

Coral L. Murrant

Advan Physiol Educ 31:180-185, 2007. doi:10.1152/advan.00061.2006

You might find this additional information useful...

Medline items on this article's topics can be found at <http://highwire.stanford.edu/lists/artbytopic.dtl> on the following topics:

- Education .. Learning Experience
- Education .. Class Size
- Education .. Creative Teaching
- Education .. Teaching Assistants

Updated information and services including high-resolution figures, can be found at:

<http://ajpadvan.physiology.org/cgi/content/full/31/2/180>

Additional material and information about *Advances in Physiology Education* can be found at:

<http://www.the-aps.org/publications/advan>

This information is current as of February 9, 2010 .

Creative teaching assistant organization to maintain an Integrative Physiology course with 440 students

Coral L. Murrant

Department of Human Health and Nutritional Science, University of Guelph, Guelph, Ontario, Canada

Submitted 7 July 2006; accepted in final form 4 February 2007

Murrant CL. Creative teaching assistant organization to maintain an Integrative Physiology course with 440 students. *Adv Physiol Educ* 31: 180–185, 2007; doi:10.1152/advan.00061.2006.—We were recently challenged with trying to maintain the integrity and learning experience of our Physiology course, which included the use of long-answer, essay-style test questions, with a class size that increased over 2 yr by ~200 students. We reorganized the teaching assistant (TA) support structure in an attempt to keep the testing style and mark (or grade) the exams accurately, in a timely fashion, and provide feedback to the students that want it. Each of four TAs became experts in two sections of the course. To assess our success, TA time allocation for specific duties was recorded. Marking (or grading) accuracy was assessed by recording test data including the number of tests returned for remarking and how much marks changed by when a grade was reassessed. Student feedback was solicited to determine whether this structure provided adequate feedback and support to the students. TAs spent an average of 115 h and 35 min \pm 7 h 21 min of a total of 140 h contracted. On average, $13.2 \pm 0.5\%$ of the tests were identified as being inaccurately graded by $4.2 \pm 0.7\%$. When asked to score whether the statement of assessment of students was fair, it scored 4.5 out of 5, where 5 equals strongly agree. When asked whether the course provided a worthwhile learning experience, the question scored 4.84 out of 5. Thus, we were successful at marking the exams accurately, in a timely fashion, and providing the necessary feedback, and we were successful at maintaining the objectives of the Physiology course with a class size of 440 students.

long-answer tests

WE WERE RECENTLY CHALLENGED with trying to maintain the integrity and objectives of our Physiology course, which focuses on integration, problem solving, and critical thinking, with a class size of 440 students. The large class size was the result of a governmental decision to decrease the number of years spent in high school from 5 to 4 yr, therefore doubling the number of students that entered university in a single year. This change increased enrollment in the class from 247 students in 2003 to 308 students in 2004 to 440 students in 2005.

The Physiology course, pioneered by Dr. Jack Barclay, is a third-year course consisting of 72 lecture hours delivered over a 12-wk semester. The course involves 6 h of lecture/wk. The course objectives include improving the following skills: critical thinking and problem-solving skills, ability to integrate material, formulate and construct logical physiological responses to challenges, and apply physiological principles to new situations. Long-answer-style tests were viewed as an integral part of achieving these objectives (1, 2). Three tests and the final exam were given at roughly 3-wk intervals. This

presented the first major problem: how to maintain using long-answer-style questions with such a large class. The second major problem was how to provide adequate support for the students to facilitate the critical thinking and problem-solving skill development demanded by the course.

The teaching assistant (TA) support structure was changed to increase the accuracy of test marking and amount of support offered to the students. Each of four TAs was assigned two sections of the course in which to become experts. In these two areas they attended lectures, read assigned readings, answered e-mails on their topics, monitored a web bulletin board, and ran question and answer discussion sessions. To assess our success, TA and professor time allocation for specific duties was recorded, marking accuracy was assessed, and student feedback was solicited regarding different aspects of their satisfaction with the course.

METHODS

Course background and student complement. The third-year human physiology course is a 72-lecture hour course spread over a 12-wk semester and is generally the students first exposure to physiology. The course is equivalent to two courses. The course's credit weighting is 1.25, which is above the normal credit weighting for two courses of 1.0. This extra credit weighting was assigned by students and is attributed to the amount of material covered and the level at which they were expected to understand and integrate the material. The students in the course are biological science students, one-half of which have unspecified majors, with the other half majoring in a wide variety of areas including human kinetics, biomedical science, nutrition, and molecular biology and genetics.

General course structure. Lectures were given every morning on Monday, Wednesday, and Friday for 50 min and Tuesday and Thursday for 80 min. Because of the compact nature of the course, the volume of information was reduced to the delivery of the appropriate information needed to be able to integrate individual systems and, ultimately, by the end of the course, to integrate systems. The course was separated into eight subject sections: 1) communication (principles), 2) communication (central nervous system), 3) communication (hormonal), 4) gastrointestinal tract, 5) blood flow (heart and vasculature), 6) blood flow (kidney), 7) blood gases (respiration), and 8) integration. Each section consisted of lectures that were designed to deliver content followed by lectures that integrated and used this content in common physiological experiences. For example, students constructed all the elements of a polysynaptic reflex after the lecture series on communication principles that included the content on membrane potential, action potentials, muscle contraction, etc. The focus of the whole section was to consider the elements of communication required in the reflex initiated when stepping on a tack. Similar approaches were taken for each section of the course. As the

Address for reprint requests and other correspondence: C. L. Murrant, Dept. of Human Health and Nutritional Sciences, Univ. of Guelph, Guelph, ON, Canada N1G 2W1 (e-mail: cmurrant@uoguelph.ca).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

complexity of systems grew throughout the course, data from human simulation computer programs were brought to class for interpretation and discussion.

Long-answer, “free-style” tests and exams were used to test for content and integration of material (including problem solving). These tests were not classified as essay examinations because writing style, sentence structure, grammar, etc., were not being assessed, and the use of point form, diagrams, and graphs were encouraged to get across relationships, ideas, and arguments. Three tests and a final exam were given after each pair of sections: *test 1* on sections 1 and 2, *test 2* on sections 3 and 4, and *test 3* on sections 5 and 6, at 3-wk intervals. The final exam covered sections 7 and 8. Tests consisted of 3 questions worth 10 marks each, and students were allowed 80 min to write (or take) each test. The final exam consisted of five questions with 120 min to write. Students were given a choice regarding the questions they could answer. Students were given legal-size paper to write their answers and were not limited to the amount of paper used per question.

Tests 1 and *2* were worth 15% of students’ final mark, *test 3* was worth 20% of the final mark, and the final exam was worth 50%. Each of *tests 1, 2,* and *3* were optional; the final exam was mandatory. Students were strongly encouraged to write *test 1*. To encourage this, students were allowed to drop their mark for *test 1* before writing (or taking) *test 2*. Students were also strongly encouraged to write *tests 2* and *3*. If students handed in *test 2* or *3* at the end of the test period and the test was marked, then they were not allowed to drop their mark but students were allowed to write any test and not hand it in. They could bring their paper in for review and feedback after the test answers were posted. The weighting of any test dropped or not written was added to the weighting of the final exam; therefore, students could have a final exam worth from 50% to 100% of their final mark depending on what they chose to do during the semester.

TA organization. Four graduate student TAs were allocated for the course, with each contracted to work 140 h. Each TA was allowed to choose a section for which they would become experts, and, based on their first choice, a second section was assigned. Section assignment ensured that one TA did not end up with two back-to-back sections and was then required to mark all of one test. The TA’s 140-h work period was contracted to be spread equally between each of the 12 wk, but this type of section assignment would not allow for this. Section assignment bundled the hours into two 3-wk periods of more intense work and two 3-wk periods of less intense work. Each TA agreed to this arrangement and actually found it more conducive to their research schedules. TAs were expected to have a depth of knowledge for each of their sections as if they had to teach the material. Their responsibilities were to attend lectures; read textbook readings; answer e-mails; monitor the web bulletin board at least once a day normally and once every few hours on the Saturday, Sunday, and Monday prior to a test (tests were always on Tuesdays); mark tests and the final exam; and run question and answer discussion sessions, which were held as two 2-h sessions following the end of each section and one 2-h session the night prior to a test.

The process of teaching the TAs to mark tests involved them marking 20 tests and then bringing them to me; I would then mark the tests while discussing my marking until we agreed on a mark. This process would go back and forth until we were both comfortable with how a question was being marked.

Correct answers (not a marking scheme) were posted on the course website, and students were able to compare their answers to the ones posted. After the answers had been posted, tests could be resubmitted for regrading simply by students indicating which question they wanted remarked and handing the test back. To encourage students to solicit feedback regarding their answers, students were told that marks would not be taken off if tests were handed back in to be regraded, and only the questions they wanted remarked would be reassessed and not the whole test. I (the professor) remarked all resubmitted tests and included feedback, in writing, on the answer in question. If further

discussion was required, students were encouraged to see me in person.

Assessment of the success of the TA reorganization. To assess the success of the reorganization of the TAs, each TA recorded daily the time spent on the various duties: attending lectures, preparation, e-mail, web bulletin board, student appointments, marking tests, and question and answer discussion sessions. Primarily, this helped to determine whether the TAs exceeded their contracted time. I also recorded the time I spent on various duties (preparation time, lecturing, time spent with TAs, marking, remarking, student appointments, office hours, e-mail, web bulletin board, time spent with students after class, administrative duties, and test preparation) to determine how much time I had spent with the TAs and where, specifically, I was spending my time.

Marking accuracy was assessed by recording test data, including the number of tests returned for remarking, how many were remarked with no mark change, how many were remarked with a mark change, and how much a mark changed when a grade was reassessed.

Whether adequate student support was provided was assessed by student feedback. As part of the normal course evaluation at the end of the year (prior to their final exam), students were given 12 statements to rank from 1 to 5, where 1 = strongly disagree and 5 = strongly agree. To try to assess whether the students thought that their assessment was fair and whether they were provided enough support/feedback during the course, questions relevant to these issues were evaluated. The relevant statements used were as follows:

1. Students were free to ask questions, disagree, and express their own opinions.
2. The instructor was available to students outside class time for discussion.
3. The assessment of students was fair.
4. Considering everything, this course has provided a worthwhile learning experience.

RESULTS

Marking accuracy. All tests were marked and returned to students within 1 wk. The statistics for each test and the final exam are shown in *Table 1*. For *test 1*, ~5.8% of tests had a grade change by 5.3% following resubmission; for *test 2*, ~14.1% of tests had a grade change by 4.0%; and for *test 3*, 13.3% of tests had a grade change by 3.3%. The percentage of tests regraded for *test 1* was lower than other tests as students were able to drop their mark for this test and many did not resubmit as they were going to drop the test regardless. If the number of people that dropped *test 1* are removed from this group, the percentage of tests regraded that had marks changed increased to 12.3%. Therefore, on average, $13.2 \pm 0.5\%$ of the tests were resubmitted for regrading and identified as being inaccurately graded by $4.2 \pm 0.7\%$, indicating 95% accuracy on 13% of tests while the remaining 87% of tests were not resubmitted for regrading.

The number of students writing (or taking) a final exam worth 50% was 34 (7.7%), a final exam worth 65% was 66 (15.0%), a final exam worth 70% was 34 (7.7%), a final exam worth 80% was 31 (7.0%), a final exam worth 85% was 113 (25.6%), and a final exam worth 100% was 163 (37.0%). Final grades had to be submitted to the administration in their final form 1 wk later; therefore, practically, there were 6 days to mark the final exam and calculate final grades. This time constraint imposed our largest pressure as all students wrote the final exam. The final exam was not returned to the students; therefore, they could not be relied upon to find errors. Thus, exams had to be marked more carefully, and exam totals were

Table 1. Tests and final exam graded and regraded

	Test 1	Test 2	Test 3	Final Exam
Test graded out of (marks)	30	30	30	50
No. of tests graded	382	170	143	440*
No. of tests resubmitted for addition errors	6	15	1	0†
Average addition errors (marks)	2.1 ± 1.6	1.2 ± 0.2	0.5	0
No. of tests resubmitted for regrading	32	35	31	40
No. of tests with no grade change	10	11	12	26
No. of tests with a grade change	22	24	19	14
Average grade change (marks)	1.6 ± 0.1	1.2 ± 0.2	1.0 ± 0.2	1.2 ± 0.5
No. of tests dropped	203			
Total time spent marking tests	60 h 30 min	29 h 30 min	42 h	101 h 20 min
Class average	63.4 ± 1.0	68.7 ± 1.4	72.3 ± 1.3	73.9 ± 0.7
Class average after tests dropped	79.0 ± 1.0			

*For the number of tests graded, 418 final exams were written in the final exam time slot; 22 deferred exams were written 1 mo later. †Final exam is double checked.

tallied at least two times to eliminate addition errors. Students could request to see their final exam; 3.2% of the final exams were regraded with a grade change of 2.5%. Students requesting that their final exam be regraded are usually students who were unhappy with their final course mark, as this is all that is posted. Many of the requests for a reassessment of the final exam were requested without students viewing their final exam first. Regrades are also requested as students are looking to find ways to bring up their overall semester average, and this course figures largely in that average. Therefore, requests for regrades of the final exam do not accurately reflect accuracy issues.

On average, 12.5 ± 3.1 min were spent marking each test, with ~4.1 min/question; 19 h and 5 min was spent remarking a total of 138 papers, which equates to 8.3 min/paper, and this time was mostly dedicated to a single question and writing feedback to the students.

TA time allocation. The above marking accuracy was achieved within the TAs' contracted hours for the course. The total time spent on all course tasks by the TAs did not exceed their 140 contracted hours allocated for the course (Table 2).

Time spent marking varied between TAs, mostly due to the random nature of how many test papers they had to mark. Students had a choice as to which questions to answer on tests, and it was not unusual that, given the choice between two questions, 75% of students answered one and 25% answered the other, meaning that one TA would have a larger number of tests to mark than another. Both experience and their comfort level with these types of tests also figured into the differences in time spent marking.

Time spent on the web bulletin board was less than expected given the frequency with which TAs were asked to monitor the board. An average of 8.9 ± 2.2% of their time was spent on addressing the bulletin board, and a semester total of 1,027 messages (questions and answers) was posted.

Professor time allocation. My time allocations are included in Table 1. I spent very little time with the TAs, especially in teaching them to mark the tests and final exam. This 12 h and 32 min was dramatically less than the time spent in previous years. The largest percentage of my time was spent on preparing and giving lectures, marking tests, and with student ap-

Table 2. Distribution of time spent on course tasks

	TA 1	TA 2	TA 3	TA 4	Professor
TA meetings	1 h 30 min	1 h 30 min	1 h 30 min	1 h 30 min	12 h 32 min*
Preparation time	19 h 30 min	2 h 35 min†	21 h 20 min	18 h 30 min	168 h‡
Class lecture time	12 h 30 min	16 h 20 min	14 h 40 min	17 h 40 min	57 h 40 min
E-mail	30 min	4 h 18 min	11 h 35 min	7 h 40 min	18 h 42 min
Web bulletin board	9 h	9 h 20 min	5 h 5 min	16 h 55 min	6 h§
Invidulating	4 h 40 min	4 h 40 min	4 h 40 min	4 h 40 min	6 h
Marking test 1		17 h	30 h		14 h 30 min
Marking test 2	13 h 30 min			16 h	
Marking test 3		31 h	11 h		
Marking the final exam	27 h	13 h 20 min	19 h	21 h	21 h
Marking deferred exams					4 h
Total marking time	40 h 30 min	61 h 20 min	60 h	37 h	39 h 30 min
Remarking					19 h 5 min
Question and answer session	10 h	13 h	13 h	14 h	12 h
Student appointments		1 h	1 h 30 min		36 h 35 min
Answering questions after class					9 h 17 min
Test design					34 h
Other (administrative duties)					13 h
Total hours	98 h 10 min	113 h 58 min	132 h 20 min	117 h 55 min	430 h 20 min

TA, teaching assistant. *Time for the TA meetings included time spent teaching the TAs how to mark exams; the TAs included this time under marking. †In terms of preparation time, TA 2 commented that coming to the lectures reduced preparation time dramatically; ‡the time spent by the professor was estimated at 3 h/lecture. §I made a conscious decision to not be involved in the bulletin board and only contributed when required; TAs indicated when this was necessary.

pointments (outside office hours). The very little time spent on the web bulletin board was intentional. I monitored it only very occasionally and only contributed when TAs alerted me as to confusion that needed my correction. I considered the web bulletin board to be student-TA territory and thought that my continued intervention would undermine the expertise and the credibility of the TAs. My other major support services were answering questions after class and office hours. The number of students attending office hours was dramatically reduced from previous years, and everyone that showed up had their questions answered. My question and answer discussion sessions were well attended, especially as they were the ones before the final exam. I still spent a considerable amount of time answering e-mails, but the time spent was dramatically reduced from previous years.

Student feedback. Statement scores from the normal course evaluation that were relevant to students' opinion regarding their assessment, their comfort level in asking questions, their impressions of my availability and approachability, and their overall satisfaction with the course learning experience are shown in Table 3. The scores for these statements from previous years are also shown in Table 3 for the sake of comparison. Students indicated strongly that they thought their assessment was fair. They also indicated strongly that they were free to express their opinions and ask questions, that they felt I was available and approachable, and that their overall satisfaction with the course learning experience was high.

DISCUSSION

We sought to maintain the integrity and the objectives of our Physiology course, which focus on integration, problem solving, and critical thinking, with a large class size by reorganizing the course TA structure. With this new TA configuration, we were able to show that 1) tests were marked accurately and in a timely fashion, 2) TAs did not overwork, and 3) students were satisfied with their assessment and the level of feedback and support given. Therefore, we were successful at maintaining the integrity of the Physiology course, including the use of long-answer-style testing, with a class size of 440 students.

The main philosophy behind the reorganization was to make the TAs experts in specific areas such that they could handle a wide variety of questions from students and answers on tests while affording me enough time to conduct all the lectures, so as to maintain a consistent philosophy of teaching throughout

the course. Not only did the reorganization allow us to maintain the integrity of the course, but the reorganization allowed for an expansion of TA duties and a reallocation of my time. In previous years, I had limited success with two TAs and a class size of 250 students and less success as the class size grew to 300 students where the TAs were used to help mark only. I spent a large amount of time teaching TAs basic material as well as the many potential issues and answers that may be given as students worked their way through the answers. It would have been more efficient to have the TAs attend lectures, but this activity would have taken more than half of their time allocated for the course due to the number of lecture hours. I spent a lot of time remarking tests, an indication of the decreased level of marking accuracy and consistency. TAs were not involved in any other support activities; therefore, I was the students primary source of support, and my time spent with TAs and remarking was time not spent with students. The reorganization of the TAs allowed me to dramatically reduce the number of hours spent teaching TAs how to mark tests (only 12 h and 32 min for all 4 TAs for the whole course), freeing me up to spend more time with students. Therefore, not only were we able to maintain the course, but we were able to focus my support efforts and introduce new support activities.

Our indexes of accuracy indicate that with 13% of the tests we achieved an accuracy of marking of 95%; therefore, in the other 87% of the tests, the accuracy was higher. A key element in being able to mark the variable nature of integrative long-answer questions was TA attendance of lectures in their selected areas. This ensured that TAs knew exactly what happened in class, planned or unplanned, and were able to see how information was integrated, so that they understood the source and context of many of the answers given on the tests that did not conform to the "standard" answer but still deserved marks for being right. This greatly enhanced their ability to mark accurately and consistently. TAs could mark for multiple and varied ideas rather than just for specific points on a marking scheme. The TAs needed very little guidance regarding course material, as reflected by the little amount of my time spent with them to achieve a high level of marking accuracy and consistency.

This accuracy relies heavily on student identification of a marking error, which, in our experience, is reliable. This reporting is made more reliable by, first, creating a "risk-free" environment. This risk-free environment was one in which regrading was done, marks were not taken away, students were

Table 3. Student feedback regarding issues relevant to fairness of assessment, student comfort levels in asking questions and seeking feedback, and whether they thought that enough support was offered in to supply them with a favorable experience

Relevant Statements Scored at the Course Evaluation	Statement Scores in Previous Years and the Current Year Offered (out of 5)			
	2002	2003	2004	2005
1. Students were free to ask questions, disagree, and express their own opinions.	4.91	4.94	4.76	4.95
2. The instructor was available to students outside of class time for discussion.	4.82	4.73	4.66	4.76
3. The assessment of students was fair.	4.47	4.14	3.78	4.50
4. Considering everything, this course provided a worthwhile learning experience.	4.78	4.74	4.61	4.84
Average score for all 12 questions	4.76	4.67	4.53	4.80
Class enrollment, no. of students	254	250	308	440
No. of respondents to the course evaluation	211	188	218	382

The evaluation asked students to score statements from 1 to 5, where 1 = strongly disagree and 5 = strongly agree.

made aware of this, and the whole test was not regraded, only the questions singled out by the student. Second, we allowed for anonymity in the remarking process. Regrading was done in the absence of the student in part to decrease the pressure on the students who did not want to be identified by coming to my office and in part so that the remarking would be consistent. If present when remarking, students often want to supplement their answers with an oral component (e.g., “what I meant was . . .”), which may unfairly bias my remarking. Finally, we created a very comfortable and approachable environment for students to solicit feedback. All of this was done to encourage students to seek feedback and discussion of the questions and answers. This system has the potential to be abused by students that resubmit all tests wanting every question remarked hoping that we will find something, but these cases were few and part of the feedback to these students was that this behavior was not appropriate.

Part of the success in our ability to mark accurately and in a timely fashion was attributed to the optional testing arrangement. The optional testing arrangement served many purposes. Previous experience indicated that when optional tests are offered, not all students will write them, thus decreasing the marking load. This decreased marking load allows for more accurate marking of the submitted tests and more time for TAs to be allocated for other support duties. However, at no time were students discouraged from writing tests; in fact, the opposite was true—it was consistently reinforced by myself and the TAs that tests should be written. To encourage students to write tests, we created an environment that taught them how to write the tests while, at the same time, not detracting from their mark during this learning process. For this process to work, students must 1) see the value in writing the tests and 2) be encouraged to seek feedback when necessary. The latter is supported by the risk-free regrading system described above. The value the students see in writing the tests is that the final exam is in a similar format to the tests, and, therefore, they have the opportunity to prepare for it in a low-risk testing environment. This low-risk environment is promoted by allowing students to write the first test and drop the mark if they are not satisfied. Eighty-five percent of the class took advantage of this opportunity, not only to write the tests but also hand them in for feedback; in some cases, only one of the three test questions was written, and, obviously, these students were hunting for feedback only. The second and third test marks could not be dropped, but only papers handed in at the end of the test period were marked, so students could, again, come in and try the tests without the fear of being penalized if they were not yet comfortable with the testing style. The primary reason for not allowing marks for *tests 2 and 3* to be dropped was that some students need consequences to seek out feedback. For example, it is very common for students to misread questions because they are reading too fast; therefore, they can get 0/10 on a question they thought they answered correctly. If they can drop it easily, they will often do so, with no attempt to find out why this happened and no attempt to change their strategy for future tests; this destines them to repeat their mistakes. However, if they cannot easily dismiss the problem (drop the mark), then the consequence of having to keep a lower mark usually drives them to seek feedback. For this reason, tests are weighted as a relatively low percentage of their final mark so that this consequence has little impact on their overall mark.

Fewer students wrote the second and third test for marks (39% and 33%, respectively). This does not include the number of students that wrote the tests and did not hand them in to be marked; this number was estimated to be at least 50 students/test. Anecdotally, students say that part of the reason they choose not to write tests was their tests/assignments in other courses. Therefore, they are incorporating the optional testing scheme into their time management and possibly doing better in other courses because of it.

Ultimately, the optional testing scheme, built to create a low-risk environment, decreased the marking load, increased accuracy of the tests marked, and allowed us to expand TA resources into other support services and focus feedback on students who wanted/needed it. This does require a mature student population, but this is a third-year university course and maturity is expected.

One of the biggest challenges was to maintain or improve our accuracy for the final exam. The heaviest pressure came from administrative deadlines regarding the submission of final marks. Most of this pressure stemmed from the placement of the final exam on the last day of the exam schedule. The guidelines indicating that final marks must be submitted 4–5 days after a final exam is written is a flexible one except when the deadline is the very last day of the semester. As our final exam was on the last day of the exam schedule, the end-of-semester deadline could not be extended. To encourage the offering of this type of course and promote the accuracy of marking, the administration could help by taking into account the testing style when scheduling final exams. For example, those using multiple-choice, computer-corrected testing styles could be placed at the end of the exam schedule and easily meet their grade submission deadlines, whereas those marking long-answer exams are pressed for time and accuracy may decrease because of this. While I realize that exam scheduling is a complex procedure, the testing/exam format should be a consideration in this process.

Evaluation of whether the feedback provided was adequate was the hardest variable to quantify. We do know that students were satisfied with their learning experience and that they considered that I was both approachable and available for feedback and that, ultimately, they thought that they were assessed fairly. All of these indicate that they had adequate support and feedback to not only manage the testing style but be successful with it. Support and feedback also extended beyond testing interactions to include e-mail support, bulletin board support, professor accessibility, and question and answer discussion sessions. Note the fall in score for the question regarding fairness of assessment since 2003, when the class size began to increase but the TA resources (2 TAs that marked exclusively) and structure had not changed. This was interpreted as an indicator that the course enrollment was getting too large with the provided resources and was not meeting students' needs adequately; this was one of the primary motivators to rethink the course structure when the enrollment jumped to 440 students.

Past experience has shown that time in office hours and on e-mail are spent answering the same question again and again. We thought the web-based bulletin board might help eliminate this. Of all the new initiatives undertaken to provide adequate student support, the initiative that ran the highest risk of failure was the use of the web-based bulletin board to replace TA

office hours. The increased risk of failure was based on this medium's ability to spread and perpetuate misinformation. Because of this risk, TAs were asked to monitor the bulletin board daily during the regular lecture series and hourly 3 days before each test. This task only occupied an average of $8.9 \pm 2.2\%$ of the TAs' total time. Anecdotal student feedback indicated that they found the bulletin board very helpful. The success was, in part, due to the diligence of the TAs at checking the board; therefore, it became known as a reliable place to get feedback fast. It was, however, very easy to get wrong information introduced; therefore, this type of TA diligence was required to control the board discussions and correct inaccurate information as quickly as possible. Although there are no direct data to support this claim, the success of this service must have contributed significantly to decreasing the number of hours spent on e-mail by myself and the TAs as well as the number of people during my office hours. Given its success, it is an initiative I would continue.

In conclusion, we were successful at maintaining the integrity and learning experience of our Physiology course in the face of increasing enrollment by reallocation of TA duties. This reallocation essentially created a team of TAs that were

experts in their field, which allowed them to mark long-answer-style tests accurately and competently as well as expand their services to support the students. This reorganization provided me the time to maintain the objectives of the teaching of the course as well as refocus my time to student support activities. Therefore, with the appropriate resources and organization, it is possible to maintain the philosophy, academic rigor, learning experience, and use of long-answer-style tests in courses with large class sizes.

ACKNOWLEDGMENTS

First, I acknowledge the founder of these physiology courses, Dr. Jack Barclay, the ultimate educator. I also acknowledge the efforts of the outstanding TAs for this course, Mark Dekker, Katie Junkin, Jasna Junuzovic, and Sean Leonard. None of the above could have been achieved without the investment of support at the departmental and college levels as well as incredible support staff, all of which I am fortunate to have.

REFERENCES

1. **Bizzell P, Singleton R Jr.** What can we do about essay exams? *Teach Sociol* 16: 177–179, 1988.
2. **Jacobs LC, Chase CI.** *Developing and Using Tests Effectively: a Guide for Faculty.* San Francisco, CA: Jossey-Bass, 1992.

