

Teaching Critical Thinking and Writing through Debates: An Experimental Evaluation

Author(s): Charles S. Green, III and Hadley G. Klug

Source: *Teaching Sociology*, Vol. 18, No. 4 (Oct., 1990), pp. 462-471

Published by: American Sociological Association

Stable URL: <http://www.jstor.org/stable/1317631>

Accessed: 30-09-2015 16:09 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Sociological Association is collaborating with JSTOR to digitize, preserve and extend access to *Teaching Sociology*.

<http://www.jstor.org>

TEACHING CRITICAL THINKING AND WRITING THROUGH DEBATES: AN EXPERIMENTAL EVALUATION*

CHARLES S. GREEN, III

University of Wisconsin-Whitewater

HADLEY G. KLUG

University of Wisconsin-Whitewater

Faced with the desire to maintain high standards of scholarship in a context of large introductory level classes and limited resources, the authors experimented with several active learning techniques designed to develop critical thinking and writing skills. The tentative conclusions drawn are that the debate format in the context of large classes is an effective way to modify students' opinions on social issues and to teach critical thinking and writing skills. Also discussed are several ethical issues involved in teaching critical thinking.

BACKGROUND

Probably all faculty think that the teaching of critical thinking and writing is a "good thing," just as motherhood and apple pie are good things. The problem, of course, is that unless one has plenty of teaching assistants, the conventional term paper isn't feasible in mass classes of 50, 100, or more because the grading load is simply too high. Moreover, if one has ambitions of having students not only write papers but rewrite them several times during a semester, the grading task becomes virtually impossible. So, perhaps reluctantly, we give up the goals of teaching critical thinking and writing and fall back on lecturing and multiple choice tests. But what if our consciences bother us?

Our consciences bothered us, so for over ten years we have tinkered around in a variety

of ways to try to give our students at least some writing experience. Green, for example, gave his students the choice of doing short papers on selected and well-defined topics or of doing reviews of books from a limited list of suitable books. Students were also offered a "revision provision": if they submitted drafts of their short papers or book reviews by a deadline early in the semester, they were provided with constructive criticisms for use in revising their work. The strategy worked reasonably well in that both writing style and substance improved markedly from rough to final drafts. But the grading load was very demanding and after several semesters he got numb reading papers on the same small number of topics over and over again! Moreover, in large classes it is impossible to give students the opportunity to improve their writing through several revision and critique cycles.¹ In addition, the proliferation of papers on a limited number of topics semester after semester invites plagiarism.

Finally, it is not at all clear that papers and reviews stimulate much critical thinking. In fact, the nature of critical thinking was not clear in our minds when we were first tinkering around with our courses and there wasn't much literature available then on teaching critical thinking in sociology. We did share the notion that critical thinking involved such things as the ability to create logical arguments based on the "sociological imagination" and to support those arguments empirically. Therefore, we concluded that one way to overcome the shortcomings of short papers and book reviews was to have small groups of students collaborate in preparing library research papers. By collaborating, students would be-

* Charles S. Green III was awarded the Hans O. Mauksch Award for Distinguished Contributions to Undergraduate Sociology in 1989 by the ASA Section on Undergraduate Education. Award recipients are required to provide a talk at the following year's meetings. Green's talk before the 1990 meetings was based on this article. The article is a revised version of a paper presented by both authors to the 1990 annual meetings of the Midwest Sociological Society in Chicago.

The authors wish to acknowledge the help their colleagues rendered to this project by grading students' essays: Ronald Berger, Eugene Grigsby, Lanny Neider, Lawrence Neuman, and Richard Salem. We also wish to thank our colleague Patricia Searles as well as Dean Dorn and Barbara Sherman Heyl for providing useful comments during various stages of the project, William Rau and Jean C. Karlen for comments during the Midwest meetings, and the three reviewers of this version: Louis E. Anderson, Eugene M. Labovitz, and Norma J. Shepelak. This article was copy edited by Donna Perdue.

come more responsible for their own learning and would teach one another, improving their thinking and writing skills (see Billson 1986; Cohen, Lotan, and Leechor 1989; Rau and Heyl 1990). In addition we anticipated that since there would be a smaller number of papers to grade, we could afford the time to critique them and allow students to revise them. We also expected that students would acquire a sense of the vast array of empirical and theoretical literature that exists in sociology and acquire the skills for finding that literature in the library. Finally, we believed oral skills could be developed by having students present their results in class.

After using such projects for several years, we concluded that they did indeed have numerous advantages over short papers and book reviews. However, the presentations the groups made in class tended to be stupefyingly dull, for the students presented their ideas in much the same fashion that professors present papers at the meetings of professional associations! A far more serious problem was that there was no direct challenge to the ideas being presented and, therefore, no need for students to defend their own arguments and rebut those of others, both of which we thought should be crucial components in the development of critical thinking skills.

In order to overcome these difficulties we decided to try a debate format. Starting in 1985, students in Green's Introductory Sociology and Marriage and the Family classes were given the choice of five debate topics, with the expectation that every student would participate in a debate.² Students were assigned at random to the "pro" and "con" sides of their debate topics. All students were given a detailed handout containing directions for researching their topic, starter bibliographies, suggestions for coordinating their efforts in preparing for both the oral in-class debate and the written version of their arguments, and the standards that would be used in grading de-

¹ The writing-across-the-curriculum approach has promised that writing skills can be improved if students have frequent writing assignments and that little increase in grading loads need result from frequent writing assignments because not all assignments need to be graded rigorously. However, Day (1989, p. 458) in an experimental evaluation of the approach found that "... numerous writing assignments are not sufficient in themselves to produce better writing, but that rigorous grading of the entire assignment is the variable which produces improvement."

bates and papers. As had been done with the earlier projects, rough drafts of the group papers were submitted relatively early in the semester and subjected to detailed critiques of logic, evidence, grammar, and composition. Students were then permitted to revise and resubmit these drafts by the end of the semester. Both the oral debate and the group paper received group grades to emphasize the importance of cooperation. However, in order to deter "free-riding," groups were empowered to eject any student who was not doing her or his share of the work.³

We expected to gain a number of advantages from this debate format. First, the added drama of debating a controversial topic was expected to generate more interest in the presentations and more post-presentation discussion than had been the case for the projects. Second, it was expected that students would not only develop research skills in the process of developing and finding evidence for their arguments but also skills in oral and written argumentation in the process of working out defenses for those arguments and rebuttals to opposition arguments. Third, it was expected that, as was the case with projects, students would learn more effectively when working cooperatively than when studying alone. Fourth, it was expected that the groups formed for the debates could be used for other class discussions, thus providing further opportunities for collaborative learning.

Early experience in using debates revealed that the first expectation was fully realized: debates were lively and elicited a great deal of post-debate discussion on the part of class members not debating. Moreover, it was clear

² In deciding upon debate topics careful consideration must be given to selecting topics which are current and controversial, but not so emotionally engaging that students are sharply polarized in their initial opinions, and to selecting topics for which a substantial and reasonably high quality body of literature exists for both sides. Students must be informed that grading of both oral debates and debate papers focuses on the quality of their reasoning and evidence, not on the side of the debate topic which they have been assigned to defend.

³ Relatively few students have been ejected by their teams, perhaps because the sanction has been so severe. Recently, on students' advice, a less severe sanction has been added. Now at the end of each semester students in each team are asked to evaluate each other's contributions to the team effort; if there is a consensus that a student's performance has been inferior, the instructor lowers that student's grade.

from both the debates and post-debate discussions that some crucial critical thinking skills were being used: students challenged each other on such matters as definition of terms, unstated value premises, the lack or misuse of evidence, reliance on "authorities" whose expertise was questionable, the pertinence, reliability, and validity of evidence, and the adequacy of logic. It was not, however, clear whether those skills had been learned in preparing for debates or learned previously in other classes such as speech, a subject which our students are required to take as part of their general education program. Finally, a number of unexpected findings turned up.

One surprise was the extent of the impact of the debates on multiple choice exam scores. Typically four such exams were given, with the third exam given after the first two debates and the fourth exam given after the last three debates. In all classes, the scores (percentage of items correct) on the last two exams were significantly higher and had lower standard deviations than did the first two exams.⁴ In contrast to what we had expected, this effect was not limited to the groups debating before each exam but rather was characteristic of most students in each class. Thus, debating and even hearing and later discussing those debates reinforced students' learning of the text material paralleling the content of the debates. A second surprise was that the student evaluations of the course improved significantly and the standard deviations of evaluation scores declined significantly compared to the scores from Green's previous courses, taught with lectures combined with group projects or with short papers and book reviews, and Klug's courses, taught in a conventional lecture format.⁵ When students in Green's sections were

asked to write on the student course evaluation instrument "What were the best and worst features of this course?" about 40 percent chose to do so. Of the latter, 50 percent wrote specific and favorable comments about debates while just 11 percent wrote unfavorable comments. No other feature of the course was mentioned as favorably.

Before analyzing data on exams and student evaluations, thereby becoming fully aware of the first two surprises, Green decided that perhaps he had been too lenient in assigning letter grades to multiple choice exam scores. So he raised the standards, hoping thereby to achieve a Pygmalion effect. In the five course sections taught since these standards were raised, exam scores (percentage correct, not letter grades) for the first two exams have gone up significantly. However, the higher expectations had no impact on scores for exams three and four. Apparently the impact of debating overwhelmed the effect of higher standards. Incidentally, although the higher standards resulted in sharply reduced proportions of As given as final course grades, this had no noticeable effects on students' evaluations of the course. In short, the major impact on those evaluations has come from organizing classes into small groups for debates and other discussion and class exercises.

In order to explore the potential of the debate format further, we designed a series of experiments to determine the impact of debates on students' opinions about social issues and on changes in thinking and writing skills.

CRITICAL THINKING AND WRITING: EXPERIMENTS

We have already noted the rough-and-ready definition of critical thinking that we used in

⁴ We regressed the grades from all four multiple choice tests against time. Before debates were introduced the slopes of all four lines were essentially zero, intercepts of all four lines were nearly identical at about 60 percent correct answers, and Pearson's R equalled roughly .70 ($p < .001$) for all four lines. After debates were introduced slopes were again found to be near zero for all four lines. Intercepts for tests one and two remained at about 60 percent but those for tests three and four increased to about 70 percent correct answers. Rs remained at about .70 for tests one and two but increased to about .80 ($p < .001$) for tests three and four. Attempts to fit a single line to all test scores were not successful: neither linear nor several curvilinear regressions produced Rs greater than .30 and none were significant. Grades in Klug's courses remained constant, with regression lines for his tests comparable to those for Green's tests one and two.

⁵ Students taking courses in our department are asked to rate instructors on fourteen items, the last of which requires an overall evaluation: "Taking everything into account, how would you rate this instructor?" The rating scale varies from a high of one to a low of five. The mean scores on this item were regressed against time. Before debates were introduced, the regression intercept (i.e., mean evaluation score) was 1.9 with a slope of zero and $R = .62$ ($p < .01$). After debates were introduced, the regression intercept was 1.6 with a slope of zero and $R = .79$ ($p < .005$). Attempts to fit linear and curvilinear lines through the mean scores of all classes yielded non-significant Rs of less than .40. Student evaluations in Klug's classes were comparable to those in Green's classes before debates were introduced.

our early tinkering with our courses. But a review of the literature suggests that a more precise and certainly more sophisticated definition is in order. For example, Dean Dorn (1987) defines critical thinking as “. . . the process of reasonably deciding what to do and/or believe. This means individuals should be able not only to assess their own and others' arguments but also construct good arguments.”⁶ Dorn points out that most experts believe that critical thinking involves acquiring a set of micro and macro level skills and dispositions:

1. Micro level intellectual abilities and skills. Examples: ability to clarify issues, ability to identify fallacies in an argument, ability to find value assumptions embedded in an argument, ability to know when statistics are misleading or absent, ability to judge whether a statement follows from a premise.
2. Macro level dispositions. Examples: an interest in seeking reasons for what to do or believe, skepticism about others' beliefs, a predisposition to ask for evidence, a predisposition to be creative, to think of counter-arguments and examples, to be sensitive to one's own biases and values.
3. Macro level values. Examples: a commitment to open-mindedness and fairness, empathy for others' positions, openness to self-criticism, appreciation of the value of looking at events from multiple points of view.

With a better idea of what critical thinking is, can we specify what are the best methods for conveying these skills, dispositions, and values? The conventional wisdom in the academic world is that lecturing and multiple choice exams are useful only for imparting the recall of facts, while Socratic dialogue, essay exams, and term papers are best for teaching higher order cognitive skills. But this dichotomy is far too simple. Socratic dialogues can be used to intimidate students, teaching them only what to think, not how to think, whereas

⁶ Dorn's approach tends to focus on the development of skills of formal argumentation as developed by philosophers and rhetoricians. Other approaches include those of cognitive psychology, epistemology, and Marxist critical theory (see Nelson 1989).

lectures can be used to challenge students to question the beliefs and values they take for granted. Similarly, objective tests can be constructed to measure higher order cognitive skills (see Howery 1987); whereas, essay exams may tap mere recall of facts, figures, and ideas. As Baker (1981) concludes “. . . critical thinking is not associated with any single pattern of teaching and testing.”

Mayer (1986) argues that “. . . critical thinking cannot be accomplished unless the goal is specifically built into course objectives and methods.” Teaching critical thinking requires what Goldsmid and Wilson (1980) call “benign disruption”: a set of techniques by which the taken-for-granted is challenged. Moreover, such techniques cannot be effective if addressed solely to intellectual concerns. Teaching techniques must instead link reason and emotion (see Baker 1981; Nelson 1989). In establishing this link “The key . . . lies in creating conditions for participation rather than passivity, and in providing opportunities for emotional engagement with the materials” (Mayer 1986). Debates appear to fit all the criteria mentioned above (see also Broderick 1982 and Huryn 1986). But do they actually work? Can it be shown that changes take place in students during a semester that can be attributed in whole or part to the debate experience?

EXPERIMENT 1: HOW DOES DEBATING AFFECT OPINIONS ON ISSUES?

We started off with a rather naive hypothesis: that debating would give students greater intellectual flexibility (see Kohn 1971; Kohn and Schooler 1983). Specifically, we predicted that those who had debated and had listened to others debate various issues pertaining to marriage and the family (Green's students) would become less extreme in their opinions than students who had also taken a marriage and family class but under more conventional teaching methods (Klug's students). To test these expectations we administered an opinion survey to two of Klug's and three of Green's classes at both the beginning and the end of the semester. To see if reactive effects were present (for example if taking the survey early had an impact on later opinions), we also administered the survey only at the end of two other classes (one of Green's, one of Klug's).

Table 1. Comparison of Opinion Change Among Non-Debaters in Experimental and Control Classes, by Debate Topic

Debate Topic:	Experimental Classes:	Control Classes	
		1	2
1. Enough love maintains marriage	1		
	2	*	*
	3	*	*
2. Singlehood preferable	1		
	2		
	3		
3. Trained parents produce model children	1	*	*
	2	*	*
	3	*	*
4. Women's liberation gone too far	1		
	2	*	*
	3	*	*
5. Family has no future	1		*
	2	*	*
	3	*	*

* The proportion of students changing opinions in the experimental class is significantly higher ($p < .05$, one-tail) than the proportion of students changing opinions in the control class; in no comparison is the proportion changing opinions in the control class significantly higher.

Binomial tests:

- For all 30 comparisons, 19 as predicted, $p = .100$ one-tail
- Excluding Debate 2: of 24 comparisons, 21 as predicted, $p = .001$ one-tail
- Excluding Debate 2 and Experimental Class 1: of 16 comparisons, all as predicted, $p = .000$ one-tail

Thus we had a Solomon four-group research design (Sellitz, Wrightsman, Cook 1976, pp. 140-48). Students were asked to express their opinions by indicating on a five item Likert scale how strongly they agreed or disagreed with each of the following debate topic statements:

1. "If two people love each other enough, they can maintain their marriage until death do them part."
2. "Singlehood is much preferable to marriage."
3. "If parents are trained in modern methods of child-rearing, they should have no problem producing model children."
4. "The women's liberation movement has gone too far."
5. "The family has no future."

What were the results of this experiment? Contrary to our initial naive hypothesis, very few opinions among either controls or non-debating experimentals shifted toward neutral ("I can see merit to both sides of this"), the position we first thought indicative of intellectual flexibility. In fact, substantial proportions of students were neutral to begin with, especially with respect to topics 2, 3, and 4. On the other hand, roughly 60 percent of students in the experimental classes and 40 percent of students in control classes changed their opin-

ions. Among those whose opinion did change, the change typically involved a shift of just one category on the Likert scale (for example from strongly agree to agree). Further, it is clear from the results in Table 1 that the experimental "treatment" did not lead to a shift of opinion for debate topic 2 nor within experimental class 1. Interestingly, it was only in the latter class that—despite directions to the contrary—students in several debates issued disclaimers to the effect that "I really don't believe this, I'm just defending this position because I have to."

When we examined opinion change among debaters we found something quite different. Those who debated a side ("pro" or "con") *opposite* to their initial opinion changed their opinion sharply, typically switching their opinion to the side upon which they debated. Those who were initially neutral were even more likely to switch their opinions to coincide with the position debated. In contrast, those who debated a side *coinciding* with their initial opinion either did not change their opinion or strengthened it (see Table 2). These differing patterns of opinion shift reflect differences in the extent to which students experienced cognitive dissonance. Furthermore, those who experienced dissonance chose one of three possible modes of resolving it (see Brehm and Cohen 1962; Michener, DeLamater, Schwartz 1986, pp.184-86). When asked to defend in public views with which they disagreed, stu-

Table 2. Change of Opinions Among Debaters in Experimental Classes, by Section and by Correspondence Between Initial Opinion and Position Debated

Opinion Change	Initial Opinion Relative to Position Debated, by Section								
	Section 1 (n = 55)			Section 2 (n = 50)			Section 3 (n = 43)		
	Opposite	Neutral	Same	Opposite	Neutral	Same	Opposite	Neutral	Same
Switched*	3	14	1	8	13	1	6	11	0
Shifted to neutral	4	0	1	3	0	0	7	0	0
Weakened**	4	0	0	2	0	4	0	0	0
No change	10	3	10	3	2	7	1	4	10
Strengthened***	0	0	5	1	0	6	0	0	4
Sub-totals	21	17	17	17	15	18	14	15	14
Gamma				.25			.51		
Chi square****				16.97			14.48		
P <				.001			.001		

* From agreement to disagreement or vice versa; or from neutral to position debated.

** From strongly agree to agree or from strongly disagree to disagree.

*** From agree to strongly agree or from disagree to strongly disagree; or from neutral to position opposite that debated.

**** For purposes of calculating Chi square, sub-tables were collapsed by dichotomizing the opinion change variable between "weakened" and "no change."

dents faced three choices: 1) they could change their opinion to conform to the view defended, which most students did; 2) they could keep their initial opinion, claiming publicly that they "don't really" feel or think "that way"—which several in experimental section 1 did, despite instructions to the contrary; or 3) they could keep their initial opinion, excusing to themselves and possibly to their significant others the apparent inconsistency between opinion and behavior by saying "I was forced to defend that side." Several students confided to Green that this is just what they did. Incidentally, we were able to find no reactive effects of taking the opinion survey at the beginning of a class: among both the experimental (debate) and control classes, the distribution of opinions at the end of a semester had no relation to whether or not a class had taken a survey at the beginning of that class.

EXPERIMENT 2: HOW DOES DEBATING AFFECT THINKING AND WRITING SKILLS?

As we noted earlier, debates have a number of advantages over other teaching techniques, even including those techniques which place some emphasis on writing (such as research projects, book reviews, short papers). But, so far we have not demonstrated whether or not debating has any impact on thinking or writing. The best demonstration would involve a careful external evaluation of both the group pa-

pers produced and the tape recordings of selected oral in class debates. However, this would be quite time consuming, and so far we have found no one willing to do it. Furthermore, such a strategy would leave unanswered the more important question of whether or not students acquired skills which would *transfer to new situations*. In fact, the literature is quite pessimistic about such a possibility (Baker 1981, p. 341) for it appears that the teaching of critical thinking and writing succeeds only when it is taught across the curriculum (Browne and Litwin 1987).

In order to examine the transferability of skills, the authors had every student in their spring 1989 Marriage and Family classes (one section each) write an essay at the beginning of the semester defending their position on one of the five debate topics outlined earlier. (Green's students were asked to pick one of the four topics they were not assigned to debate.) We also had them write on those same topics at the end of class. Thus, we had "before" and "after" essays from each student which we planned to have graded blindly by other colleagues in our department.

In order to keep the grading load for these faculty within some reasonable bounds, we each picked a 25 percent sample of our students at random: 11 from Green's class, 10 from Klug's. Thus from these 21 students there were 42 essays to be evaluated (21 "before" and 21 "after") by each of two faculty members. We assigned essays to faculty at random so that each would grade essays from

Table 3. Reliability of Faculty Evaluations of the Quality of Thinking and Writing Displayed in Students' Essays

	Thinking	Writing
1. Differences in absolute scores between the two graders of each essay:*		
a. Mean difference	12.64	7.02
b. Median difference	10.00	5.00
2. Rank order correlation (Kendall's tau) between each grader's scores and those assigned by other graders of those same essays: **		
Grader 1	.804	.841
Grader 2	.921	1.000
Grader 3	.724	.871
Grader 4	.889	.946
Grader 5	.902	.935

* Scoring was based on a 100 point scale; see Appendix for scoring directions given faculty.

** For all taus, $p < .0000$ one-tail.

both experimental and control classes and each would grade "before" and "after" essays. Each faculty member was provided with a set of guidelines for grading the quality of both students' writing and critical thinking (these guidelines may be found in the Appendix).

Before turning to the results of the faculty grading efforts, it is important to discuss whether or not the grading of the essays was reliable; for example, do faculty agree on what is good critical thinking and good writing? Note the results reported in Table 3. Part 1 of that Table suggests that faculty have quite different standards in assigning absolute numerical scores, especially to the quality of thinking. They are almost 13 points apart, the equivalent of a grade and a half or the difference between an A and a B-. On the other hand, Part 2 of the Table suggests a very high degree of reliability, for faculty do agree on the rank ordering of the quality of students' essays. For example, the rank order correlation between the scores assigned by Grader 1 for writing and the scores assigned by the other graders of 1's set of essays is .804. With some assurance that the grading is reliable, we turn to the actual research questions of interest.

The results are summarized in Table 4. Part 1 of that Table shows that the experimental class improved the quality of its thinking by a full ten points, a highly significant change, whereas the control class improved barely four points, which is not significant. It is important to note that the scores reported in Table 4 were

Table 4. Faculty Evaluations of the Quality of Thinking and Writing in Beginning and Final Student Essays, by Experimental Condition

	Experimental	Control
1. The Quality of Thinking		
Mean score of two graders:		
Beginning essay	65.6	63.6
Final essay	75.9	67.1
Mean change in scores:	10.4	3.5
	$p < .0005$ one-tail	p NS one-tail
2. The Quality of Writing		
Mean score of two graders:		
Beginning essay	71.6	68.7
Final essay	77.3	72.8
Mean change in scores:	5.6	4.1
	$p < .0005$ one-tail	p NS one-tail

obtained by averaging for each student the scores assigned by each faculty grader. The statistical significance of the results is not changed if, instead of averaging the scores of the two graders, the scores of just the first or just the second graders are analyzed, or if the scores of the graders assigning the worst grades are analyzed, or if the scores of the graders assigning the best grades are analyzed, or if the scores of graders picked at random are analyzed. Therefore it doesn't seem to matter which set of scores are chosen for analysis.

Part 2 of Table 4 shows that students in the experimental class improved their writing by almost six points, a statistically significant change; whereas, students in the control section improved by just four points, a difference which is not a statistically significant result. Again, tinkering around with scoring methods does not change the significance of the overall result.

In summary, the results look quite good: debates do seem to work in that they produce improvements in both the quality of thinking and the quality of writing. However, a more cautious, even pessimistic, interpretation can be made of the results in Table 4. Even in the experimental class, final essays were judged to be about C quality work with respect to the quality of thinking (a score of 75.9) and about the same quality with respect to writing (a score of 77.3). We would be much more encouraged by B quality work.

DISCUSSION

We have shown in this paper that debates can be effective in improving students' perfor-

mance on multiple choice exams covering material related to the debates, in raising students' participation in and enthusiasm for class discussions, and in raising students' evaluations of instructors. Furthermore, there is evidence for modest gains in writing and critical thinking skills and evidence that these skills are transferable. However, the study does need to be hedged with several qualifications.

From a methodological perspective, our experiments are not true ones: students were not assigned at random to experimental and control classes. Therefore, there is the risk that the results are due merely to differences in composition between classes. We have tried to rule out the effects of several potential sources of bias in the composition of classes. For example, the distributions of students by year in college and by major field are the same between our classes. Moreover, initial opinions about the five debate topics are similar. But we cannot rule out more subtle differences: might students have chosen Green and Klug for different reasons which might have affected their performance in class? In fact, there is a hint that such is the case, for a somewhat higher (but not quite statistically significant) percentage of Green's classes are composed of females. Differences in the way we conduct our classes also may have had an impact on one or more aspects of students' performance. One such difference, our textbooks, can probably be ruled out: Green used several different textbooks before introducing debates with no noticeable effects on student evaluations or on grades on multiple choice tests. In addition, we have no way of knowing to what extent opinion change may have been influenced by students knowing or guessing that Green expected such change. Nor do we know how those debaters who faced a dissonance arousing situation were influenced by the presence of friends in the audience for their debate.

Our conclusions with respect to improvements in thinking and writing skills are based on just 21 students. If we had more faculty on our staff or could have induced our present faculty to grade far more essays, we would have greater confidence in our results. Thus, the study cries out for replication.

This study provides no evidence concerning the permanence of either the changes in opinions induced or of the improvements in writing and critical thinking skills. We also

do not know whether the latter skills will transfer to other sociology courses, let alone to course work in other disciplines. It should be obvious that follow-up studies are needed.

Debates seem to be like other methods of stimulating critical thinking in that they 1) encounter students with views contrary to their own and, in so doing, 2) induce students either to change those views or learn to defend their own views with better logic and more substantial evidence. We wish to emphasize, however, that debates are not the only way to get students to think more critically, and they probably are not the best way—of course the “best way” has yet to be found. Nevertheless, debates are especially advantageous in large classes. By subdividing a large class into smaller groups, instructors can encourage cooperative learning, not only for the debates per se but also for other class discussions or exercises and for preparing for exams. A final advantage of debates is that instead of grading (or giving up on grading) 50 or more papers, the instructor has to grade just one paper from each debate team.⁷

We wish to close this article by noting two ethical issues involved in teaching critical thinking. First, are we engaged in a form of brainwashing when we get students' opinions to change sharply, especially under the condition in which they must defend in public views to which they are initially opposed? The second issue, one raised by a reviewer, concerns whether our teaching of critical thinking merely creates amoral monsters: persons adept at criticism but lacking both a sense of values and a commitment to acting on those values. Thus, could it be that critical thinking and commitment are actually incommensurate,

⁷ The writing of formal, stylistically correct papers may not be essential in the teaching of critical thinking. In teaching during summer sessions often just three weeks in length, Green has found that it is almost impossible for students to write up their debate papers; consequently, he no longer requires papers during the summer. However, adequate preparation for the oral debates requires considerable notetaking, the sharing of those notes with one's debate team members, and the reorganization of data and potential points of argument from those notes into a logically interrelated set of arguments. Each debate team typically has a well developed outline and arranges to have each member present certain arguments in a specified sequence. The better prepared teams will also have members assigned in advance to rebut their opposition. Therefore, considerable writing (albeit informal) does take place, together with much discussion of how to organize arguments and mobilize evidence for best effect.

even mutually destructive? Or is it possible that students can be taught to act as responsible citizens while fully and freely acknowledging the weaknesses in the basis for that action (see Paul 1984)?

- value assumptions left unstated and unrecognized.
- cites the opinions of persons whose expertise is dubious or undocumented.

APPENDIX

DIRECTIONS FOR FACULTY GRADING THE ESSAYS

1. Please assign two grades to each essay, one for the quality of writing and one for the quality of argumentation or critical thinking. Guidelines for each of these are provided below.

a. Use a numerical scale for each grade:

Numerical Scale:	Rough Equivalency:
91-100	A or excellent
81-90	B or above average
71-80	C or average
61-70	D or below average
60 and less	F or well below average

b. NOTE—you do not have to provide any written comments, critique, or rationale for your grade. In fact, you should not write anything at all on the blue books.

2. Some guidelines to use in grading the essays:

a. Quality of writing

- i. Is the essay well organized and cohesive—are arguments clearly stated and is an outline clearly discernible, albeit implicit?
- ii. Are the sentences complete or are they mere fragments? Are sentences linked together into meaningful paragraphs? Are there effective transitions between paragraphs?
- iii. Is grammatical usage proper (e.g., punctuation, use of tense, avoidance of split infinitives)?
- iv. Are word choices suitable? Are words spelled correctly?

b. Quality of argumentation or of "critical thinking."

- i. Are sociological concepts used to illuminate and analyze the issue? Does the writer indicate an awareness of historical, cultural, and social structural (e.g., class, gender, age, race) contexts?
- ii. How adequate is the logic of the argument? Some common logical goofs to look for:
 - uses evidence selectively, or uses out of date evidence, or uses examples of dubious pertinence.
 - generalizes from personal experience or uses ad hominem arguments.
 - fails to define key terms or uses circular definitions.
 - insensitive to weaknesses and contradictions in own logic and/or evidence.
 - appeals to numbers or prevailing opinion (e.g., "we all know that . . .," "since most people think that . . . it follows that . . ."—the bandwagon rationale!)
 - not sensitive to questions of reliability and validity of evidence.
 - oversimplifies or misrepresents opponents' arguments—the straw person fallacy.

REFERENCES

- Baker, P.J. 1981. "Learning Sociology and Assessing Critical Thinking." *Teaching Sociology* 8:325-363.
- Billson, J.M. 1986. "The College Classroom as a Small Group: Some Implications for Teaching and Learning." *Teaching Sociology* 14:143-151.
- Brehm, J.W., and A.R. Cohen. 1962. *Explorations in Cognitive Dissonance*. New York: Wiley.
- Broderick, J.J. 1982. "Using Debate to Teach Sociology." *ASA Teaching Newsletter* 7:8-9.
- Browne, M.N., and J. L. Litwin. 1987. "Critical Thinking in the Sociology Classroom: Facilitating Movement from Vague Objective to Explicit Achievement." *Teaching Sociology* 15:384-391.
- Cohen, E.G., R.A. Lotan, and C. Leechor. 1989. "Can Classrooms Learn?" *Sociology of Education* 62:75-94.
- Day, S. 1989. "Producing Better Writers in Sociology Classes: A Test of the Writing-Across-the-Curriculum Approach." *Teaching Sociology* 17:458-464.
- Dorn, D.S. 1987. "Critical Thinking." Presentation at the Spring Teaching Seminar, University of Wisconsin-Whitewater.
- Goldsmid, C.A., and E.K. Wilson. 1980. *Passing on Sociology*. Belmont, CA: Wadsworth.
- Howery, C.B. 1987. "Measuring Up to Writing Good Multiple Choice Tests." *ASA Footnotes* 15 (February):9.
- Huryn, J.S. 1986. "Debating as a Teaching Technique." *Teaching Sociology* 14:266-269.
- Kohn, M. M. 1971. "Bureaucratic Man: A Portrait and An Interpretation." *American Sociological Review* 36:461-474.
- _____, and C. Schooler. 1983. *Work and Personality*. New York: Aldine.
- Mayer, J. 1986. "Teaching Critical Awareness in an Introductory Course." *Teaching Sociology* 14:249-256.
- Michener, H.A., J.D. DeLamater, and S.H. Schwartz. 1986. *Social Psychology*. New York: Harcourt Brace.
- Nelson, L. 1989. "Critical Thinking About Critical Thinking." Undergraduate Teaching Improvement Council, University of Wisconsin System. *Teaching Forum* 11(April):1-3.
- Paul, R.W. "Critical Thinking: Fundamental to Education for a Free Society." *Educational Leadership* 42:4-14.
- Rau, W., and B.S. Heyl. 1990. "Humanizing the College Classroom: Collaborative Learning and Social Organization Among Students." *Teaching Sociology* 18:141-155.
- Selltiz, C., L.S. Wrightsman, and S.W. Cook. 1976. *Research Methods in Social Relations*. 3rd ed. New York: Holt, Rinehart.

Charles S. Green III is Professor of Sociology at the University of Wisconsin-Whitewater. His interests include marriage and the family, critical thinking, and applied sociology. Address correspondence to Charles S. Green III, Department of Sociology, University of Wisconsin-Whitewater, Whitewater, WI 53190.

Hadley S. Klug is Associate Professor of Sociology at the University of Wisconsin-Whitewater. He has interests in marriage and the family, rural sociology, and intentional communal groups. He is doing research on changes in the curricular offerings of those communal groups known as "graduate departments of sociology."

For Faculty and Professionals with Authority to Recommend Journals for Their Library . . .

Would you like your library to subscribe to this journal? If you have the authority to recommend journals to your library, use the form below to order a complimentary sample copy for your librarian's review.

Yes, please send a sample copy of:

_____ (journal name)

I will submit the copy to our library for subscription consideration. The name of my institution and library are:

Name _____

Institution _____

Address _____

City _____ State _____ Zip _____

**Return form to: Publications Department, American Sociological Association,
1722 N Street, N.W., Washington, DC 20036**