

Collaborative Examinations: A Way to Help Students Learn

Douglas R. Haddock

Experimenting with Collaborative Exams

Law students have significant professional incentives to do well in law school. Because the breadth of their career options is often determined by their grade point averages, grades are immensely important to most students, especially during the first year of law school.¹ As the primary basis of grades, examinations carry great weight; exams are potentially a significant source of learning, as well as a means of evaluating student performance. So I have frequently experimented with examinations, trying to make them an important part of my teaching and a significant component of my students' learning process.²

My most successful experiments with exams change the testing process from a three-hour essay and short-answer exam to a more creative and educational exercise that occurs over a period of several weeks. For the past sixteen years I have used a simple process to motivate students to review the subject matter and to create pertinent legal problems as they study for the examination. During the semester I distribute to students a document containing factual and legal information.³ It presents a story or stories with the potential for numerous legal disputes and problems relevant to the subject matter of the course. It also

Douglas R. Haddock is a professor of law at St. Mary's University.

I thank my colleague Mike Ariens for his useful comments on early drafts of this article. I also express my gratitude to my secretary, Aurelia Vincent, for her consistent and patient help in administering the exams discussed here, and to Yvonne Olfers for her effective and cheerful help in compiling information. Most of all, I thank the many wonderful students who have tolerated, and sometimes even enjoyed, my experiments with examinations for almost three decades.

1. Employers of lawyers have unfortunately pushed professional evaluation of students more and more to the first and second semesters of law school, placing extraordinary emphasis on grades earned by students at the very beginning of the law school experience.
2. I recognize that "experimenting" with examinations is risky business precisely because grades *are* so important to law students. I have always been frustrated by the emphasis most students place on exams, but it is difficult to fault them. I wonder whether drastic modifications in the way we educate law students are in order. I agree with many of the points made by Bethany Rubin Henderson, *Asking the Lost Question: What Is the Purpose of Law School?* 53 J. Legal Educ. 48 (2003).
3. I ordinarily distribute this document about two weeks before the last class session. Other teachers who use a similar process deliver the information at the beginning of the semester.

contains a variety of statements about the law, usually including a number of statutory provisions pertinent to the course.⁴ In class and in the instructions that accompany the document, I explain that some or all of the examination problems will be based on these facts and governed by applicable legal doctrine. I encourage students to study the document, alone and in groups, and to practice imagining legal problems that could arise from the situations presented. Students understandably take this advice seriously and often become intensely engaged in the process, realizing that they have an opportunity to anticipate and thus prepare for actual examination problems.

I invite students to ask questions about the information during the last two weeks of class. I respond candidly to their questions, helping them understand the problems and issues they identify.⁵ From numerous conversations, I believe that many of them begin to see the variety of potential legal problems and to understand some of the complexities of resolving those problems. Many students work in groups, brainstorming at some length to extract potential exam problems from the document. In studying the information, and in discussing the material with other members of the class, they develop a better sense of a lawyer's work. I believe that, for many of them, this focused review becomes the most valuable learning experience of the semester. Their imagination often takes them far beyond what I could cover in a three-hour examination.

In many courses I have used this focused-review approach to exams combined with a traditional three-hour final examination. Some years ago I began to consider collaboration *during* the exam as a simple extension of the process I had seen working well in focused review. Students had often worked in groups, and many had commented on how much they had learned as they tried together to anticipate and resolve problems they might encounter on the exam. I thought that the benefits of collaboration might be magnified if I allowed students to work together in composing their responses to the actual examination.

During the past nine years I have given many take-home exams on which I allow students to collaborate in composing answers. I usually combine the take-home exam with the focused review, distributing information on facts and law during the last few weeks of the semester. During the exam period students receive the problems and have a few days, usually forty-eight hours, to complete their responses and turn them in. In some courses I have divided the exam into two parts, requiring individual work on one part while allowing collaboration on the other part. On other occasions I have permitted collaboration on the entire exam. In all but one instance, I have imposed limits on the

4. I believe an emphasis on statutes is especially important in teaching first-year students. The case method of study sometimes creates the feeling among impressionable first-year students that answers are generally found only in judicial opinions and that the interpretation of statutes is not a task for lawyers and law students.
5. If students come up with a problem I had intended to use on the exam, I still respond to the question. Depending on the discussion that ensues, I will decide whether to use the problem, not use the problem, or modify the problem for the exam.

number of students who can work together. In no case have I *required* any student to collaborate with others.⁶

The Process

Our school assigns each student a four-digit exam number each semester. Until all grades have been turned in for the term, faculty are not permitted to match exam numbers with students. I decided that during the time I was grading the exam I also did not want to know whether the paper I was reading had been written by an individual or by a group of students. With the help of a patient secretary, I use a fairly simple procedure to accomplish that objective. Students are instructed not to put any information on the exam papers that would communicate to me either the identity of the author or authors or the number of persons who worked on a particular paper. When they finish the exam, they simply turn the paper in to my secretary. She assigns it an arbitrary control number and writes that number on the paper. Students give her their exam numbers, and she records them on a chart linking the exam numbers to the appropriate control numbers. After I have finished grading the papers, she gives me that chart and I can then assign the appropriate grade to each exam number.⁷

In my two most recent experiments with collaborative exams, I have broken the exams into two major components. In the first part (about 70 percent of the grade) students were permitted to work with others. In the second part, however, collaboration was not allowed. Students turned in separate papers for part 1 and part 2 of the exam, whether they worked with others or not on part 1. They placed their regular four-digit exam number on part 2 and the control numbers were placed on part 1.

The first course in which I permitted collaboration was first-year Property. I devote some time in Property to law and economics, and I believe it was during that segment of the course that I first considered the possibility of permitting students to work together not only in studying the information on which the exam would be based but also in composing answers. It occurred to me that a collaborative examination would teach the students a great deal about some economic concepts, such as transaction costs and efficiency. It would also, I thought, give them more accurate glimpses of what the practice of law is like and the essential process of integrating practice and theory.

I assumed that my colleagues would find my idea controversial, even radical and inappropriate. I was also interested in my students' reaction. Initially, I said nothing to faculty colleagues but discussed the idea with my students on

6. During a presentation I made on this subject at the 2001 AALS conference on New Ideas for Experienced Teachers in Calgary, one participant suggested that perhaps I should require that all students collaborate with others on the exam and assign them to groups. I am reluctant to require collaboration. One of my concerns about my current process, however, is the possibility that some students may be unable to find willing partners, and assigning partners would resolve that problem.
7. I have instructed my secretary that I am not permitted to see the chart that links exam numbers to control numbers before I have finished grading the papers. She has taken this instruction very seriously, essentially requiring me to establish to her satisfaction that I have finished the grading before she will turn over the chart.

several occasions. Many students enthusiastically embraced the idea, apparently assuming there would be strength in numbers. Others reacted negatively, and some remained undecided. The major concerns expressed were variations on a free-rider theme: quite a few students assumed that collaboration would benefit lazy students who could receive a better grade than they deserved on the strength of their collaborators' efforts. This was also one of my primary concerns about permitting collaboration, but I tend to give students the benefit of the doubt, assuming that each student in any particular class is committed and diligent. By the end of each semester I have reconsidered that position, but I begin each course as an optimist.

Whatever the truth about students' diligence and capabilities, I assume that permitting them to collaborate on examinations or other graded exercises compromises, to some degree, the evaluative function of exams and grading.⁸ When four students work together and turn in one exam paper, it may be impossible to prove or disprove empirically that every person in the group actually *deserves* the grade earned by group effort. In fact, in my experience it appears clear that the grades some students receive by working with others are different from the grades they would have received had collaboration not been permitted. But I also believe that, for a substantial majority of students, the impact of collaboration on grades is not very significant.

In that first experiment with collaboration, I thought about whether to limit the number of students who could work together during the exam. I decided not to impose any limit. At that time our school had a requirement that the median grade in large classes must be B-. No other grades were absolutely required.⁹ Theoretically, then, some students might convince 100 percent of the class to work together on the exam, guaranteeing each student a B- grade. But that seemed unlikely. From our classroom discussions of transaction costs and other economic considerations, I assumed that only a few students would be interested in such an idea and that it would be impossible for those students to persuade even a large number of others—let alone the entire class—to work together on the exam. I was also quite certain that some confident students would not settle for a B-.¹⁰

8. This depends, I suppose, on what we are evaluating. The traditional approach to law school work is highly individualistic, and it is this concept that is compromised. I believe there is a need for individual evaluation in law school, and at this point I would not favor basing all student evaluation entirely on collaborative exercises. An obvious purpose of including individual work on the exam is to mitigate the free-rider problem. In fact, a participant in the Calgary conference pointed out that I am a free rider among my colleagues, who have more exams to grade.
9. Our grading policies have changed. In first-year courses we are now *required* to give grades of A, A-, or B+ (as a group) to no fewer than 10 percent and no more than 20 percent of the students in the class. The same percentage requirements also apply to grades of C-, D, and F, our three lowest grades.
10. As I recall my thinking at the time, I assumed it was unlikely that more than six or seven students would agree to work together. I also thought that no individual or group would seriously try to persuade the whole class to work together on the exam. In retrospect, those assumptions were naive, for a fair number of people in a large group might be sufficiently risk-averse to accept a virtually guaranteed B- rather than face the possibility of a lower grade.

Relatively few exam papers were turned in for that first collaborative exam, and I was a bit surprised to learn that a few students had indeed orchestrated somewhat vigorous efforts to get the entire class to turn in one exam paper. As I expected, they did not succeed. But to my amazement they did persuade a group of thirty-three students, over a third of the class, to work together. As one would expect, the paper produced by that large group did fall right in the middle of the pack for grading purposes, and I was not faced with a potentially difficult grading problem. From conversations with a number of those thirty-three students, I sensed that the test was by no means an entirely pleasant experience for the group, but it did teach them a lot about working with others and the costs of large-group collaboration.

Since my first experiment, I have imposed a limit on the number of students who may work together.¹¹ I also offer many comments in class about the advantages and disadvantages of collaboration on such a project. I try to remain impartial, but I have been somewhat surprised at how many students choose to work with others. Sometimes I have candidly told them that I probably would not work with other students if I were in their shoes.

Evaluating the Experiment

In attempting to assess the effects of collaborative examinations, I have considered several questions. Using examples, data, and feedback from several different exams, I address here some of the issues and offer some conclusions under the general topics of student decisions on collaboration, the impact of collaboration on grades and students, learning theory, the students' evaluation of the process, and my tentative judgment about the experiment.

Student Decisions on Collaboration

During the 2000–01 academic year, students in my Property course were permitted to work with up to four other students on both fall and spring semester exams.¹² Collaboration was not required, but most students chose to work with others. In the fall semester only 21 of 156 students (in two sections) worked alone on the exam. The Property exam was their last exam of the semester, and many of these first-year students seemed grateful for the change of pace from three-hour exams to a three-day collaboration with colleagues, which, they initially thought, would allow them some escape from the pressures of time. In this same group of students, down to 152 in the spring semester, only six worked alone on the spring semester exam. A probable reason for the shift from solo efforts to collaboration was a perception that the fall semester exam was too long and too difficult for one student.¹³ The

11. But I haven't ruled out the possibility of another exam without limits.
12. Unlike many schools, St. Mary's still has yearlong courses in torts, contracts, and property, as well as legal research and writing.
13. Students were allowed 72 hours in the fall and 48 hours in the spring to complete the examinations. Invariably, some students have said that these tests are too long and too difficult for one person to handle alone. I try not to make them much "longer" than an exam I would give in a three-hour time period, although some of them probably were too long to be completed in three hours. I explain to the students that many legal problems, including exam questions, will expand to fill whatever time is available, which is one reason my exam

makeup of the groups changed substantially from fall to spring. The difference is particularly noticeable at the extremes, with only a third as many students working alone in the spring, compared with the fall, and eight groups of five in the spring semester compared to only one in the fall.

The dynamic of choices about collaboration probably changes in significant ways after a student's first semester of law school. Starting with the second semester, students have track records, and the choice of collaborators becomes a different problem. Once students have received some grades, they are probably placed in the minds of their fellow students in some type of order, ranging from the apparently superior students to those who are failing or close to it. I assume that some students who have established a strong record in the first semester would be reluctant to collaborate with those who did considerably less well. And students with relatively low grades in the first semester might have difficulty finding partners during the second semester. In my conversations with a number of students, I have noticed that those who received a lower grade in my course than in their other courses tend to assume, perhaps with justification, that they were hurt by their collaborators. I believe the process of shifting alliances from fall to spring semesters in a yearlong course is interesting and deserves more study.¹⁴

I did not teach Property in 2002 or 2003, but in the fall semester of those years I taught first-year Federal Civil Procedure, a four-credit, one-semester course. I taught two sections of Procedure in 2002 and one section in 2003. Compared to the 2000–01 Property course, there was a significant shift in 2002 and 2003 from relatively large groups to individual effort. Whereas only 4 percent of the spring semester Property students worked alone on the exam, 24 percent worked alone on the Procedure exams. At the other end of the spectrum, only 35 and 37 percent of the Procedure students worked in groups of four or more, while half of the students did so in Property. One significant difference between the courses was that individual work was required on one part of both Procedure exams; that may explain, in part, the difference in choices made by students. Perhaps some students decided that they might as well work alone on the entire exam since they were required to do so on part of it.

seems substantially longer than their other exams. I also suggest that working on the exam with a group of people will usually consume more time than working alone. The observation that the test is too long and too difficult for one person almost always comes from people who collaborated with other students. I seldom hear that comment from a student who worked alone. And certainly I have had good exams from students who worked by themselves.

14. Although I have used collaborative exams in second- and third-year courses, the dynamics of choosing collaborators after students have received law school grades, which is really an issue for all courses after the first semester, concerns me. I don't want to contribute to an invidious class system, and I wonder whether the use of collaborative exams after the first semester might exaggerate distinctions based on grades. But I have not heard negative comments on this point from students. Some of my colleagues have expressed serious reservations about collaboration in first-year courses. Their concerns seem to be based on a perceived need to identify individual students who may not be willing and/or able to handle law school work on their own. They assume that collaborative exams could reduce attrition and keep students in law school who should not be there. I suppose this might be a problem if most or all first-year grades were based on collaborative work, but I don't think it has been a significant problem with my students up to this point.

The Impact of Collaboration on Grades and Students

Going into my experiment, I had three general questions about the effects of allowing collaboration on law school exams.

- What impact does collaboration have on the distribution of grades?
- How does collaboration affect students, in terms of grades and otherwise?
- Is there merit to the basic thesis underlying my experiment, which is that collaboration on exams can enhance students' learning experience?

The Effect of Collaboration on the Distribution of Grades

I was concerned by the possibility that grades could be affected in such a way that I would have difficulty fitting them into institutional grading requirements. St. Mary's uses a ten-step system with grades of A, A-, B+, B, B-, C+, C, C-, D, and F. In the fall of 2000–01 our required median grade for first-year courses was B-. For the spring semester it was changed to C+. In fall 2000 there was no required distribution of grades other than a limit on high grades and low grades; no more than 15 percent of the students in a course could receive grades of A and A- or grades of C-, D, and F. There was no mandatory curve as such, and no requirement that we give any of these grades (other than at least one appropriate median grade).¹⁵ Although grading constraints do not make collaborative exams unworkable, they do raise several issues. For example, if I am limited by school policy in a particular class to giving no more than four A grades, and one group of four students writes the best exam in the course but another student or group of students writes an exam that I believe also deserves an A, I have a problem (shared by the second group of students).

Property, 2000–01. I have not wanted to discourage students from working alone, but I am concerned that students might feel pressured into collaboration because they think groups will fare better on the test than individuals will. As I've said, many of my 2000–01 Property students apparently assumed this was the case. In fact, I believe that assumption is not generally justified.

In the two Property exams in 2000–01, the highest grade received by an individual was a B+ in the fall and a B in the spring. These results arguably call into question my assumption that individuals would fare as well as groups. In the fall semester individuals received none of the highest grades and a disproportionate share of the low grades. Two groups of 4 students received grades of A; one group of 4 students and one group of 2 students received grades of A-. In the spring semester those working in groups of 2, 3, or 4 (23 students in all) received all of the top three grades. On that test, however, 8 students working in groups of 2 and 4 also received all of the D's and F's, calling into question another of my assumptions, which was that collaboration would cause performance to gravitate slightly toward the mean.

15. This requirement has now been changed. See *supra* note 9.

In the fall 21 students worked alone, but in the spring that number fell to 6. I wondered whether students had been discouraged from working alone in the spring semester because of the fall semester exam. I think some students thought that exam was too difficult, and perhaps it was.¹⁶ Theoretically, if I am correct in assuming that collaboration requires more time than individual work, students should not have been prompted to shift to collaboration because of the pressures of time. Some students have realized this. Unfortunately, what has happened in some groups is a division of labor rather than collaboration: students have divided up various parts of the exam and then pooled their results. In most cases, I think, there is some discussion about the work each student is producing, but it may be minimal. True collaboration takes time. In any event, although all of the spring grades at the two highest levels and the two lowest levels resulted from collaborative effort, I suspect that these numbers have little statistical significance. Very few students worked alone on that exam, and those who worked in groups that did not fare well did not have particularly good records in other courses.¹⁷

In general, the distribution of grades in my 2000–01 Property courses was similar to the bell-shaped curve I have usually found in courses where collaboration is not permitted. As with many noncollaborative exams, there are slight deviations from the bell shape, and some of those deviations might be related to collaboration. For example, in the fall semester more students received A than A-, and more students received F than D. Again in the spring more students received F than D, and 18 of 152 students received A-, compared to only 3 receiving A and 2 receiving B+.

Procedure, fall 2002 and fall 2003. In both 2002 and 2003 students in my first-year Procedure course were required to work alone on part of the exam. The exams involved essay problems, short discussion problems, and statement evaluation problems. Students were permitted to collaborate on part 1 of each exam, which constituted about two-thirds of the exam for grading purposes. Because there was an individual component to these exams, students who worked together did not necessarily receive the same grade for the course. For example, in 50 percent (18 out of 36) of the collaborations in 2002, grades of one or more individuals within a group differed from those of other members of the group.

Compared to the grades in the 2000–01 Property courses, the curve of the grades in the two Procedure courses was closer to the shape of a bell. This slight

16. Students generally think my exams are difficult. I agree that they are difficult, but I think they are fair and effective. Comments I have heard suggest that many of my students agree that the exams are fair. Like many law teachers, I have discovered that it is hard to compose a good exam that is “easy” for students and relatively easy to produce a good exam that many students consider to be quite difficult. In any event, I do not seek to produce or to assign low or failing grades to students.
17. Groups of four and two students received failing grades, and another group of two received a D. The grade point averages of the four failing students in the other courses they took that semester were 1.9, 2.1, 2.3, and 2.4 on a 4-point scale. The averages of the failing group of two students on their other exams were 2.4 and 2.8. The two students who worked together and received a D grade had averages in their other courses of 2.3 and 2.1.

difference may not be significant, but the individual component of the Procedure exams could have had some effect. In both Procedure courses, working in groups arguably did push grades toward the mean. In 2003, for example, the top three papers and five of the best six papers were from students who worked alone on the entire exam. Nine of the 19 students (47%) who worked alone received grades at or below the median, whereas 33 of the 58 (57%) who collaborated with others received such grades. It is also interesting that the grades seem to decline somewhat as the collaborators increase in number. This information doesn't prove or establish much, if anything, about the effect of collaboration on grades, but it does suggest that collaboration will not necessarily help students get a better grade. In contrast to the results in the Property exams of 2000–01, the slightly lower grades of those who collaborated in 2003 may indicate that some of the better students chose not to collaborate and those who were less well prepared were inclined to collaborate.

The Impact of Collaboration on Students

I have been concerned about two kinds of effects collaboration could have on students. For one thing, I assumed that collaboration would have some effect on some students' grades. I have examined that assumption by comparing grades in my courses with grades the students received in other courses during the same semester. I was also interested in the effects of collaboration on the students in more subjective ways. Were their experiences in working with others positive or negative? Did the collaboration teach them valuable lessons of any kind? Did it enhance their understanding of the subject matter? Were their attitudes toward other people affected? In an attempt to gather information on this second type of effect, I have used surveys asking students to respond to various questions and to offer comments on the testing process. Those surveys have given me interesting insights as to the effect collaboration has on students and also as to the possibility that collaboration can enhance students' learning experience.

Property, 2000–01. I believe my experiment suggests that collaboration does not have a particularly significant effect on grades for *most* students but probably affects *some* students' grades. In many cases, I believe the difference—if any—between the grade a student would have received without collaboration and the grade she received working with other students would be a small one.¹⁸ But some students have received a grade in my course that is

18. I assume the reality is more complex and more nuanced than the picture I present here. I focus here on a few of the most significant variations between the students' grades in my course and their grades in other courses. Slight differences in some students' grades could be caused by almost every collaboration. It would be difficult, if not impossible, to devise any means of exposing the precise effects of each collaboration. There are, of course, various possible explanations for grade variations, including such things as subject matter, mood, and interest. I would be surprised, however, to learn that collaboration doesn't affect grades. I assume those who believe that grading is a rather precise science and that individual assessment is universally necessary would not accept collaborative exams, but the probability that grades are sometimes affected by collaboration doesn't concern me very much. As I have suggested, my primary motivation for this experiment is to explore the

quite a bit different from grades they received in other courses. In at least some of those cases, it seems likely that collaboration is at least partially responsible for the grade discrepancies.

Consider discrepancies between several students' Property grades and the other grades they received in the fall of 2000. One student made an A in Property but only a D, C-, and C in Torts, Contracts, and Procedure. At the other end of the spectrum was a student who received an F in Property along with a B in Torts and B- in Contracts and Procedure. In two other cases involving disparate grades, students were quite consistent in courses other than mine. One received A's in Contracts, Torts, and Procedure and a B- in my Property course. Another made F in Property and C+ in every other course.

These are somewhat striking examples of grade disparity. In our grading climate, I would describe grades of C and below as poor grades. Our best students, on the other hand, can compete with students anywhere, and those who receive A grades in my courses have done extremely good work. So I think that a few of my students have been very much helped by collaboration. Others were presumably hurt badly by collaboration, having received acceptable grades in other courses but a failing grade in Property.

To further understand the impact of collaboration on grades, I have also looked at all of the semester grades of individuals within collaborating groups. Looking at the grades of each participant, one gets some sense of the group dynamics. For example, three students collaborated with the student mentioned above who made an A in Property and a D, C-, and C in other courses. In courses other than Property, those three students received grades of A, A-, B+ (2), B (3) and B- (2), for an average of 3.2, compared with the fourth student's average of 1.56. The three strong students presumably carried the weakest one, and together the group was able to produce superior work.

As for the student who failed Property but had fair grades (B and B-) in other courses, the collaboration produced very poor work. That student's grade point average in other courses was 2.8, while the three companions' combined GPA in other courses was 1.9. But two of that student's three collaborators also had respectable grades (five C+ and one B-) in other courses. The fourth student, on the other hand, had an average grade in the other three courses of 1.0 (C, D, and F). Perhaps this student, for whatever reason, in some fashion dominated the collaboration, with disastrous results.¹⁹

For the spring 2001 exam, I looked at grade variations in courses other than mine, to determine whether the variations between grades in my course and

possibility that examinations can be better used to enhance learning, in addition to measuring performance. Although grades are certainly important, I don't assume they are necessarily a precise measurement of each individual student's mastery of a subject. I think subjective student evaluation, which is discussed later, may be the most helpful indicator of the merits of collaboration in terms of learning.

19. These four students each received a B- on the spring semester Property exam. Three of them, including the student with low grades in other courses in the fall semester, worked together again, adding two other students to the group. The fourth student collaborated with three other students.

grades in other courses were somewhat unusual, as I expected them to be. In fact, the variations among other grades were more substantial than I expected.

- For about two-thirds of the students, the highest grade received in courses other than Property was *at least one full grade point* above the lowest grade (C and B, for example).
- For more than one-third of the students, the highest grade received in courses other than Property was *more than one full grade point* above the lowest grade (C and B+, for example).
- For 17 of the students, the highest grade received in courses other than Property was *at least two full grade points* above the lowest grade (C and A, for example).

I then compared the average grades for other courses with the students' Property grades. Only 23 students had a variation of at least one grade point between their Property grades and their average grade in other courses, and 21 students had a variation of more than one grade point. In general, the Property grades were not necessarily out of line with the students' other grades, suggesting that collaboration may not have a particularly significant effect on grades.

- For 61 of the 152 students, the Property grade was the same as at least one other grade received that semester.
- For 62 of the students, their grades in other courses were both higher and lower than their grades in Property.
- Only 10 students' Property grades were one grade point or more away from their grades in all of their other courses.

I believe that these numbers and those for the Property I course suggest that for many students collaboration does not affect grades to any substantial degree. In some instances, however, it appears that there is an impact, and in a few cases the effect can be dramatic. It appears that the effect of collaboration on grades, if any, goes in both directions. Some students seem to be helped and others are apparently hurt by collaborative effort.

Procedure, fall 2002 and fall 2003. During the fall semester of 2002 I taught two sections (138 students) of first-year Procedure, and in fall 2003 I taught one section (77 students). On both Procedure exams I permitted collaboration, but on one part of each of them, constituting about one-third of the grade, individual work was required. Again, collaboration seemed to hurt as often as it helped. In 2002, of nine students with a relatively large disparity between the Procedure grade and their other grades, five made a higher grade in Procedure and four made a lower grade. In 2003, of seven students showing a similar disparity, three made a higher grade in Procedure and four made a lower.

There are many interesting questions raised by these and other bits of information I have gathered. Based on the information discussed above, one might expect that a few students would choose to associate with a different

group or work on their own the second time around. For example, by collaborating with two other students, one student in the fall 2000 Property class received a grade of C+ in Property I but made an A, working alone, in all three other courses that semester. Students who seemingly benefited by the collaboration, on the other hand, might have concluded that they needed to collaborate with the same people (and/or others) to improve their chances of getting a good grade. I assume these disparities resulted, at least in part, from the group dynamics at work. Perhaps in some groups the better students were rather passive, letting less well-prepared collaborators exercise too much control. In other cases, it may be that the collaboration actually dragged the individuals down, resulting in a whole that was less than the sum of the parts.

Learning Theory

Late in 2000 the National Research Council published *How People Learn: Brain, Mind, Experience, and School*, an expanded version of *How People Learn: Bridging Research and Practice*, published in 1999. As explained in its Preface, the 2000 publication, produced by the Committee on Developments in the Science of Learning, expands on the earlier work of the Committee on Learning Research and Educational Practice in exploring “the critical issue of how better to link the findings of research on the science of learning to actual practice in the classroom.”²⁰ The committee’s report focuses “especially on learning research that has implications for the design of formal instructional environments, primarily preschools, kindergarten through high schools (K–12), and colleges,”²¹ but its findings are also relevant for law schools. It was the primary text used for the AALS conference on New Ideas for Experienced Teachers in 2001.

Judging from my experience, many points made in the NRC report, which is based on the results of fundamental studies on how people learn, are pertinent to my experiment with collaborative examinations. One obvious problem for learning in most law schools is the student-faculty ratio and the large number of students in each class, particularly in the first year. It may be impossible to change the numbers, but we need better strategies for improving the learning environment. The NRC report makes the following observations about a “reciprocal teaching” approach that has been used to increase reading comprehension for some students:

The three major components of reciprocal teaching are instruction and practice with strategies that enable students to monitor their understanding; provision, initially by a teacher, of an expert model of metacognitive processes; and a social setting that enables joint negotiation for understanding. The knowledge-acquisition strategies the students learn in working on a specific text are not acquired as abstract memorized procedures, but as skills instrumental in achieving subject-area knowledge and understanding. The instructional procedure is reciprocal in the sense that a teacher and a group

20. National Research Council, *How People Learn: Brain, Mind, Experience, and School* at vi (Washington, 2000).

21. *Id.* at 5.

of students take turns in leading the group to discuss and use strategies for comprehending and remembering text content.²²

Although it is difficult to manage in large law school courses, something akin to “reciprocal teaching”—involving more small-group work, including interactions among students with and without the instructor—should be encouraged. Collaborative exams are, perhaps, a relatively extreme example of such small-group work designed to enhance learning.

The NRC addresses cooperative learning in a number of contexts. Noting that learning “is influenced in fundamental ways by the context in which it takes place,” the report recommends a learning community:

Teachers must attend to . . . helping students organize their work in ways that promote the kind of intellectual camaraderie and the attitudes toward learning that build a sense of community. In such a community, students might help one another solve problems by building on each other’s knowledge Both cooperation in problem solving and argumentation among students in such an intellectual community enhance cognitive development.²³

The NRC further says:

[T]he ultimate goal of schooling is to help students transfer what they have learned in school to everyday settings of home, community, and workplace. . . . [A]n important strategy for enhancing transfer from schools to other settings may be to better understand the nonschool environments in which students must function. . . .

The question of how people function in a number of practical settings has been examined by many scientists, including cognitive anthropologists, sociologists, and psychologists. One major contrast between everyday settings and school environments is that the latter place much more emphasis on individual work than most other environments. A study of navigation on U.S. ships found that no individual can pilot the ship alone; people must work collaboratively and share their expertise. More recent studies of collaboration confirm its importance.²⁴

Many students’ reactions to the collaborative exam, in both written evaluations and personal conversations, have been consistent with the NRC’s report. In that regard, the experiment with collaboration on exams has exceeded my expectations. Many students apparently come away from the experience thinking that they learned much more because of the collaboration.

The NRC notes that “teachers need to pay attention to the incomplete understandings, the false beliefs, and the naive renditions of concepts that learners bring with them to a given subject.”²⁵ I assume all law teachers are familiar with incomplete understandings, false beliefs, and naive renditions of concepts. As noted below, many students have indicated that some of their incomplete or false understandings and naive renditions of concepts have been corrected in the course of collaborating on the examination.

22. *Id.* at 67.

23. *Id.* at 25 (references omitted).

24. *Id.* at 73–74 (references omitted).

25. *Id.* at 10.

The NRC suggests that “effectively designed learning environments” must be not only “learner centered” and “knowledge centered” but also “assessment centered.”²⁶ The report notes that feedback is extremely important to effective teaching. The traditional law school model, in which students receive very little useful feedback during a course, is not necessarily unusual:

Most teacher feedback—grades on tests, papers, worksheets, homework, and on report cards—represent summative assessments that are intended to measure the results of learning. . . . Feedback is most valuable when students have the opportunity to use it to revise their thinking as they are working on a unit or project. Opportunities to work collaboratively in groups can also increase the quality of the feedback available to students, although many students must be helped to learn how to work collaboratively.²⁷

Student Reaction to and Evaluation of Collaboration

I have had conversations with many students about my experiments with examinations. Their reactions have generally been thoughtful and reasonable. One first-year student was even concerned that distributing information about facts and law several weeks before a three-hour examination was inappropriate: he thought that would help students who hadn't been studying during the semester catch up with the more diligent students just before the exam. I assured him that diligence throughout the semester was generally necessary to succeed in law school, and after the exam he agreed with me.

The feedback I have received from students suggests that collaborative models of examination have significant potential as a means of helping students learn. In addition to informal conversations, I have asked students to participate in several surveys concerning their opinions about the collaborative exam process. The responses have been quite consistent from one year to the next. I summarize here some of the results of such surveys with responses received from 136 of the 156 first-year students who took the Property exam at the end of the fall semester of 2000 and from 52 of the 77 students who took the Procedure exam at the end of the fall semester of 2003.

One set of questions in the survey of Property students asked for their reactions at different points in time to the opportunity to collaborate with other students on the examination—before the exam, during the two- or three-day examination period, and after the exam was over. Opinion varies some, but the initial reaction was nearly identical to the reaction after the exam. Interestingly, the number of students who were opposed or very opposed to the method was at its highest after grades were received (16 students, as opposed to 11 students after the exam but before grades were received). Because Procedure is only a one-semester course, I did not learn the opinions of the 2003 Procedure students after they had received their grades.

26. *Id.* at 139.

27. *Id.* at 140–41. My experience is certainly consistent with the conclusion that many students need help in learning how to work collaboratively. In my first-year courses I have given a midterm practice exam that uses the same format as the final examination. Collaboration is permitted. Practice seems to help and in some cases, after collaborating on the midterm exam, students have decided against collaboration on the final exam.

In each course, the students were asked to fit their reaction into one of five descriptions. The following opinions were expressed by the 136 Property students and 52 procedure students in response to a question like the following: Which of the following terms best describes your opinion about the fact that collaboration among students was permitted on part I of the Procedure I examination?

Property I, Fall 2000		Procedure I, Fall 2003	
Very enthusiastic	34%	Very enthusiastic	29%
Enthusiastic	36%	Enthusiastic	38%
Undecided or neutral	18%	Undecided or neutral	19%
Opposed	7%	Opposed	10%
Very opposed	4%	Very opposed	4%

In both courses, I also asked the students to rate the exam “as a teaching/learning device” and “as a means of evaluating students’ work.” The responses were as follows.

Property I, Fall 2000		Procedure I, Fall 2003	
<i>Learning device</i>		<i>Learning device</i>	
Excellent	34%	Excellent	31%
Good	46%	Good	46%
Neutral	12%	Neutral	15%
Poor	5%	Poor	6%
Very poor	2%	Very poor	2%
<i>Means of evaluation</i>		<i>Means of evaluation</i>	
Excellent	17%	Excellent	10%
Good	40%	Good	36%
Neutral	20%	Neutral	36%
Poor	13%	Poor	12%
Very poor	10%	Very poor	6%

I am somewhat surprised by these results. I expected criticism from a larger number of students, and I did not anticipate as much enthusiasm as I have received in both written and oral communication. Combining the two courses summarized above, for example, and reducing the reaction to positive (“excellent” and “good”) versus negative (“poor” and “very poor”), student opinion as to collaboration as a learning device is remarkable. Of the students who expressed positive or negative opinions on the process in those two exams, 91 percent concluded that the collaborative exam was a good or excellent learning device.

A Tentative Conclusion About the Experiment

Although I occasionally have mixed feelings about these experiments in collaborative examination, there seems to be more promise there than I thought I would find, and I am more enthusiastic about the process than I expected to be nine years ago. An image that keeps coming back to me is of three male students in my first-semester Property course a few years ago, probably in 2000–01. They seemed to be diligent students, and during our

discussions, over a period of several weeks, they vigorously contended with me over the prospect of a collaborative exam. They expressed principled and justifiable reasons for their opposition to it.

Notwithstanding the arguments of that trio, I decided to give a collaborative exam at the end of that semester. A month or two after the exam, one of the three students came to my office. He exuberantly told me he had changed his mind about collaboration on exams. He explained that he and his two friends who had argued against collaboration had decided to work together. They had spent the exam period in a cabin somewhere in the Texas hill country, discussing the issues and formulating their responses to the examination problems. He told me that in all his years of formal education, the time spent in that cabin was the best educational experience he had ever had. I have heard less dramatic but similar stories from many students who have collaborated with others on my exams.

There have certainly been negative remarks as well, but not as many as I expected. And even some of those whose collaborations did not produce high grades have commented positively on the learning experience. In some cases, unfortunately, I am quite sure that bad things have happened because of some of the dynamics of the group work.²⁸ From conversations with students, I believe that even in those situations students have sometimes garnered valuable learning experiences. For the time being, I continue this experiment because I am persuaded that collaboration on exams has promise as an effective tool in the education of law students, without necessarily compromising the teacher's responsibility to evaluate performance.

Comments from Students

In addition to specific questions, my surveys have invited students to share with me their general thoughts and opinions about the examination process. They seem to have appreciated that opportunity, and reading their written comments has been enlightening. As a conclusion to this paper, I present comments made by Property students during the spring semester of 2000–01 about their experience with the fall term final examination. Almost half of the students enrolled in the two sections wrote comments. Although I have omitted some repetitious comments and paraphrased others, the substance of virtually all student remarks is included here. I have grouped the comments as either “positive” or “negative and/or cautionary.”

Negative and/or Cautionary Comments

There was no examination of individual performance.

The only unfairness is students who “coast” through, relying on others.

This exam format discourages many from studying during the semester.

28. “Bad things” can certainly be educational and in that sense somewhat worthwhile, but further study is warranted on the social effects of collaboration in an examination context.

If you chose your partners poorly, you were stuck.

The exam was very time-consuming and almost required a group-work setting.

Students divided the work and therefore weren't tested on all of the exam.

Opportunity to collaborate was good practice but my knowledge was not reflected in my grade.

Some students may not give a full effort but rather rely on fellow students who are more serious or aggressive.

Those who worked in groups were more likely to make higher grades.

It was rumored that the larger your group, the lower your grade.

The grade received doesn't reflect, in the slightest, the endless hours of study and hard work I put into the course material. The professor should seek more objective means of evaluating a student's true knowledge of the material.

My enthusiasm for collaboration dwindled when my fellow collaborators were crying and yelling.

Positive Comments

The exam allowed us to bounce ideas off of one another.

We were able to explore potential answers to their fullest.

We learned more about Property and collaboration.

The exam taught us how to work together in negotiation situations.

We received feedback and reinforcement from one another.

This was a refreshing format for an exam.

The exam forced reevaluation of my understanding and interpretation and required me to defend my position or compromise and caused me to look at the analysis on a deeper level. There was a psychological benefit because I felt a little less pressure.

The process allowed opportunity for discussion, during which we realized that law is not black and white.

This exam really helped me learn the material.

This exam was more similar to "real-life" situations.

Working through problems, issue by issue, forced us to a level of specificity and a depth of analysis that we wouldn't have gotten to without the seriousness that creating exam answers moved us to.

Collaboration is good because it forces students to work together and exposes them to other ideas and perspectives.

The exam gave me a better understanding of property concepts. Things I thought were clear became unclear as others questioned the concept.

This exam separated correctly the students who worked hard all semester and those who did not take the class seriously.

The exam helped me clarify issues that were “blurry” to me.

The exchange of ideas and discussion of thoughts helped us understand the concepts.

Collaboration allowed me to analyze and ponder the issues. This was an excellent opportunity to apply the law.

My partner and I were able to teach each other. The test also helped me put things together and see the “big picture.”

Although I didn’t collaborate, I think the opportunity to do so was a good thing.

We learned by sharing and exchanging ideas.

The exam was more like practice, and collaboration helped me understand and retain the material.

“Traditional” exams test speed, not mastery of the material which the collaborative exam required.

I had time to really think about the subject and thus learned a lot.

Although skeptical at first, we learned a lot from each other through collaboration.

The process allowed a comprehensive review and helped me learn the material in more depth.