th>
<th>Contrast2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SuEv</td>
<td>Ev</td>
<td>Un</td>
<td>Gr</td>
</tr>
<tr>
<td></td>
<td>2003</td>
<td>2004</td>
<td>2005</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7.3b</td>
<td>7.9a</td>
<td>7.2b</td>
<td>8.6a</td>
</tr>
<tr>
<td>17</td>
<td>6.7b</td>
<td>6.0c</td>
<td>5.6b</td>
<td>8.4a</td>
</tr>
<tr>
<td>18</td>
<td>7.6a</td>
<td>7.8a</td>
<td>7.3a</td>
<td>8.8c</td>
</tr>
<tr>
<td>24</td>
<td>5.6a</td>
<td>6.2a</td>
<td>6.2a</td>
<td>4.8a</td>
</tr>
<tr>
<td>38</td>
<td>7.1a</td>
<td>7.6a</td>
<td>6.8a</td>
<td>8.7a</td>
</tr>
</tbody>
</table>

a,b Means between Su score and Ev score within yearly student cohort or student standing differed (P < 0.01).

c,d Means between Su score and Ev score within yearly student cohort or student standing differed (P < 0.05).

e,f Means between Su score and Ev score within yearly student cohort or student standing differed (P < 0.10).

1 Item 1 = My level of interest/curiosity for nutrition is; item 17 = The reading assignments help me learn; item 18 = The in-class discussions help me learn; item 24 = I would prefer the professor lecture; item 38 = I am learning a lot in this class.

2 P-value for student effect was P < 0.001 for all items, except for item 24 (P = 0.02).

3 NA = Data not available.

In-class discussion was through answering online quizzes or mini-essays, scores were lower in the evaluations compared with the surveys.

**Item 18: Learning with In-Class Discussions.** The contribution of in-class discussion to students' self-assessment of learning was higher in graduate students than in undergraduate students and higher in 2005 than in 2004 (Table 4). Thus, compared with the 2004 discussion format, the mini-lectures and other changes made in the 2005 format improved students' assessment of its effectiveness. In addition, the decrease in score between the survey and evaluation in 2004 (Table 5) may have reflected disenchantment with the discussion format despite changes made midsemester. In contrast, a positive dynamic in the class of 2005 may have contributed to the numerical increase between survey and evaluation scores.

**Item 24: Preference for Lecture.** The students' desire for instruction in the form of lecture was lower in graduate students than in undergraduate students and more pronounced in 2004 than in 2005 (Table 4). As discussed above, the desire for lecture was inversely

Table 5. Change between survey (Su) and evaluation (Ev) scores with yearly student cohort (2003, 2004, and 2005) and student standing [(undergraduate (Un) vs. graduate (Gr)] for items measured on a scale of 1 (not at all) to 10 (a great deal)

<table>
<thead>
<tr>
<th>Item</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>Un</th>
<th>Gr</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Su</td>
<td>Ev</td>
<td>Su</td>
<td>Ev</td>
<td>Su</td>
<td>Ev</td>
</tr>
<tr>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>17</td>
<td>6.1b</td>
<td>6.7a</td>
<td>6.9c</td>
<td>6.5a</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>18</td>
<td>NA2</td>
<td>NA2</td>
<td>7.4a</td>
<td>6.7a</td>
<td>8.2</td>
<td>—</td>
</tr>
<tr>
<td>24</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>5.6a</td>
<td>6.8a</td>
</tr>
<tr>
<td>38</td>
<td>6.4b</td>
<td>8.6a</td>
<td>6.9a</td>
<td>6.2a</td>
<td>7.4</td>
<td>7.7</td>
</tr>
</tbody>
</table>

a,b Means between Su score and Ev score within yearly student cohort or student standing differed (P < 0.01).

c,d Means between Su score and Ev score within yearly student cohort or student standing differed (P < 0.05).

e,f Means between Su score and Ev score within yearly student cohort or student standing differed (P < 0.10).

1 Item 1 = My level of interest/curiosity for nutrition is; item 17 = The reading assignments help me learn; item 18 = The in-class discussions help me learn; item 24 = I would prefer the professor lecture; item 38 = I am learning a lot in this class.

2 NA = Data not available.
proportional to self-assessed interest in the topic, which varied with student standing. Thus, the greater proportion of graduate students in the 2005 cohort may have contributed to a lower preference for lecture compared with the 2004 cohort. Fisher’s LSD indicated also that the desire for lecture was higher in undergraduates, but numerically lower for graduate students in the evaluation compared with the survey (Table 5).

**Item 38: Overall Learning.** Overall, students’ self-assessment of learning was higher in graduate students than in undergraduate students, higher in the evaluation than in the survey and higher in 2005 and 2003 than in 2004 (Table 4). The higher scores in the student cohorts of 2003 and 2005 were associated with a positive difference between evaluation and survey scores, but the opposite was true for the 2004 student cohort (Table 5). It was not in the design of this study to relate students’ perception of learning with a measure of student performance (such as final grade) or departmental course evaluation. However, grade distribution remained relatively unchanged, whereas evaluation of instructor’s effectiveness was much higher in 2003 and 2005 compared with 2004 (data not shown).

**Implications for Teaching Improvement in Upper-Level Undergraduate Courses**

**Student Cohort.** When cohorts of students are studied, by design or by necessity due to small enrollment as in this study, the effect should be interpreted as a series of interactions among a variety of factors. Each cohort is unique and may respond differently to the learning environment not only because of differences in student standing and level of interests as found here, but also because of variation in learning styles (Hartel, 1995), personality type, and individual response to group dynamic (peer pressure). In addition, other factors will most likely vary from year to year. For example, in this study, there were slight differences in course content, weight of various class activities, and degree of involvement of teaching assistants. Nevertheless, our results agreed with Palmer (1998), who indicated that, despite relatively small changes in format, students’ response to class discussion might vary considerably from year to year. In addition, our results highlighted the importance of considering multiple cohorts in classroom research project designed to provide instructors with insights into how students learn in their class because of their method of instruction (Cross and Steadman, 1996).

**Survey vs. Evaluation.** Educational research has suggested that students tend not to change their opinions over the “ups and downs” of a semester and therefore their evaluation should be considered reliable (Cross and Steadman, 1996). Nevertheless, despite students’ resilience in their way of scoring, the survey in this study may have captured students’ initial reaction to a novel learning environment (essentially free of lecture), whereas the evaluation may have captured a longer-term impact of having to adjust to a different set of responsibilities and expectations (being prepared for class discussion and building one’s own understanding). However, the best mode and time of administration of the instrument for data quality remain uncertain. Avoiding the last week of class for administration of the evaluation may be desirable because of end-of-semester pressure and conflict with regular course evaluation.

**Classroom Discussion.** Other studies in upper-level courses have reported success in adopting discussion with almost total and unanimous student support as measured by qualitative feedback (Glazer, 2000) or end-of-semester evaluation (Cameron, 2003). However, results of this study indicated that classroom discussion of a preassigned reading assignment was perceived as a beneficial activity primarily by students with a higher motivation as measured by level of interest and curiosity for the topic. Level of interest varied from year to year in proportion to the presence of graduate students in the class (Tables 1 and 5). In addition, the negative correlations of Table 3 indicated that students with the least interest in the topic were more likely to prefer lecture as a mode of instruction. For these students, discussion may be less appealing not only because of apparent loss of effectiveness in content delivery, class structure, organization, and use of classroom time (Brookfield and Preskill, 1999), but also in part because the path to a good grade is no longer obvious. In general, graduate students appeared to welcome the opportunity to construct their knowledge from reading and discussing papers (i.e., subject-centered discussion), but undergraduates appeared to respond more positively to a discussion-based classroom that was more student-centered. For example, the self-assessed overall learning (item 38) increased dramatically during the semester when the discussions were driven primarily by students’ own questions about the readings (2003, Table 5). Additional evidence was found in written comments of undergraduates in 2003 and 2005 indicating that classroom discussions were beneficial to the extent that students’ own questions were being answered. Some of the disenchantment with discussion-related activities (items 17 and 18) in 2004 was associated not only with less interest in the topic, but also with a discussion format that did not encourage students to submit their own questions for class discussion. Taken together, these findings suggested that discussion-based courses might be more difficult to implement successfully in
lower-level, introductory courses compared with upper-level, major-specific courses.

Dairy science majors benefit from dairy cattle judging (Guthrie and Majeskie, 1997) or the Dairy Challenge (Weber Nielsen et al., 2003) in part because of the opportunity to exercise critical thinking, decision-making, and communication skills. Designing effective classroom discussions would contribute to each one of these skills and would help students take ownership of their learning within an academic discipline (Crouch and Mazur, 2001).

CONCLUSIONS

Pairing students’ responses of an instrument administered anonymously early and late in the semester may contribute to investigating students’ response to the learning environment and changes made in a class during a semester. This study indicated that in-class discussion of preassigned reading material in an upper level nutrition course contributed positively to the learning environment. However, discussion was perceived as more beneficial to students with higher curiosity and interest in the topic. The items used in this study could be used in an early semester class survey to help instructors gauge students’ motivation. Discussion format could then be altered toward more student-centered discussions when the level of motivation is low and toward more subject-centered discussions when the level of motivation is high.

ACKNOWLEDGMENTS

The authors wish to thank Eric Nordheim (Department of Statistics and Department of Forest Ecology and Management) and James Wollack (UW-Madison testing and evaluation services) for constructive comments during the preparation of this manuscript. Partial funding was provided by USDA Higher Education Grant # 2003-38411-13465.

REFERENCES


McMahon, M. J. 2000. Student-centered learning in agriculture and natural resources at The Ohio State University. NACTA J. 44:41–47.


