*Science and Engineering Ethics (1996) **2**, 345-366*

Evaluating Teaching and Students' Learning of Academic Research Ethics

Deni Elliott, *Mansfield Professor of Ethics in Public Affairs, University of Montana*
Judy E. Stern, *Associate Professor, Dartmouth-Hitchcock Medical Center, USA*

Keywords: evaluation, ethics, teaching, course development, pre-test, post-test, learning environment, assessment

ABSTRACT: *A team of philosophers and scientists at Dartmouth College worked for three years to create, train faculty and pilot test an adequate and exportable class in research methods for graduate students of science and engineering. Developing and testing methods for evaluating students' progress in learning research ethics were part of the project goals. Failure of methods tried in the first year led to the refinement of methods for the second year. These were used successfully in the pilot course and in one university setting external to Dartmouth. The process of development and justification for the final methods are discussed here.*

Successful evaluation of applied ethics teaching depends on clear and direct connection between measurable instructional goals and the assessment techniques that give students an opportunity to show that they have achieved those goals. That seemingly obvious point gets lost in the teaching of ethics, a discipline that some doubt can be taught at all. Furthermore, some scholars claim that successful ethics teaching cannot be measured even if it can be taught. Any true effect of a course, they would argue, cannot be measured until long after the intervention. As a team of scientists and philosophers[1,2] from Dartmouth College discovered in the pilot teaching and evaluating of a graduate level seminar in academic research ethics in the 1993-94 and 1994-95 academic years, the task of assessment is not easy, even if instructional goals are clear. However, evaluation of one's success in teaching ethics and the evaluation of students' learning in the class is both possible and necessary.

While it is unusual in most scholarly disciplines to discuss failed experiments in the literature, we do so here because many other institutions are also in the early stages of developing graduate level courses in academic research ethics. Alternatively, they

Addresses for correspondence: Deni Elliott, Mansfield Professor of Ethics in Public Affairs, University of Montana, Missoula, MT 59847, USA.
Judy E. Stern, Associate Professor, Department of Obstetrics and Gynecology, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, USA.
Paper received 21 September 1995; revised 11 June 1996; accepted 12 June 1996.
1353-3452/96

are experimenting with other ways to meet the mandate of Federal funding agencies (e.g., the National Institutes of Health)[3] to include ethics instruction for students on training grants. Because we have achieved important understandings through our evaluation attempts and mistakes, we present the failures here along with our ultimate success so that others can learn from them as well. Specifically, we discuss our theoretical assumptions about teaching and learning ethics, which we believe to be correct, and will describe how our first attempt to put these assumptions into practice failed. We also describe how we modified our instruction style and instruments in the second year and achieved substantially better results.

If Ethics Can Be Taught, Then the Results Can Be Measured

We start with the assumption that applied ethics is an academic discipline. As such, it has a body of knowledge as well as legitimate processes or skills that can be used, taught and learned. A list of criteria to guide evaluation and assessment that was introduced by the Hastings Center in 1980[4] still provides a standard set of criteria from which one can choose for judging student success in the ethics course:

1. *Quality of arguments for moral issues*. Students can be judged on the basis of how well they argue a case rather than on the position they take.
2. *Mastery of theories and principles of ethics*. Students can demonstrate knowledge of ethical traditions and concepts.
3. *Identification of moral issues*. Students can show their ability to isolate moral issues from questions of law, economics, religion, politics, prudence and aesthetics.
4. *Ability to argue both sides of a position*. Students can show their understanding of various arguments and can evaluate the soundness of the arguments.

Because the student's success in accomplishing any or all of these can be measured, the results of the ethics course can be determined. However, this is a field in which some teachers and scholars are not at all sure that the positive aspects of ethics teaching are quantifiable. Our first step in developing evaluative techniques was to deal with that skepticism.

Pedagogical Hope vs. Instructional Objective

The conceptual confusion that leads to evaluative skepticism is the confusion between pedagogical hope and instructional objectives. One might have pedagogical hope that one's students become highly ethical practitioners in their careers and become highly ethical people in their private and public lives as well. But, that is not an objective that can guide the teaching plan for a specific class period. The difference between pedagogical hope and an instructional objective is understood in other areas of the academy. For example, a literature professor might hope that his students become producers of fine literature and a science professor might hope that her students advance scientific knowledge through their future work. Both professors know that none of these hopes can be fulfilled by the end of a semester and none lend themselves to evaluative measure except in the most longitudinal and non-correlative sense.

Regardless of the pedagogical hope for one's students, it is the immediate instructional objectives for the course that provide criteria for determining whether the instruction has been successful or not.

The confusion of pedagogical hope with instructional objective in the teaching of applied ethics is demonstrated by some of the important, early champions of the movement.

For example, Derek Bok, past president of Harvard University, says that the "positive results" of teaching applied ethics "may be impossible to document". But, he adds that "the consequences of doing nothing are plainly intolerable."(p. 126)[5]

Bok believes that something of importance is happening in the ethics classroom. He says, "It does seem plausible to suppose that such classes will help students become more alert in perceiving ethical issues, more aware of the reasons underlying moral principles, and more equipped to reason carefully in applying these principles to concrete cases (p. 135)."

However, because he believes that the ultimate goal of teaching applied ethics is to help students behave more ethically, Bok says that the results of ethics teaching are impossible to measure. Bok asks if students can be expected to behave more ethically because of a course in applied ethics. He answers himself this way:

> One would suppose so, provided we assume that most students have a desire to lead ethical lives and share the basic moral values common to our society. Granted, we cannot prove that such results will be achieved. Even so, the prospects are surely great enough to warrant a determined effort, not only because the subject matter is interesting and the problems intellectually challenging but also because the goal is so important to the quality of the society in which we live (p. 135).

Through his focus on his pedagogical hope, Bok fails to consider that it may be his instructional objectives (alertness in perceiving ethical issues, awareness of the reasons underlying moral principles, ability to reason carefully in applying principles to concrete cases), that provide the basis for analyzing the 'positive results' of the course.

The confusion between pedagogical hope and instructional objective is also evident in the recent Federal requirements that training grants would be provided only if the grant includes a program for teaching "responsible conduct" and learning "principles of scientific integrity".[6] When the National Institutes of Health mandated training in ethics, the expectation was that funded organizations do what they could to produce ethical scientists: — that is, scientists who were committed to performing their work in a responsible manner. But, of course, no training grant can produce 'ethical scientists' any more than it can produce 'successful scientists'. What ethics training can do is to identify a narrow scope of skills and knowledge that students should learn. Through evaluation, ethics professors can then decide if these skills and knowledge have been learned.

The Problem with Moral Development Measures

Ethics instruction should not be concerned with the moral behavior of students, even if good moral behavior is a pedagogical hope for people who take the class. Yet, it is not surprising that the Federal guidelines would include this misunderstanding about the

nature of ethics instruction: this is a misunderstanding that permeates teaching of applied and professional ethics in a broad sense. For example, a well known and often used instrument for assessing the success of ethics instruction, the Defining Issues Test (DIT) developed by James Rest and colleagues at the University of Minnesota, seeks to answer the following question: "Can moral judgment development be affected or improved by deliberate intervention?" (p. 204)

The test is designed to measure how students make moral judgments in hypothetical situations. Based on the moral development stages developed by Harvard psychologist Lawrence Kohlberg, the test measures change by administration of a pre- and post-test. The DIT has measured moral growth in students that the testers attribute to ethics instruction (Rest, p. 207).[7]

Using the DIT to evaluate ethics teaching and learning seems to us to miss the mark. First, it does not take into account the other variables in a young adult's life which may contribute to accelerated growth. The moral tests that confront students outside of the ethics classroom confound the results of such testing. It seems arrogant, at best, to believe that teaching a system of decision making over an 8-14 week period alone creates new moral capacities in the student that would have any lasting effect. Furthermore, given that Kohlberg postulated that it takes an individual an average of 12 years to mature through one moral developmental stage, it is not even clear what such accelerated growth might mean.

Moral behavior does not provide an adequate basis for evaluating the teaching of ethics. Caplan notes that focusing on behavioral change "skews the assessment in the direction of the effects, rather than their causes."(p. 138)[8]

Students learn the conventions and expectations of scientific conduct by functioning in the laboratory environment and by daily observations of *how* the more experienced others in the lab act qua scientist. They learn how to act, but not necessarily *why* they should act or not act in a particular way. The purpose of ethics education is not indoctrination or inculcation of values, but analysis of those professional values.[9] Appeal to rationality is the hallmark of ethics instruction as it is the hallmark of higher education as a whole.

As Caplan says,

> An emphasis on the overt, observable effects of teaching on students would, for many teachers, inappropriately undervalue the vitally important role of reasoning in coming to act upon or maintain a set of value beliefs. 'Right' reasons, rather than 'right' answers or 'right' conduct, is the measure likely to concern instructors who have only limited time and access to students (p. 144).

Defensible Measures for the Evaluation of Ethics Teaching and Learning

If 'ethical behavior' is removed as a basis for the evaluation of the teaching of ethics, it seems to us that the teacher is left with two bases from which to measure the success of the course: the skills and content taught in the course and the learning environment in which the teaching takes place.

In addition to the standard student evaluation measure, we initially chose evaluation strategies that included a pre- and post-test essay to measure the students' success in incorporating content and skills and a pair of instruments developed by the

Center for the Study of Intellectual Development to provide data concerning students' learning styles and students' perceptions of how well the learning environment met their needs (Appendices E & F). The first measure was substantially modified between our two years' experience. The second measure was finally determined to be inadequate to truly measure what we had in mind and we found that the standard student evaluation form served somewhat the same purpose in a less complex way.

Standard Measure - Student Evaluations

The standard measure for determining the success of college courses is a student evaluation. We had our students complete a standard evaluation form at the same time as we administered our other evaluation instruments. In this way we could compare student perceptions of the success of the course with success of the course as defined by faculty goals and standards.

Students rated the course on a scale of 1 (strongly disagree) to 5 (strongly agree) on a variety of measures. In the first year, students rated the instructors between 3.4 and 4.3 in the areas of instructor preparedness, effectiveness of leading discussions and enthusiasm. However, students rated the difficulty of the course as only 1.5. All of the students believed that the course content had been easy to understand.

In the second year, the student ratings for course organization and instructors were between 4.3 and 5. These students rated the difficulty of material at only 2 on the 5 point scale.

The consensus among students that the course was easy provided a warning flag for faculty as we attempted to interpret other measures. We attribute part of the failure of our first year's pre-test/post-test essay measure to the students' erroneous belief that there was really nothing new to learn in this course, even if they enjoyed the exercise.

Direct Measure - Content and Skills

Content and skills provide an obvious starting point for student evaluation, but, as the Dartmouth team learned, they are measurable only if the content and skills are specified in instructional objectives and only if these objectives then provide the skeleton for the course itself.

A pilot course was taught to graduate students of science and engineering in the Winter 1993 term at Dartmouth College. This 8 week 2 hour/per week course was taught by a team of five scientists after two years of interacting with philosophers in faculty seminars on issues in ethics and academic research. The instructional objectives seemed to us, prior to the course, quite clear. They included:

1) be able to clearly describe relevant scientific conventions including: laboratory practice, institutional responsibility, etc.;
2) be able to describe what leads to ethical problems including causes inherent in the social context of the practice of science;
3) be able to identify ideal scientific practice and consider how to bring scientific conventions more in line with the ideal;
4) be able to separate behaviors into four categories: morally prohibited, required, permitted, and encouraged, thus illustrating an understanding of the role of the scientist in society.

To measure the students' accomplishment of these objectives in the first year, we

constructed a pre-test/post-test essay that consisted of an edited version of a case fraught with ethical implications that was published in *Science* (see Appendix A). Our instructions to the students were as follows: "Identify the ethics problems in this case. Discuss what the individuals involved did right. Discuss what the individuals involved could have or should have done differently."

The pre- and post-tests were scored by three ethics professors external to the institution; the results from these scorers were then compiled and analyzed by an evaluator, another ethics professor who was also external to Dartmouth. All four readers had ten or more years of experience teaching and researching in applied ethics, but none specialized in scientific research ethics. We had purposefully avoided teachers of scientific ethics as a way of avoiding scoring bias on how the case 'should' be decided. However, we decided in retrospect that familiarity with issues in scientific research and with science students might have helped to increase inter-scorer reliability.

The raw pre-tests were sent separately from the raw post-tests for scoring, which itself may have biased the reviews. Comparisons among readers showed that overall student performance rated between a 20% gain to a 4% drop in achievement between the pre- and post-test. There was no inter-rater reliability, even with what we initially perceived to be the easiest question to score (Count the number of moral agents that the student identified in the scenario.) It wasn't until we analyzed the raw data of the students' responses ourselves that we understood how the readers could disagree.

When the students' pre-and post- tests were coded and combined by a research assistant so that the faculty could not identify the test-takers or tell which test was a pre-test and which was a post-test, it became clear to the faculty that there was no significant difference between how students approached the vignette at the beginning of the term and how they approached it at the end. Most papers were poorly written and showed little attention to detail or apparent concern for content. There was no evidence of the more sophisticated and systematic approach that we expected students to take in the post-test. We realized that even if students had learned the material they had not been encouraged to express what they had learned. The post-tests were not graded; indeed the students' grade of 'Pass' was determined solely on attendance.

Another consistent problem with the students' essays was that the students understood the test directions differently from the way in which the faculty intended them. In designing the instructions the intent had been for students to discuss each ethical problem and to elaborate with regard to that problem what individuals involved did that was right, and what they could have done better. Indeed, the scoring sheet developed by the faculty (Appendix B) reflected this intent and sought to determine the development of a student's understanding of the complexities of each issue.

However, the students made lists instead of elaborating on each issue and writing thoughtful analyses of the roles of individual moral agents within the context of that issue. Students listed what they saw as ethical problems; they listed a few individuals with clauses noting what they did that was right; and they listed one or two individuals with a clause describing what they could have done better. Few students made any attempt to integrate the three questions and to evaluate the responsibilities of individual moral agents as complex people. As a result many answers reflected a 'good guy', 'bad guy' approach.

Most interesting in our findings was the realization that what the faculty identified as the primary ethical questions for the content analysis form used by scorers (see Appendix B) were not what students perceived to be primary ethical questions. Scoring of student responses on the content analysis forms was impossible because the students were literally off the chart. In a very straightforward way, the students let us know that they were considering issues other than the issues identified by faculty as 'the' ethical problems. The case given to the students involved post-doctoral fellows who reportedly fabricated and falsified data for publication. While faculty members considered the post-docs to be autonomous agents despite poor mentoring; many students perceived them only as victims.

For example, one student wrote, "I feel for the student because his advisor or mentor should have been there to help in constructing and checking the figures."

Another wrote, "In both cases, some fault should be found with the researcher in charge of the lab, Hood. Was he putting undo pressure on these students to publish at the cost of the integrity of the science being done?"

Yet another wrote, "Essentially, Kumar's problems stem from his inadequate training. He evidently didn't know that duplicating lanes as he did was improper. The fact that his professors pushed him along from stage to stage without requiring him to complete his work, and, indeed, performing Kumar's work for him seems unethical."

We realized at the end of the seminar that we had failed to fully incorporate our objectives into the teaching methods employed. We did not provide the students with any explicit instruction in conducting systematic moral analysis. We instead expected students to intuit the process by examining a series of cases. The application of method by instructors was often too subtle for students to discern that any system or generalizable procedure was being used. Despite consensus among members of the faculty team that supervised practice was necessary for students to learn how to think systematically about moral problems, many of us lapsed into a traditional lecture mode when teaching the class, often leaving less than one-quarter of the class time free for student practice.

The realization that the students framed the ethical issues differently from the faculty is significant in terms of the teaching of the class. It is imperative that the teacher not assume that *the ethical issue* is obvious. Indeed, how a student frames the ethical issue can tell us something about their own concerns.

Thus, our direct measurement tool failed in three ways:

(1) the scoring was biased by our distribution to the readers of pre- and post-tests as separate groups. The pre- and post-tests should have been distributed at one time without indication of which test was taken first;
(2) we failed to provide a motivational tool to encourage the students to take the pre- and post test seriously, and
(3) we failed to understand that the students might perceive the case for analysis differently from the way in which the faculty perceived it.

We made the following adjustments for the Year Two Direct Measure:

(1) We provided short vignettes (Appendix C).[10] Two contained issues of ethical importance (from the faculty's perspective) and one did not.

(2) More explicit instructions were given for the pre-test and students were told that a 'high quality' response to the pre-test/post-test was necessary to receive a 'Pass' in the class.
(3) At the time of the post-test, students received back their pre-tests, with instructions to analyze how well they had responded to the pre-test (Appendix D).

Students were, thereby, asked to perform a meta-analysis and the results were far more satisfying from an evaluation standpoint. As we did not involve external evaluators, the report of the results are anecdotal but compelling.

For example, in Scenario #1,* Student A reflected that on the pre-test, he "seemed to think that the PI was the supreme sovereign and the graduate student was subject to his/her whims rather than being an active participant in decision making." Student A noted that in addition to consulting the first author (which he had included in the pre-test), the PI should also consult the graduate student. Giving the student authorship, Student A reasoned, would make the student accountable for the data contained in the article.

In discussing the same Scenario, Student B said in the pre-test that the student should "extract herself from this mess". In the post-test, Student B provided a generalizable method for addressing the problem: "Would you want everyone else to do this? Clearly authorship becomes meaningless in more ways than just as an evaluatory measure if anyone's name could appear on the paper out of courtesy. It is also deceptive to claim authorship that one is not entitled to."

In discussing Scenario #2,* Student C discusses only Michelle in the pre-test. "The ethical thing to do would be for Michelle to publish the original micrographs from the very beginning. In the post-test, Student C says "all in all, there are more options available and more ramifications involved in this scenario than I had originally considered." She also identifies John as having moral responsibilities saying that if Michelle is unwilling to retract the photos, John should "report it to someone in a higher position."

Students generally realized that there were no specific ethical problems in Scenario #3,* however, Student D included in the post-test a note that if students come to the lab agreeing to those conditions, then they are bound to follow them. Another student, after identifying no ethical problem in the pre-test analysis, suggests that the Lab Director is acting immorally by stifling students' development and by presenting the work of students as his own.

This pre-test analysis/post-test meta-analysis has also been used successfully in five courses at the University of Montana during the Spring, 1995 term: Introduction to Ethics, Ethics in Public Affairs, Ethics in Journalism, Ethics in Social Work and Ethics in Anthropology. Three of these courses were taught by a co-author (DE); two were not.

Indirect Measure - Analysis of the Learning Environment

Learning applied ethics requires students to take intellectual risks; it requires students to give close examination to their beliefs, values and methods of thinking about adequate and inadequate professional behavior. Clearly, some environments encourage

* Scenarios 1,2 and 3 are reproduced in Appendix C, pp. 360-1 of this issue.

this kind of risk taking and other environments discourage it. Our attempt to analyze the learning environment reflected our desire to create the kind of environment that encouraged, rather than discouraged ethics learning.

Of the many variables that can impact on the student's ability to learn, the classroom environment is one of the few that is within the teacher's control. If an instructional objective is to help students become more autonomous decision makers about the moral aspects of scientific practice, then it is important to create the kind of classroom environment in which such growth is likely to occur. Constructing this kind of classroom environment at the post-secondary level, generally, includes:

1) faculty who express doubt rather than authoritarian 'right' answers;
2) faculty who are open to the ideas of others and who have a rational defense for their own position;
3) faculty who expect students to defend their ideas through critical analysis;
4) faculty who are available for interactions with students and settings that encourage discussion about ideas.[11]

Measure most of the student's perceptions of the environment paired with an analysis of the student's learning style has been used to help assess the success of other innovative post-secondary curricula. For example, the University of Wisconsin Medical School developed an instrument to monitor a non-traditional medical curriculum. Harvard Medical School further refined this tool and added it to others to assess the impact of their own non-traditional medical curriculum, the New Pathway program.[12]

For the Dartmouth pilot course in Year One, we used instruments developed by William Moore at the Center for the Study of Intellectual Development, which follow the work of Knefelcamp and Perry.[13] A nationally-normed measure, Learning Environment Preference Index, (see Appendix E) was used as a pre-test; a new measure, tailored for the evaluation of the ethics classroom, Perceptions of Learning Environment, (see Appendix F) was administered as a post-test.

Learning Environment Preferences

The *Learning Environment Preferences* (LEP) percentile indicates the pattern of complexity that the learner prefers in a classroom environment. Generally speaking, the mature learners that one would expect in a graduate level course will prefer less structure and more abstraction.[14] The learner's development is rated on the scale from Position 2 (prefers great structure and little abstraction) to Position 5 (prefers great abstraction and little structure). An examination of the individual positions can provide information on the style of instruction that is likely to be most comfortable (and the styles that are likely to be the most challenging) for the student or group of students.

For example, a high degree of structure is supportive for students who are in Position 2, but is difficult for those who are in Positions 4 and 5. A high degree of structure gets in the way of these more mature learners.

A diverse set of perspectives is difficult for those in Position 2. Those in Position 3 appreciate diversity. They would prefer to hear more rather than fewer perspectives

(because that gives them more practice in determining the right answer); those in Position 4 appreciate receiving an increasingly complex set of alternatives.

Abstraction is a challenge for those in Position 2, who are waiting to hear the RIGHT answer, but abstraction becomes a support for those in Position 5, who are continually seeking more adequate ways of judging and interpreting. Experiential learning is a support for those in Position 2, but can be a challenge for those in Position 5. The more mature learners would prefer to think about what makes something right or wrong rather than be put into a learning situation that requires them to actually make the right/wrong determination.

Personalism, the creation of a 'safe' environment for the taking of intellectual risk, is important for learners whatever their developmental stage, but the students' perception of who is responsible for their safety changes dramatically. Students in Position 2 require strong support from their teachers; learners are increasingly less teacher-dependent and more dependent upon themselves and their peers as they mature. By the time a learner is in Position 5, the role of the teacher is best one of facilitation rather than as a direct source of student support or learning.

We used the information from the LEP to compare with the level of student satisfaction described in the indirect measures post-test, *Perceptions of Learning Environments* (PLE). However, we now believe that the information from the LEP would be better used early in the term so that faculty can reconsider the appropriateness of planned course content and activities. The challenge for faculty, as most teachers intuitively suspect, is to make students feel uncomfortable enough with their current understandings that the students will feel compelled to take on new knowledge and understandings, and at the same time, create an environment that is safe enough for students to take the necessary intellectual risk.

The LEP and PLE indicated that our group of students had evolved past the need for a highly structured classroom experience and would be responsive to a discussion-style approach with focus on generalizable principles. Individual student profiles, however, showed ambiguity about their teaching style needs. One would expect individual student responses to create a bell curve around a single maturity level. A few students showed a distribution of this type but others displayed a much wider distribution of responses. Two of the students had a bimodal distribution of responses, as did other students we tested in other settings. Their responses indicated both a desire for extensive structure and a desire for considerable independence.

While Moore declined to draw conclusions relating to the bimodel distribution, aside from noting that it shouldn't happen,[15] we suggest that it might be particularly difficult to interpret the developmental learning stage of graduate students of science. The picture of our students as mature learners, ready for independence and facilitation rather than direct instruction, fits with what one would expect of students who are more than 22 years of age. However, the expectations one has of the graduate student qua student of science is quite different from the expectations one might have of a person that age outside of the lab. The graduate student is often the lowest ranking member of the lab. As such, she may be expected to be a passive learner and to acquiesce to authority. Often this means puttting aside independent judgment and research desires. While the graduate students enrolled in a research ethics course are likely to be developmentally capable of independence, they may be operating in a laboratory setting that can function to discourage such maturity. The rigid hierarchical

structure of many labs may create some dissonance for the students when they are asked to operate independently in the ethics classroom.

The Perceptions of Learning Environments Data

The *Perceptions of Learning Environments* (PLE) was administered as a post-test to the student's enrolled in Dartmouth's pilot seminar. The instrument asked students to reflect on four characteristics of the learning environment created by the course in terms of its level of diversity, structure, abstraction, and personalism. **Diversity** refers to the number and complexity of alternatives or perspectives presented. It also includes the process of sequencing and contextualizing diverse aspects of the course. **Structure** refers to the amount of framework and direction provided to the students. Examples of structure include providing a clear context for the course, defining basic terms, providing specific guidelines and using specific examples from student's experiences. **Abstraction** refers to the continuum of direct and concrete 'hands on' experiences to abstract, vicarious ones. Students who are more mature do better with abstract learning than do students who are less mature. **Personalism** refers to the level of safety in the learning environment, i.e., the encouragement of risk-taking and of students' listening to one another.[16]

Unfortunately, ambiguity of the wording of questions in the PLE post-test and the lack of norms made results from the post-test unusable. Discussions with the Center Director, Bill Moore, raised doubts about the ability of this test to accurately reflect students' perceptions of the learning environment or to correlate with the LEP.[17] While we believe that an adequate classroom environment is a necessary, but not sufficient, condition for teaching ethics, we concluded that the standard student evaluation form can provide some of this essential information without using an hour of class time for administration of the LEP/PLE pre and post-tests. However, we also believe that it is important for ethics teachers to be particularly sensitive to learning environment issues. In further research, we will be testing other less cumbersome instruments in an effort to better measure the appropriateness of the learning environment for ethics teaching.

Conclusions

Our conclusion is that the success of the teaching of academic research ethics can be easily evaluated, but only if it is treated as an academic discipline by both faculty and students. Changing the pre-test/post-test essay to a self-assessment meta-analysis helped us better unite the theory and practice of evaluation. Even if a traditional pre-test, post-retest 'worked' in terms of showing us that the students had achieved more sophisticated approaches to considering ethics problems, it would not have provided all of the information we needed to know. Greater consciousness is required for more consistent ethical reasoning. Only a student self-assessment can reveal this.

We were also surprised to realize that while our evaluation tools failed during our first attempt, our evaluation process was still a success. That is, if the purpose of honest evaluation of teaching (or of research, for that matter) is to further one's learning, we achieved that goal.

We learned how our tools needed to be refined, but far more importantly, we learned more about the students who we teach. Successful ethics teaching begins with teachers and students sharing common ground in determining the difference between people who are "victims of the system" and those who are moral agents who are responsible and accountable for their actions. They also need to share a belief that the systematic analysis of ethical problems in the lab is a complex but important task to undertake.

REFERENCES AND NOTES

1 This work was made possible by a grant from the Fund for the Improvement of Post-Secondary Education, Department of Education, #P116B20871.

2 The faculty research, development and teaching team includes Edward Berger, Department of Biology, Marilyn Brown, Animal Care and Use Program, Deni Elliott, then Institute for the Study of Applied and Professional Ethics, Bernard Gert, Department of Philosophy, Allan Munck, Department of Physiology, Judy Stern, Department of Obstetrics and Gynecology at Dartmouth and Karen Lomax, Clinical Ethics Center, Veterans Administration.

3 *Human Research Report*, Vol. 5, No. 6, June 1990, p. 1

4 Caplan, A. (1980). "Evaluation and the Teaching of Ethics," in: Callahan, D. and Bok, S. (eds) *Ethics Teaching in Higher Education*, Plenum Press, New York.

5 Bok, D. (1982). *Beyond the Ivory Tower, Social Responsibilities of the Modern University*. Harvard University Press, Cambridge, MA, USA.

6 *Human Research Report*, Vol. 5, No. 6, June 1990, p. 1.

7 Rest, J. (1979) *Development in Judging Moral Issues*. University of Minnesota Press, Minneapolis, MN, USA.

8 Caplan, A. (1980) "Evaluation and the Teaching of Ethics," in: Callahan, D. and Bok, S. (eds) *Ethics Teaching in Higher Education*, Plenum Press, New York.

9 For more detailed description about the difference between ethics education and inculcation of professional value, see Elliott, D. (1984) "Education: Avoiding Indoctrination Through Ethics Instruction", *1984-85 Journalism Ethics Report*. National Ethics Committee, Society of Professional Journalists, Sigma Delta Chi. p. 31.

10 Upon further reflection, we think that short hypotheticals do not provide the richness and complexity of real cases. Future plans include comparison of student responses to thse two types of cases.

11 Kurfiss, J. (1986) *Intellectual, Psychological and Moral Development in College: Four Major Theories*. Unpublished paper.

12 Moore, G.T., Block, S.D., Style, C.B. and Mitchell, R. (1994) "The influence of the New Pathway Curriculum on Harvard Medical Students" *Academic Medicine* **69 (12)**: 983-989.

13 Moore, Wm. S. (1991) *The Perry Scheme of Intellectual and Ethical Development: An Introduction to the Model and Major Assessment Approaches*. Paper presented at the American Educational Research Association annual meeting, Chicago, April 3-7.

14 Ibid

15 Personal communication betwen co-authors and William S. Moore.

16 Knefelkamp, L., Moore, W., Taylor, K. (1991). *Dimensions of Effective Teaching: Four Classroom Design Elements.* Center for Intellectual Development, Seattle, WA, USA.

17 Personal communication between co-authors and William S. Moore.

APPENDIX A
Misconduct: Caltech's Trial by Fire

Edited extract from *Science* **253**: 1344-1347*

Vipin Kumar and James Urban joined the Hood lab several years ago; Kumar from a postdoc at Harvard, Urban from the University of Chicago. They began working, at first together and then independently, in an especially hot area of immunology research, looking at the molecular biology of and possible treatments for autoimmune diseases such as multiple sclerosis. Pressure was intense, as it is throughout the huge Hood lab, which numbers 65—especially because Hugh McDevitt's gorup at Stanford was pursuing the same tack. Both Kumar and Urban were ambitious, logging long hours and winning high marks from Hood in the process.

But not everyone shared Hood's opinion—several people in the group went to him with their suspicions, not about fraud, *per se*, but about sloppy science, says Hood. He investigated each accusation and turned up nothing solid, chalking the problems up to personality conflicts and inexpereicne. "I had complete faith," he recalls.

That faith began to crumble in late May of 1990, when Dennis Zaller, a senior member of Hood's group who is now at Merck, Sharpe, & Dohme Research Laboratories, and a colleague went to Hood with what they thought was clear evidence of wrongdoing. Zaller had been trying to extend some of Kumar's work, and in the process tried to repeat one of his experiments. He couldn't. He then showed Kumar's original paper, which had been published in the December 1989 *Journal of Experimental Medicine (JEM)* to Mike Nishimura of the Hood group. Nishimura was struck by what everyone in the lab, including Hood, anf the *JEM* peer reviewers had missed the first time around: a key figure appeared to be falsified.

Says Zaller: "If you look at the [Southern] blot it is unmistakable." It was supposed to show DNA from several different cell lines that all had essentially the same pattern—namely, a rearrangement in the T cell receptor gene locus. But Zaller and Nishimura could tell by looking at the artifacts, the little spots that crop up on the gels, that Kumar had used data from just a few cell lines—one lane in each—duplicated repeatedly and labeled as if they came from many more cell lines.

A stunned Hood immediately informed the chairman of the biology division and other university officials, who began an inquiry into the allegations—the first step to see whether a full investigation is warranted. While the inquiry was getting under way, Hood enlisted the senior scientists in his group to perform an internal review of all of Kumar's work; Hood later gave their report to the investigation committee. He also asked others to the lab to try to repeat the *JEM* experiment. They couldn't.

But that wasn't the only devastating finding. Im the process of reviewing Kumar's data, the Hood group looked into some of Urban's work as well, as he was a coauthor on some of Kumar's papers. To their dismay, they quickly spotted what looked like a problem in his work, too—a problem that appeared to be unrelated to Kumar's misdeed. Hood found himself in the unenviable position of telling university officials that his lab might have a second case of misconduct on its hands. Caltech vice-president and provost Paul Jennings launched a separate inquiry, which got underway on 20 August 1990.

The Kumar Investigation

When Hood confronted Kumar, asking him to provide the original data and explain how he had constructed the Southern blot, Kumar reportedly did not deny doctoring the figure but did deny any intentional fraud. Instead, he insisted that he had only been trying to create a more attractive image and that he did not know this sort of duplication was unacceptable practice.

Indeed, Sercarz says Kumar sought advice from Urban, his "mentor" in the lab, on the propriety of duplicating lanes but apparently misunderstood what Urban told him. Says one Caltech source: "His rationale was essentially, 'I was young and naive.'"

Sercarz, for one, buys that argument, explaining that "Vipin had never prepared a paper before." In India, where he studied at the Institute of Science in Bangalore, his adviser wrote most of his

---

* Roberts L. (1991) Misconduct: Trial by Fire, *Science*, **253**: 1344-7. Reprinted with permission from *Science* Copyright 1991 American Association for the Advancement of Science.

thesis, says Sercarz. And when Kumar went to Harvard for his first postdoc, says Sercarz, his professor, Debajit Biswas, prepared all the papers and figures—a fact Biswas confirms. Says Sercarz, "Vipin arrived at Caltech a very, very green fellow. Vipin did not know what to do with lanes that were irregular. He wanted to rationalize it to produce an aesthetic figure." Sercarz notes that Kumar made no effort to hide the telltale artifacts.

But the inquiry committee, which met with Kumar, decided just one week later, on 8 June, that a full-blown invesigation was warranted. Jennings set up a committee of four members of the biology division to investigate. It began working on 13 June.

During the investigation, Kumar was relieved of his duties in Hood's lab, though he retained his appointment there. Says Jennings, "We did not want to act until the investigation was complete." About that time, Hood and Jennings decided that, rather than wait for the results of the investigation, Hood should retract the *JEM* paper, since Kumar had admitted duplicating the lanes, though he denied fraud.

The investigation had come at an extremely awkward time for both Kumar and Caltech. Nearing the end of his postdoc, Kumar had applied for several jobs, with strong recommendations from Hood. After considerable soul-searching, Hood and Jennings decided they had no choice but to notify the universities to which Kunar had applied, along with the journals that had published the suspect work, coauthors, the National Multiple Sclerosis Society, which had given him a fellowship, and, as required, the NIH, which had funded the work, and the National Science Foundation, which supports Hood. "We tried to do it as confidentially as possible," says Jennings, but before long the community was abuzz. Wasnigton University in St. Louis, which had already offered Kumar a job, withdrew its offer.

The investigation turned up problems with another altered Southern blot in a paper submitted to *Cell*; that paper was withdrawn before publication. Perhaps the biggest problem was that much of Kumar's data was missing, say several sources. Kumar maintains that two of his lab notebooks, containing his scintillation counter data, were stolen—in fact, he informed Hood about the theft before the fraud investigations arose. Even so, says Hood, Kumar never informed him that all of his data were missing for one paper, a paper on autoimmunity published in the February 1990 *Proceedings of the National Academy of Sciences*.

The Urban Investigation

The Urban case was more complicated, almost from the outset, says Caltech sources. Hood's group first turned up what looked like a fairly minor problem with a figure, which prompted the inquiry, says Hood. But the investigators soon encountered more serious problems as well. Caltech decided to extend the inquiry to give Urban time to respnd to all the charges. Thus, the full investigation did not get underway ... until 16 November, with a committee of five faculty members, four from inside the department and one from without. Those familiar with the case say the committee found no sign of collusion between the two postdocs but multiple problems in Urban's work, as in Kumar's.

The problems with Urban proved trickier to nail down initially. "The smoking gun did not exist in the same way," says one source, referring to Kumar's Southern blot. As in Kumar's case, much of Urban's data was missing—discarded, Urban told the committee, when he moved to the University of Chicago in March. What the investigators did find was the final version of a paper exploring a mouse model for multiple sclerosis, published in the 20 October 1989 issue of *Cell*, containing different data from the one that had been submitted for review. The committee concluded that the data in the first draft were fabricated—a charge Urban reportedly did not deny, though he did deny any intent to deceive. According to one official close to the case, Urban said he intended to do the work and assumed he knew how it would turn out. "All along, he claimed he was just trying, because of pressure, to speed the review process and that he never intended to publish without real data." No one in the lab, including Hood, noticed thet discrepancy between the review draft and the final version, not did *Cell*.

Appendix B

Ethics and Scientific Research
Pre/Post-Test Evaluation Form

Scorer's Name_____ Student's Name_____

For each question, detemine if student has achieved NO MASTERY (1), SOME MASTERY (2), or COMPLETE MASTERY (3)

1. Student has identified relevant moral agents: Kumat, Hood, Zaller, Cal Tech, journal staff or reviewers, Washington University.

1	2	3
(fewer than 3)	(3-5 agents identified)	(more than 5)

2. Student states that Kumar's duplication of gel lanes violates accepted scientific practice.

1	2	3

3. Student identifies Kumar's lack of honesty to his lab colleagues and journal reviewers in that he didn't reveal the manipulations used in the presentation of the data.

1	2	3

4. Student notes that Kumar should have written his own thesis and should have prepared his own papers during his first post-doc.

1	2	3
(doesn't note)	(notes)	(also notes improper action of thesis advisor in writing thesis for him)

5. Student states that Urban violated accepted scientific practice in discarding data relevant to an investigation in progress.

1	2	3

6. Student states that Urban violated accepted scientific practice in presenting fabricated data in the review copy of his journal article.

1	2	3

7. Student acknowledges that Hood has a responsibility for training and oversight of post docs and for the quality of the work produced in the lab.

1	2	3

8. Student recognizes that Hood has a responsibility to follow up on the initial accusations of sloppy science on the part of Urban and Kumar.

1	2	3

9. Student discusses to what extent Hood was meeting his responsibility to follow up after the fabricated gel was discovered: initiating investigation in lab,

1	2	3
(not mentioned)	(some mention)	(recognizes responsibility)

notifying department chair,

1	2	3
(not mentioned)	(some mention)	(recognizes responsibility)

notifying funders, notifying journals,

1	2	3
(not mentioned)	(recognizes that journals must be notified eventually)	(recognizes that notifying journals prior to investigation may at times be morally problematic)

notifying institutions to which Kumar applied.

1	2	3
(not mentioned)	(recognized that institutions should be notified after someone is found guilty of misconduct)	(recognizes that notifying institutions prior to investigation may itself be morally problematic)

10. Student discusses whether Hood had a responsibility to know that Urban's data had changed from the review to the publication copy of his article.

1	2	3
(not mentioned)	(talks about Hood but does not recognize that he has a role in oversite)	(recognizes responsibility)

11. Student discusses Zaller's action in notifying Hood that he could not follow up on Kumar's work and that Kumar's figure was falsified.

1 2 3

12. Student understands Cal Tech's responsibility to investigate and report.

1 2 3

13. Student understands Kunar's gel lane duplication could have been noticed by Journal editors or reviewers.

1 2 3

14. Student understands that Urban's changed data could have been noticed by Journal editors prior to publication.

1 2 3

15. Student appreciates Washington University's difficult position in determining whether or not to rescind Kumar's job offer when he was under investigation.

1 2 3

16. Student discusses additional morally relevant features of the case (that we haven't thought of ourselves).

1 2 3

17. Student discusses additional and totally irrelevant features of the case.

1	2	3
(discusses many)	(discusses some)	(discusses none)

APPENDIX C
Science and Ethics Diagnostic Test

Scenario #1

A graduate student on the verge of finishing her degree is competing for a research and teaching position at a major university. Her publications, on which her qualifications for the job are somewhat dependent, are sparse.

To help her out, her thesis advisor adds the student's name as co-author to a manuscript reporting research to which the student has not significantly contributed. That extra publication tips the scales in her favor so that she rises to the top of the list of candidates and is offered the job.

Scenario #2

John Sato, a young technician in an immunology lab, looked forward to a paper recently published by Michelle Greer, one of the postdoctoral fellows in his lab. John assisted Michelle in preparing the electron micrographs thst provided the empirical support needed for Michelle's central thesis. John was to receive photo credits and he was excited about seeing his name in print.

When the journal arrived in the lab, John noticed that the micrographs were not quite as he remembered them. He recalled that he and Michelle had had some difficulty in interpreting which of the surface structures visible in the micrographs were irrelevant to the thesis of

Michelle's research and which were important. In the published images, the differences in the structures seemed greatly exaggerated and more uniformly corresponded to their description in the text than John recalled being the case. Knowing that Michelle had the original photographic materials, he wondered if Michelle hadn't done something to enhance them.

John was reluctant to challenge Michelle, but the elctron micrograph was credited to John and he believed that it was unfair for Michelle to alter his work without consultation. When John asked Michelle, she admitted rather unabashedly that she had employed some "enhancement techniques" to improve the clarity of the images. The micrographs were for publication, after all, Michelle said, and needed to be of the best quality and clarity. There was nothing wrong with this, Michelle asserted, as long as the images were merely "enhanced" and had not altered to change their fundamental characteristics.

Scenario #3
(The case which follows was produced in 1993 by a research group at the Illinois Institute of Technology in the course of considering its own ground rules.)

The following memo is presented at the beginning of the academic year to graduate students at a large and productive lab in a mid-western university:

To: New and Used Graduate Students in the Laboratory of Last Resorts:
From: Director Drake
Subject: General Rules

Welcome to our laboratory. As you know, research in this lab is funded by grants from the NIH, NSF and other agencies. The projects have specific aims and a detailed research plan stated in the grant applications. Departure from these aims and plans requires reapplication for the grant funds. We would only do this if the original ideas prove early to be without merit.

Students are not free to pursue activities of their own design, unless they fit the aims and plans of the project that supports them. In accepting this fact, you are surrendering a significant amount of intellectual freedom.

I agree to provide, as long as grant funds are available:
1. Your tuition.
2. A stipend for you to live on.
3. Excellent laboratory facilities, including all necessary computers, instruments, equipment, tools, supplies and desk space.
4. Superior research training.
5. Thesis ideas and guidance.
6. A long term commitment to your career goals.

You agree, as the Laboratory's highest priority is continued funding,
I may:
1. Set your daily work schedule.
2. Determine your research.
3. Personally present your work where and when I deem appropriate.
4. Decide what and when to publish.
5. Decide the authorship and order of names on all publications.
6. Determine your readiness for PhD qualifying, and final examinations.
7. Approve your committee membership.
8. Approve any communication you have with other laboratories.
9. Have excusive ownership of your data — before and after you leave the laboratory.
10. Restrict your lunches to the usual banana and tuna sandwich.

APPENDIX D
Ethics
Final Exam Instructions

The purpose of this final exam is to help assess what difference this class has made in the way that you think, dealing with ethical problems.

The diagnostic test that you completed at the beginning of the term is attached.

1) Please review the case, the instructions you received at the beginning of term and your responses.

2) Analyze your initial response. Describe how your thinking has changed. Be sure to discuss understandings or information that you have now that you didn't have at the beginning of term.

   This is your opportunity to consider how your thinking has changed. Please notice changes in HOW you think as well as any changes in WHAT you think. It may be that you reach the same conclusion now that you did in the beginning of the term, but that you think about the situation in a different way.

3) Please attach your diagnostic test to the final exam.

   Please keep in mind that you are NOT being asked to repeat the assignment from the beginning of the term. You are being asked to analyze how you initially responded to that assignment.

APPENDIX E
Learning Environment Preferences

This survey asks you to describe aspects of your IDEAL learning environment. Your opinions are important to a study of teaching and learning in college. Please take this seriously and give your responses some thought.

The survey consists of five sections, each representing a different "domain". For each area, please rate each statement in terms of its significance or importance on a scale of 1-4 (with 1 meaning "not at all significant" and 4 meaning "very significant"). Once you have rated all of the items in a section go back through the list and rank the three items most significant to you. Remember to think about the **IDEAL** learning environment for **YOU**.

Please mark all your answers on the separate sheet provided.

DOMAIN ONE: COURSE CONTENT/VIEW OF LEARNING
My Ideal Learning Environment Would:

1. Emphasize basic facts and definitions.
2. Focus more on having the right answers than on discussing methods or solving problems.
3. Insure that I get all the course knowledge from the professor.
4. Provide me with an opportunity to learn methods and solve problems.
5. Allow me a chance to think and reason, applying facts to support my opinions.
6. Emphasize learning simply for the sake of learning or gaining new experience.
7. Let me decide for myself whether issues discussed in class are right or wrong, based on my own interpretations and ideas.
8. Stress the practical applications of the materials.
9. Focus on the socio-psycho, cultural and historical implications and ramifications of the subject matter.
10. Serve primarily as a catalyst for research and learning on my own, integrating the knowledge gained into my thinking.
11. Stress learning and thinking on my own, not being spoonfed learning by the instructor.
12. Provide me with appropriate learning situations for thinking about and seeking personal truths.
13. Emphasize a good positive relationship among the students and between the students and teacher.

DOMAIN TWO: ROLE OF THE INSTRUCTOR
In My Ideal Learning Environment, the Teacher Would:

1. Teach me all the facts and information I am supposed to learn.
2. Use up-to-date textbooks and materials and teach from them, not ignore them.
3. Give clear directions and guidance for all course activities and assignments.
4. Have only a minimal role in the class, turning much of the control of course content and class discussions over to the students.
5. Be not just an instructor, but more an explainer, entertainer and friend.
6. Recognize that learning is mutual—individual class members contribute fully to the teaching and learning in class.
7. Provide a model for conceptualizing living and learning rather than solving problems.
8. Utilize his/her expertise to provide me with a critique of my work.
9. Demonstrate a way to think about the subject matter and then help me explore the issues and come to my own conclusions.
10. Offer extensive comments and reactions about my performance.
11. Challenge students to present their own ideas, argue with positions taken, and demand evidence for their beliefs.
12. Put a lot of effort into the class, making it interesting and worthwhile.
13. Present arguments on course issues based on his/her expertise to stimulate active debate among class members.

DOMAIN THREE: ROLE OF STUDENT/PEERS
In My Ideal Learninig Environment, as a Student I would:

1. Study and memorize the subject matter—the teacher is there to teach it.
2. Take good notes on what is presented it class and reproduce that information on the tests.
3. Enjoy having my friends in class, but other than that classmates don't add much to what I would get from a class.
4. Hope to develop my ability to reason and judge based on standards defined by the subject.
5. Prefer to do independent research allowing me to produce my own ideas and arguments.
6. Expect to be challenged to work hard in class.
7. Prefer that my classmates be concerned with increasing their awareness of themselves in relation to the world.
8. Anticipate that my classmates would contribute significantly to the course learning through their own expertise in the content.
9. Want opportunities to think on my own, making connections between the issues discussed in class and other areas I'm studying.
10. Take some leadership, along with my classmates, in deciding how the class will be run.
11. Participate actively with my peers in class discussions and ask as many questions as necessary to fully understand the topic.
12. Expect to take learning seriously and be personally motivated to learn the subject.
13. Want to learn methods and procedures related to the subject—learn how to learn.

DOMAIN FOUR: CLASSROOM ATMOSPHERE/ACTIVITIES
In My Ideal Learning Environment, the Classroom Atmosphere and Activities Would:

1. Be organized and well-structured—there should be clear expectations set (like a structured syllabus that's followed).
2. Consist of lectures (with a chance to ask questions) because I can get all the facts I need to know more efficiently that way.
3. Include specific, detailed instructions for all activities and assignments.
4. Focus on step-by-step procedures so that if you did the procedure correctly each time, your answer would be correct.
5. Provide opportunities for me to pull together connections among various subject areas and then construct an aequate argument.
6. Be only loosely structured, with the students taking most of the responsibility for the structure there is.
7. Include research papers, since they demand that I consult sources and then offer my own interpretation and thinking.
8. Have enough variety in content areas and learning experiences to keep me interested.
9. Be practiced and internalized but be balanced by group experimentation, intuition, comprehension and imagination.
10. Consist of a seminar format, providing an exchange of ideas so that I can critique my own perspectives on the subject matter.

11. Emphasize discussions of personal answers based on relevant evidence rather than just right and wrong answers.
12. Be an intellectual dialogue and debate among a small group of peers motivated to learn for the sake of learning.
13. Include lots of projects and assignments with practical, everyday applications.

DOMAIN FIVE: EVALUATION PROCEDURES
Evaluation Procedures in My Ideal Learning Environment Would:

1. Include straightforward not "tricky" tests, covering only what I had been taught.
2. Be up to the teacher, since s/he knows the material best.
3. Consist of objective-style tests because they have clearcut right or wrong answers.
4. Be based on how much students have improved in the class and on how hard they have worked in class.
5. Provide an opportunity for me to judge my own work along with the teacher and learn from the critique at the same time.
6. Not include grades, since there aren't really any objective standards teachers can use to evaluate students' thinking.
7. Include grading by a prearranged point system (homework, participation, tests, etc.) since I think it seems the most fair.
8. Represent a synthesis of internal and external opportunities for judgment and learning enhancing the quality of the class.
9. Consist of thoughtful criticism of my work by someone with appropriate expertise.
10. Emphasize essay exams, papers, etc. rather than objective style tests so that I can show how much I've learned.
11. Allow students to demonstrate that they can think on their own and make connections not made in class.
12. Include judgments of the quality of my oral and written work as a way to enhance my learning in class.
13. Emphasize independent thinking be each student, but include some focus on the quality of one's arguments and evidence.

LEARNING ENVIRONMENT PREFERENCES ANSWER SHEET*

COURSE NAME AND NUMBER_____DATE_____
FRESHMAN_____SOPHOMORE_____JUNIOR_____SENIOR_____GRAD_____
AGE_____ SEX:MALE_____FEMALE_____
ETHNIC HERITAGE: AFRICAM-AM_____ASIAN-AM_____HISPANIC_____
NATIVE-AM_____WHITE_____OTHER (please specify)_____

RATING SCALE:

	1 Not at all significant	2 Somewhat significant	3 Moderately significant	4 Very significant
DOMAIN 1 COURSE CONTENT	**DOMAIN 2 ROLE OF INSTRUCTOR**	**DOMAIN 3 ROLE OF STUDENTS**	**DOMAIN 4 CLASSROOM ATMOSPHERE**	**DOMAIN 5 EVALUATION PROCEDURES**
1.____	1.____	1.____	1.____	1.____
2.____	2.____	2.____	2.____	2.____
3.____	3.____	3.____	3.____	3.____
4.____	4.____	2.____	4.____	4.____
5.____	5.____	5.____	5.____	5.____
6.____	6.____	6.____	6.____	6.____
7.____	7.____	7.____	7.____	7.____
8.____	8.____	8.____	8.____	8.____
9.____	9.____	9.____	9.____	9.____
10.____	10.____	10.____	10.____	10.____
11.____	11.____	11.____	11.____	11.____
12.____	12.____	12.____	12.____	12.____
13.____	13.____	13.____	13.____	13.____

INDICATE TOP THREE CHOICES IN EACH AREA (MARK ITEM NUMBER)

COURSE CONTENT	ROLE OF INSTRUCTOR	ROLE OF STUDENTS	CLASSROOM ATMOSPHERE	EVALUATION PROCEDURES
1st_____	1st_____	1st_____	1st_____	1st_____
2nd_____	2nd_____	2nd_____	2nd_____	2nd_____
3rd _____	3rd _____	3rd _____	3rd _____	3rd _____

Appendix E, Learning Environment Preferences

APPENDIX F

PERCEPTIONS OF LEARNING ENVIRONMENTS (PLE)

1. The course included a wide variety of content areas.
2. A clear, detailed syllabus was provided and followed.
3. The class seldom used concrete, 'hands-on' activities for students (like field trips, role plays, simulations, student case discovery, etc.).
4. The teacher seemed to know and address the students by name.
5. Different teaching methods — lecture, small group discussions, films, etc. — were used.
6. The course material was organized in a way that was understandable.
7. Students were rarely encouraged to share with their classmates' personal experiences related to the course learning tasks.
8. Students were provided specific opportunities to get to know each other early in the class.
9. There was a range of types of assignments and learning experiences — for example, papers, journals, individual and group presentations, etc.
10. Students were told at the beginning of the course what kind of performance tasks — for example, comprehension, analysis, evaluation — was expected of them.
11. The course material or assignments had little to do with students' lives.
12. The course included in-class activities — like specific tasks assigned to small groups — to help students learn to trust and work with each other.
13. Reading assignments were frequent and moderate-to-long in terms of amount required.
14. The course activities seemed well planned and organized by the teacher.
15. Students were expected to read and write about theories and abstract concepts.
16. The teacher encouraged and modeled respect for diverse viewpoints; the class felt "safe" in terms of expressing different opinions.
17. Students were encouraged to contribute ideas and discuss each other's views in class.
18. Individual class sessions were organized and structured to make it easy to follow what was going on.
19. Students were asked to understand cource material with little, if any, immediate or explicit connection to their lives.
20. The atmosphere of the course was informal and comfortable — relaxed pace, informal discussion encouraged when appropriate, etc.
21. Students were asked to consider a diversity of points of view within the subject matter of the course.
22. Visual aids — transparencies, handouts, notes on the board — were used regularly.
23. Information in the course was presented in a lecture format based largely on the formal course materials (textbooks, readings, handouts, etc.).
24. Students had plenty of opportunities for one-to-one interaction with the teacher.
25. Coursework included an emphasis on the processes and methods involved in the subject area.
26. Students were given an opportunity to learn the basic concepts of the subject areas before being asked to analyze or intepret this material.
27. The teacher rarely made connections between the course material and current events or issues facing students.
28. Class members — students and teacher — felt comfortable sharing personal information when it was relevant to course materials or discussions.

29. Students were expected to take an active role in determining what went on in class.
30. Some class time was used to help students practice and learn the skills emphasized in the course.
31. Students were not asked to collaborate with their classmates in course projects or activities.
32. The course included assignments or activities which called for personal reflection — for example, journals, logbooks.
33. Students were given tasks that asked them to provide evidence to back up their opinions.
34. Material covered on the exams had been discussed in class previously.
35. Evaluation of student progress rarely focused on non-written student performance tasks (like small groups, participation, in-class presentation).
36. The teacher encouraged students to contact him/her outside of class and seemed to be accessible.
37. Testing in the course included tasks that asked students to evaluate different points of view and make a personal judgment about what s/he believed regarding the material.
38. Specific detailed instructions were given for class activities.
39. Students were asked to assess their own learning and performance.
40. The teacher provided extensive feedback to students concerning their performance in class and on tests.

PERCEPTIONS OF LEARNING ENVIRONMENTS (PLE)

Think about this course. We are interesed in your thoughtful response to the question: what kind of learning environment existed in the course? The term *learning environment* refers to a wide variety of specific aspects in the classroom, including:

* the kind of activities and assignments
* the role of the teacher in the course
* the expectations and performance requirements of the class
* the role and typical behaviors expected of the students.

We are interested in gathering your perceptions as a way of helping us to understand the climate for learning established in this class. Please take your time and give careful consideration to your answers.

COURSE ASSESSED (Dept./Number) _____ TITLE _____
CLASSIFICATION: Fr.____So.____Jr.____Sr.____Grad._____Other_____
SEX: Male _____ Female _____ AGE: _____

For each item on the instrument, use this answer sheet to record your ratings in the space provided below.

RATING SCALE:

Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree
1	2	3	4	5
Never	Occasionally	Sometimes	Often	Regularly

* Please note that answer spaces run horizontally.

1. _____	**2.** _____	**3.** _____	**4.** _____
5. _____	**6.** _____	**7.** _____	**8.** _____
9. _____	**10.** _____	**11.** _____	**12.** _____
13. _____	**14.** _____	**15.** _____	**16.** _____
17. _____	**18.** _____	**19.** _____	**20.** _____
21. _____	**22.** _____	**23.** _____	**24.** _____
25. _____	**26.** _____	**27.** _____	**28.** _____
29. _____	**30.** _____	**31.** _____	**32.** _____
33. _____	**34.** _____	**35.** _____	**36.** _____
37. _____	**38.** _____	**39.** _____	**40.** _____