Previous analysis of rights has been inextricably tied to ideology (e.g., Tushnet 1984). For instance, it is often argued that the assertion of rights is a highly individualistic, even capitalistic position (e.g., Lynd 1984; Gabel 1984). Milner (1989, p. 638), summarizing this position, asserts: "An emphasis on rights encourages a politics based on selfish individualism and discourages the communitarian and egalitarian values that the rights advocates often want to pursue." From this perspective, those more strongly committed to right-wing ideology should hold a broader understanding of individual rights.

On the other hand, there is good reason to expect that those on the political left in Europe hold a more expansive conceptualization of rights, especially those who are not on the extreme left. Conservatives are more likely to value social order over individual liberty, and are more likely to equate extreme rights consciousness with licentiousness and lack of individual discipline. Thus, we hypothesize that there is a negative relationship between right-wing ideological identification and rights consciousness.

Those who are better educated are also expected to subscribe to a broader definition of individual rights. Rights consciousness must be acquired—people do not naturally assert rights; they must be taught such assertiveness.[7] Formal educational systems are institutions that might contribute to the learning of rights awareness (at least in fairly democratic societies), and we therefore hypothesize a positive correlation with years of education.

Our earlier research on public opinion in Western Europe has shown that opinion leaders differ from ordinary citizens in several important ways (Gibson and Duch 1991b), and, more generally, elites are typically more committed to democratic values than ordinary citizens. Just as we hypothesize that more formal education contributes to an expanded sense of rights, we suspect that those who are more informed and active in politics—opinion leaders—will have a broader sense of rights.

We further expect that those who are deeply religious are less likely to claim secular rights than those who are not religious. Traditionally, the role of religion has often been to encourage the acceptance of authority, and is thus not supportive of the development of a broad sense of rights entitlement. Though different religious sects may vary to some degree on this score, generally those who do not look forward to some "pie in the sky" are hypothesized to be more likely to claim rights in the temporal world.

At an even more exploratory level, we will also investigate the degree to which rights consciousness varies across social class, gender, and age groups in Europe. Our general hypothesis is that those who are more socially, economically, and politically marginal in society will have learned a less expansive definition of the rights to which they are entitled.

Table 2 reports the relationships between these various predictors and the index of rights consciousness within each country.[8] Several conclusions can be drawn on the basis of these data. First, some of the hypotheses must be completely rejected. It is clear that those with higher levels of education,

women, those claiming higher class affiliations, and opinion leaders are no more assertive of rights than their counterparts. Nor is there much of an effect on rights consciousness from levels of religiosity (except perhaps in the the Netherlands and Portugal, where the more religious are likely to claim fewer rights). The lack of effect of education is particularly noteworthy. It appears that rights consciousness permeates the political culture of Western European countries, even among those who have had little formal education. Learning about rights probably reflects widespread social and cultural learning rather than formal education.[9]

Age is of some but not great consequence for rights consciousness. In some countries there is some tendency for older citizens to claim rights (e.g., Luxembourg and France), but in a few countries—Great Britain in particular—the relationship is reversed. In light of the variability of the signs of the coefficients, it is unlikely that the effect of age is actually an effect of aging.

On the other hand, there are reasonably strong relationships between support for democracy and rights consciousness. Those who are more likely to reject democracy as a superior form of government claim fewer rights for themselves, and this is true in nearly every country. This seems to suggest that perceptions of rights are bound up in the ideology of democracy—those who believe that government ought to be accountable to the people are also more likely to assert rights. It is not surprising, therefore, that the particular right that best differentiates supporters of democracy from others is the right to freedom of speech. In most countries, supporters of democracy are considerably more likely to claim the right of free speech than are those who oppose democracy. Rights and democracy seem to go hand-in-hand within the minds of many Europeans.

We also observe some relationship between ideological self-identifications and rights consciousness in most of the countries. Those who perceive themselves as on the *left* are more likely to claim rights, not those who are conservative. This is particularly true in Portugal, although in a few countries, especially Ireland, Luxembourg, and Belgium, there is no difference in the number of rights claimed by those on the left and on the right. This probably reflects the long association of far right politics in Europe with antidemocratic tendencies.

In sum, our efforts to explain variation in rights consciousness within countries has met with some success. In most countries, awareness of rights seems to be tied up (loosely) in a bundle of values related to democracy. Beyond that, there is occasionally a tendency for those who see themselves on the left to claim more rights, although we do not observe this tendency throughout the entire European Community.

We have also found surprisingly weak correlations with basic class and demographic attributes of the respondents, and opinion leaders and ordinary citizens hold similar views of rights. Unlike many political values, it seems that there is actually little variance across the traditional cleavage lines that have

incorporate them into their basic belief systems. Second, the value to individuals of basic democratic rights is especially pronounced during periods of transition from nondemocratic to democratic rule: the mass public has an immediate memory of circumstances—the *ancien regime*—under which these rights had been violated. Third, during the early period of democratization the mass public has little experience with the complexities and trade-offs implied by many of the basic democratic rights. They have yet had to confront the costs and value conflicts associated with supporting democratic rights. For example, government assurances of absolute freedom of speech inevitably mean that certain groups are permitted to promote ideas that are fundamentally unacceptable to certain segments of society.

We expect some degeneration of democratic values the longer a regime enjoys democracy. We have earlier shown that political tolerance declines in the most mature democracies (Duch and Gibson 1992), and we expect a similar fate for rights consciousness. Citizens of mature democracies tend to take many rights for granted and are less likely to be unequivocal in their support for them. As the costs of democratic rights become clear, we expect more conditional support for the exercise of such rights. Thus, contrary to some prevailing models of democratization, we hypothesize that rights consciousness will be highest in nations that recently have undergone democratic revolutions and will be lower in more established democracies.

Thus, we expect to see a negative correlation between experience with democracy and rights consciousness.[11] The plot in Figure 1 of average rights consciousness (as shown in Table 1, above) against the number of years of continuous democratic government in the 12 nations in our sample strongly supports the hypothesis. The bivariate correlation between these two variables is -.78—a strong correlation that is highly statistically significant and in the predicted direction. *Rights consciousness is highest in those countries with the most recent experience with democracy, not among the more mature democracies.*

This negative correlation between years of democracy and rights consciousness also suggests that cultural and institutional understandings of rights may be related in unexpected ways. Undemocratic norms may persist in nations with relatively democratic political institutions, or vice versa. To test this hypothesis, we have borrowed two conventional measures of the degree of institutional protection of civil and political liberties in our twelve European nations.[12] These two measures are also strongly correlated with aggregate rights consciousness, but, as with years of democracy, in an unexpected direction. Where institutional guarantees of political and civil rights are stronger, citizens tend to claim *fewer* rights ($r = -.72$ and $-.71$, respectively). Both of these coefficients are highly significant statistically, despite the small number of cases. Again, this finding indicates that cultural and institutional understandings of right covary, but that citizens' expectations of rights are strongest where institutional guarantees of rights are weakest.

variables is fairly restricted, and Greece, Spain, and Portugal can hardly be thought of as representative of the variety of newly democratized political systems in the world. Moreover, in each of these countries, some democratic traditions predated authoritarian rule, and data from a 1988 survey can barely reveal the vestiges of the democratization period. To get additional insight into the processes by which citizens come to demand rights, it is perhaps useful to turn to a political system currently undergoing democratization, and one without strong democratic traditions. The USSR can serve this purpose nicely.

IV. RIGHTS CONSCIOUSNESS IN A DEVELOPING DEMOCRACY

The analysis above suggests that rights consciousness is most prevalent in countries that are relatively newly democratic. It is perhaps interesting to extend this analysis a bit further by comparing rights consciousness in Western Europe with some newly available evidence from a country just now in the throes of democratization—the Soviet Union.[13] These Soviet data are based on a representative sample of the European portion of the Soviet Union. In-person interviews were conducted with approximately 1,560 Soviet citizens in May 1990.[14] A look at these data can perhaps tell us something about rights consciousness in a society that has traditionally not viewed government as limited or directly accountable to ordinary citizens.

With just a few exceptions, the figures in Table 3 are remarkably similar to the overall EC figures in Table 1. For instance, 74 percent of the (European) Soviets asserted that freedom of speech ought to always be protected, compared to an EC average of 74 percent. The Soviets look uncannily like their Western neighbors in terms of their willingness to assert basic citizenship rights.

The differences between the Soviet respondents and the Western Europeans are also interesting. For instance, the Soviets place less emphasis on freedom of association, no doubt reflecting their long history of weak voluntary interest groups and political parties. The second biggest difference is in the right to cultural autonomy ("the right of people to their own language and culture"), where the Soviets are more jealous of the right than the Western Europeans. This too makes sense given the extraordinary range and legitimacy of cultural diversity within the Soviet Union and the many nationalist interest groups that have emerged in recent years.

Roughly one-third of the Soviet respondents laid claim to all the rights about which they were queried. This places the Soviet Union as among the more rights conscious polities when compared to Western Europe. In comparison to the established democracies of Western Europe—and especially, for instance, to the United Kingdom—the level of rights consciousness in the USSR is remarkably high.

The key question to be asked of these data, though, is whether the same factors influence variation in individual levels of rights consciousness in the

Table 4. Origins of Rights Consciousness, European USSR, 1990

	Pearson Correlation	*Beta*
Support for democracy—I	.21	.09*
Support for democracy—II	.28	.21*
Conservative ideological attachment	-.18	-.12*
Level of education	.21	.11*
Age	-.11	-.02
Gender	-.10	-.04
Religiosity	-.09	-.07
Opinion Leadership	.09	.03

Note: * Regression coefficient is significant at .01 or less.
$R^2 = .14$. Minimum pairwise $N = 880$; average N per variable = 1,472.

Nor do Soviet opinion leaders assert more rights than ordinary citizens. It is possible that this reflects the broad diffusion of rights consciousness throughout the Soviet population. That is, perhaps our survey was conducted too late; perhaps rights consciousness has been fully developed in the Soviet Union following nearly five years of democratization. We expect that opinion leaders are most likely to learn new conceptions of democratic citizenship, and that these ideas percolate down to the mass public fairly slowly. The absence of a difference between leaders and citizens in these data suggests that this process may have been completed in the USSR.

V. CONCLUSIONS

This paper has reported an analysis of the correlates of rights consciousness in Western Europe and the Soviet Union. Generally, we find that at the individual level, those who are more strongly committed to democratic values are more likely to assert rights for themselves. We interpret this to mean that democratic ideology legitimizes citizen rights. It does this by advancing notions of both the limits of government and the inviolability of individuals. Those who value democracy more seem to have a better understanding of the role of the citizen in democratic polities.

At the same time, aggregate levels of rights consciousness are highest in the *newer* democracies of Western Europe. We understand this as being due to the skepticism that citizens hold of government and other major institutions in polities that have only recently emerged from repressive regimes. It also seems that citizens in more mature democracies become less vigilant about protecting their rights. In some respects, this may simply mean that *citizens value rights more when rights are more valuable.*

Our analysis of rights consciousness in the (European) Soviet Union has revealed remarkable similarities between the Soviet people and Western Europeans in terms of rights claimed. At the same time, however, Soviet

Thus, the specific geographical universe from which the sample was drawn includes residents of the Republics of: Estonia, Latvia, Lithuania, Belorussia, Ukraine, Moldavia, Georgia, Armenia, and the portion of Russia west of the Ural Mountains.

2. *The Sample:* A primary sample of 1,590 respondents was selected. In anticipation of the nonresponse rate, a supplementary sample of 410 respondents was also drawn. Respondents from each of the Republics identified above were included in the sample.

This sample is a four-stage stratified sample. At the initial stage of the sampling geographical units were classified according to four statistical indices: (1) the level of industrial development (e.g., the number of plants and factories); (2) the level of well-being of the population (e.g., income per family); (3) ethnocultural peculiarities (e.g., the degree of homogeneity of the distribution of nationalities); and (4) accessibility of cultural amenities (e.g., the numbers of libraries, cinemas, theaters, etc.). These criteria were selected due to their assumed importance in structuring the beliefs of ordinary respondents. The geographical units classified were the oblasts in Russia and the Ukraine, and the Republics elsewhere. A total of 85 units were classified within these 18 homogeneous strata.

At the second stage of the sampling, eight substrata were identified, including: (1) capitals of the union republics with a population more than 1 million; (2) capitals of the union republics with a population less than 1 million; (3) regional centers with a population more than 1 million; (4) regional centers with a population less than 1 million; (5) centers of the autonomous republics, regions, districts; (6) peripheral towns of the autonomous republics, regions, districts; (7) Moscow and Leningrad; and (8) villages.

At the third stage of the sampling, each geographical unit fitting within the 18 * 8 matrix was enumerated, and units were sampled. Sampling points were drawn from each cell in the matrix. Where there was discretion on which of several units to select, experts on the area were consulted in an effort to select the most representative unit. The strategy resulted in 62 sampling points

3. *Respondent Selection:* Within each of these sampling points, respondents were selected through random procedures. Using the records of the address bureaus and farm records, specific named respondents were identified. Thus, unlike many western samples, there was no need to select individual respondents within households using household enumeration methods.

It is important to consider whether the records of the address and farm bureaus constitute a useful sampling frame. Certainly these records are superior to using voting lists, which are derived mainly from the address and farm records themselves. But using these records clearly has some disadvantages, in addition to their many important advantages.

The first question is whether these records fairly completely enumerate the population. We believe they probably do. There is a strong incentive for all

in the project, 236 completed the questionnaire and returned it to Moscow. Like interviewers in the West, these interviewers were overwhelmingly female (71 percent). The average age was 35 years old. Reflecting their associations with various institutes, the interviewers are quite well educated, even though two-thirds report an average monthly income of less than 150 rubles per family member. One-fourth of the interviewers are members of the Communist Party, while two-thirds are members of a trade union. Thus, the Soviet interviewers are better educated than western interviewers, but in most other respects are similar to their western colleagues.

Several steps were taken to discourage cheating on the interviews. First, a fairly elaborate system of supervision was put in place. Second, each questionnaire was carefully checked by a supervisor. Third, with just a few exceptions, interviewers were not allowed to conduct a great number of interviews. The average number of interviews conducted per interviewer is 5.9 (minimum = 1 interview, 3 interviewers; maximum = 22 interviews, 1 interviewer). Finally, a small percentage of cases was selected for verification. No instances of falsification were discovered.

7. *Data Reliability:* As with all surveys, there can be no absolute guarantee of the reliability of all the data collected. All that we can guarantee is that every reasonable step to insure the quality of the data was in fact taken.

ACKNOWLEDGMENT

This is a revised version of a paper delivered at the 1991 Annual Meeting of the International Political Science Association, Buenos Aires, Argentina, July 1991. The Western European survey data analyzed in this paper are drawn primarily from the Fall, 1988, European Omnibus Survey. We would like to thank Karlheinz Reif and the Commission of European Communities for making these data available. We are also grateful for the efforts of Hermann Schmitt of the Zentrum für Europaeische Umfrageanalysen und Studien, Institute for Social Sciences, University of Mannheim, who expeditiously prepared a cleaned version of the data. Some of the variables analyzed here are drawn from the 1989 European Election Study, the first wave of which was fielded as part of the same Omnibus Survey as contained our questions. The 1989 European Election Study is directed by Roland Cayrol, Cees van der Eijk, Mark Franklin, Manfred Kuechler, Renato Mannheimer, and Hermann Schmitt; and funded in part by the British Economic and Social Science Research Council. We are grateful to the project directors for making their data available to us. The Center for Public Policy at the University of Houston and AT&T General Business Systems provided much needed support for this project, and we are indebted to them. Our research efforts were also supported by a Limited-Grant-in-Aid from the University of Houston, for which we are grateful. We also appreciate the comments of Stuart Scheingold on an earlier version of this article.

For the Soviet portion of the paper, we are grateful to several agencies for support, including the National Science Foundation (SES-9003868), the College of Social Sciences and the Limited-Grant-in-Aid program at the University of Houston, the

approved by society (or significant segments of it)" (1983, p. 28). Social learning theory typically emphasizes childhood learning experiences, and we think these important, but we would also treat democratization as a social learning process. The symbols and rhetoric of democratic revolutions—a rhetoric that invariably emphasizes the inviolability of the rights of citizens—no doubt creates in citizens a sense of rights entitlement.

10. Weil (1991) found precisely this sort of relationship in his analysis of change in German public opinion.

11. This is of course a cross-sectional analysis so inferences about longitudinal processes are weak at best. For a longitudinal analysis West German public opinion see Weil (1991).

12. The measures are reported in Bilson (1982, p. 111). Countries that rank high on support for political rights have a score of 3, countries ranked intermediately receive a 2, and low ranking countries have a score of 1. Similarly, the civil rights index ranges from 3 (strong support for civil rights) to 1 (comparatively low support for civil rights).

13. Change is occurring so quickly in the USSR that events since the original data collection (May 1990) have altered virtually every aspect of the Soviet state, including its name. In order to simplify matters, and because that was the name of the country when the data were collected, I will refer to the country as the USSR.

14. For technical details on the survey see the appendix. For earlier reports based on these data see Gibson and Duch (1991c, 1992a, 1992b).

15. Support for democratic values was measured by two questions:

> (1). There is too much democracy in the Soviet Union today. (Agree strongly—Disagree strongly)
>
> (2). Are you in favor of democratic government even if that may lead to a certain amount of insecurity and disruption, or are you in favor of strong government control even if that may lead to a certain amount of regimentation and loss of individual expression?
>
> 1. Democratic government
> 2. Government control
>
> It should be noted that in light of very difficult problems of measurement, no indicator of class identification is available in for the Soviet sample.

16. The European USSR questionnaire represented the joint efforts of the American and Soviet research teams, and the *Los Angeles Times*, which paid for a portion of the research. The late I.A. ("Bud") Lewis, Director of *The Los Angeles Times Poll*, contributed substantially to the design and execution of the survey.

REFERENCES

Almond, G.A. and S. Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations.* Princeton, NJ: Princeton University Press.

_____. 1980. *The Civic Culture Revisited.* Boston: Little, Brown and Company.

Bilson, J.F. 1982. "Civil Liberty: An Econometric Investigation." *Kyklos* 35:94.

Dahl, R.A. 1971. *Polyarchy.* New Haven, CT: Yale University Press.

_____. 1980. "The Moscow Discourse: Fundamental Rights in a Democratic Order." *Government and Opposition* 15:3.

_____. 1982. *Dilemmas of Pluralist Democracy: Autonomy and Control.* New Haven, CT: Yale University Press.

_____. 1989. *Democracy and its Critics.* New Haven, CT: Yale University Press.

Duch, R.M., and J.L. Gibson. 1992. "Putting Up With" Fascists in Western Europe: A Comparative, Cross-Level Analysis of Political Tolerance." *Western Political Quarterly* 45:237.

Sullivan, J.L., M. Shamir, P. Walsh, and N.S. Roberts. 1985. *Political Tolerance in Context: Support for Unpopular Minorities in Israel, New Zealand, and the United States.* Boulder, CO: Westview Press.

Tweedie, J. 1989. "The Dilemma of Clients' Rights in Social Programs." *Law and Society Review* 23:175.

Tushnet, M. 1984. "An Essay on Rights." *Texas Law Review* 62:1363. Weil, F.D. 1985. "The Variable Effects of Education on Liberal Attitudes: A Comparative-Historical Analysis of Anti-Semitism Using Public Opinion Survey Data." *American Sociological Review* 50:458.

Weil, F.D. 1985. "The Variable Effects of Education on Liberal Attitudes: A Comparative-Historical Analysis of Anti-Semitism Using Public Opinion Survey Data." *American Sociological Review* 50:458.

———. 1991. "Structural Determinants of Political Tolerance: Regime Change and the Party System in West Germany Since World War II." Pp. 299-332 in *Research in Political Sociology*, Vol. 5, edited by P.C. Wasburn. Greenwich, CT: JAI Press.

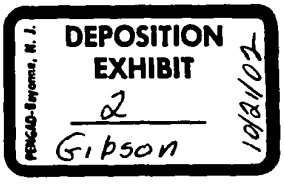
Appendix A: Coding the Commercials

Appendix A shows how the content of the commercials discussed in the book were coded. First, the coding protocol students at ASU used to code the ads is shown, followed by the list of campaign themes from which coders chose answers for questions 23-26, and finally, we show how items were re-coded to correspond to the characteristics shown in the book.

Coding CMAG Storyboards

Please circle the relevant response or enter it in the blank provided. Remember we are not coding commercials for or against ballot propositions or public service announcements.

1. Numeric code identifying ad: _____
2. Does the ad direct the viewer to take any action (as opposed to merely providing information)?
   1. No
   2. Yes
3. (If yes to #2)) What is that action?
   1. To **vote** for someone
   2. To **elect** or **re-elect** someone
   3. To **support** someone
   4. To **vote against** someone
   5. To **defeat** someone
   6. To **reject** someone
   7. To **write, call** or **tell** someone instructing them to do something (e.g. to vote for or against a bill)
   8. To **urge** action or attention to a particular matter
   9. To **send a message** or **call** someone to express yourself on a particular subject
   10. Other
4. (If a an ad asks people to contact a public official [choices 7 through 9 above]) Does it provide a specific bill number to discuss or urge action on?
   1. No
   2. Yes
5. (If an ad asks people to contact a public official) Does it provide a phone number or address to help them do so?
   1. No
   2. Toll-free telephone number listed
   3. Toll number listed
   4. Address listed
6. In your opinion, is the purpose of this ad to provide information about or urge action on a bill or issue, or is it to generate support or opposition for a **particular candidate?**
   1. Provide information or urge action (If so, skip to question #19)
   2. Generate support/opposition for candidate
   3. Unsure/unclear
7. Is the favored candidate...
   1. Mentioned by name in the text of an ad?
   2. Pictured in the ad?
   3. Both mentioned and pictured in the ad?
   4. Not identified at all?
8. Is the favored candidate's opponent...
   1. Mentioned by name in the text of an ad?
   2. Pictured in the ad?
   3. Both mentioned and picture in the ad?
   4. Not identified at all?
9. In your judgement, is the primary purpose of the ad to **promote** a specific candidate ("In his distinguished career, Senator Jones has brought millions of dollars home. We need Senator Jones"), to **attack** a candidate ("In his long years in Washington, Senator Jones has raised your taxes over and over. We can't afford 6 more years of Jones.") or to **contrast** the candidates ("While Senator Jones has been raising your taxes, Representative Smith has been cutting them.")?
   1. Promote
   2. Attack
   3. Contrast
   4. Unsure/unclear
10. Does the favored candidate appear on screen narrating his or her ad?
    1. No
    2. Yes
11. Is the office at stake mentioned in the ad?
    1. No
    2. Yes - referred to in text of the ad
    3. Yes - written in one (or more) of the visual frames of the ad
    4. Yes - referred to in both the text and in the visuals of the ad



12. Is an opponent's commercial mentioned or shown on screen?
    1. No
    2. Yes—opponent's commercial is referred to in the text of the ad
    3. Yes—opponent's commercial is shown on screen during segment of the ad
    4. Yes—opponent's commercial is referred to in text and on screen
13. Does the ad use any of the following adjectives to characterize the favored candidate? (First mention)
    1. Honest
    2. Caring/Compassionate
    3. Conservative
    4. Moderate/middle of the road/mainstream
    5. Proven/tested/experienced
    6. Independent
    7. Competent/knows how to get things done
    8. Tough/a fighter
    9. Protector
    10. Friend of Clinton
    11. Values (shares them, has American ones...)
    12. Other
    20. No adjectives or descriptions of candidate
14. Does the ad use any of the following adjectives to characterize the favored candidate?

    (Second mention)
    Enter code (see #13 above) _____
15. Does the ad use any of the following adjectives to characterize the opposing candidate? (First mention)
    1. Liberal
    2. Reactionary/right-wing
    3. Extremist/radical
    4. Dishonest/corrupt
    5. Incompetent
    6. Heartless (may be used in reference to Social Security)
    7. Friend of Newt Gingrich
    8. Taxing (or some version of liking taxes)
    9. Friend of special interests
    10. Friend of Clinton
    11. Friend of Pat Robertson/religious right
    12. Other
    20. No adjectives or descriptions
16. Does the ad use any of the following adjectives to characterize the opposing candidate? (Second mention)

    Enter code (See #15 above) _____
17. Does the ad mention the party label (i.e. Democrat or Republican) of the favored candidate or the opponent?
    1. No
    2. Yes—favored candidate's party
    3. Yes—opposing candidate's party
    4. Yes—both candidates' party affiliations are mentioned
18. Does the ad use technology to distort (i.e. "morph") the opposing candidate's face?
    1. No
    2. Yes
19. Is the ad funny or is it intended to be humorous?
    1. No
    2. Yes
20. Does the ad refer to newspaper stories or editorials?
    1. No
    2. Yes—in the text part of the text of the ad
    3. Yes—in the visuals of the ad
    4. Yes—in both the text and visuals
21. Does the ad cite supporting sources (including as footnotes) to bolster various claims?
    1. No
    2. Yes—in the text part of the text of the ad
    3. Yes—in the visuals of the ad
22. In your judgement, is the primary focus of this ad on the personal characteristics of either candidate or on policy matters?
    1. Personal characteristics
    2. Policy matters
    3. Both
    4. Neither

Campaign themes: Attached are a list of issues that frequently come up in the course of political commercials, such as education, the environment, and taxes. Please list the ones that come up in this ad in the order in which they appear. (Note: We are currently compiling a list from the NES's most important problem codes.)

23. First mention _____
25. Third mention _____
24. Second mention _____
26. Fourth mention _____

Campaign Themes

Personal characteristics of the candidate(s)

1. Background
2. Political record
3. Attendance record
4. Ideology
5. Personal values
6. Honesty/Integrity
7. Special Interests

Policy issues

Economy

10. Taxes
11. Deficit/surplus/budget/debt
12. Government spending
13. Minimum wage
14. Farming (e.g. friend of)
15. Business (e.g. friend of)
16. Employment/jobs
17. Poverty
18. Trade/NAFTA
19. Other economic reference

Social issues

20. Abortion
21. Homosexuality
22. Moral values
23. Tobacco
24. Affirmative action
25. Gambling
26. Assisted suicide
27. Gun control
28. Other reference to social issues

Law and Order

30. Crime
31. Drugs
32. Death penalty
33. Other reference to law and order

Children

40. Education
41. Lottery for education
42. Child care
43. Other child-related issue

Foreign Policy/Defense

50. Defense
51. Missile defense/Star wars
52. Veterans
53. Foreign Policy
54. Bosnia
55. China
59. Other defense/foreign policy issue

Clinton

60. Clinton
61. Ken Starr
62. Whitewater
63. Impeachment
64. Sexual harassment/Paula Jones

Other/assorted

70. Environment
71. Immigration
72. Health care
73. Social security
74. Medicare
75. Welfare
76. Civil rights/race relations
77. Campaign finance reform
78. Government ethics
95. Other
99. None

Re-coding the Storyboards in Chapters 1 through 7

Chapters 1 through 7 provide analysis of the content of political advertising using the coding system shown in Appendix A. This coding, in turn, was used to determine the characteristics of ads examined throughout the volume. Note: MD refers to missing data, usually the result of coding item being inappropriate for a given storyboard.

Item	Coding
Use of "Magic Words" (i.e. "vote for, vote against, support...")	Yes: Response 1-6 to question 3. No: Response 7-10 to question 3, MD.
Urge action (Did ad urge viewers to take an action?)	Yes: Response 2 to question 2. No: Response 1 to question 2, MD.
Mention candidate[1] (Did the ad mention any candidate?)	Yes: Responses 1-3 to question 7 or question 8. No: Responses 4 to question 7 and question 8, MD.
Office mentioned (Did ad mention the political office that the candidates were competing for?)	Yes: Response 2-4 to question 11. No: Response 1 to question 11, MD.
Tone (Did the ad promote, attack, or contrast the candidates?)	Promote: Response 1 to question 9. Attack: Response 2 to question 9. Contrast: Response 3 to question 9. Don't know: Response 4 to question 9, MD.
Candidate narration (Did a candidate narrate the ad?)	Yes: Response 2 to question 10. No: Response 1 to question 10, MD.
Political party mentioned (Was a political party referred in the ad?)	Yes: Responses 2-4 to question 17. No: Response 1 to question 17, MD.
Purpose (Was the purpose of the ad to promote a candidate or an issue?)	Promote candidate: Response 2 to question 6. Promote issue: Response 1 to question 6. Don't know: Response 3 to question 6, MD.
Focus (Did the ad focus mainly on personal characteristics of a candidate or policy?)	Personal: Response 1 to question 22. Policy: Response 2 to question 22. Don't know/both: Responses 3-4 to question 22, MD.
Identify bill (Did the ad identify a specific bill?)	Yes: Response 2 to question 4. No: Response 1 to question 4, MD.
Toll-free number (Did the ad provide viewers with a toll-free phone number?)	Yes: Response 2 to question 5. No: Responses 1,3,4 to question 5, MD.
Use of adjectives (Did the ad use adjectives to characterize any candidates?)	Yes: Responses 1-12 to questions 13-15. No: Response 20 to questions 13-15, MD.
Themes[2] (What were the main issues raised by a commercial?)	Codes 1-99 in questions 23-26. See Appendix 2.

[1] Candidate mentions were initially coded only for ads whose purpose was perceived as generating support or opposition to a candidate. Figure 4.22, however, is based on a subsequent examination of all ads, regardless of their purpose.

[2] Ads may have multiple themes, such as "Mr. Smith is a pro-life conservative fighting to keep taxes down."

Appendix B: Measurement Issues

Appendix B explores a few of the most important measurement issues that affect the results in *Buying Time*. Some of the issues involved are internal to the data themselves, while others arise in the course of coding and analyzing the information. Our intent in all cases is to make our decisions as transparent as possible by providing this appendix.

I. Viewership and Cost Estimates

The price of an individual television spot is a function of a show's popularity (rating) and the size of the market where it is airing. Knowing on what show a spot was aired and in which market, CMAG uses information from a variety of sources to provide measures of the cost and viewership—measured in gross ratings points—of each commercial aired. Both of these measures are estimates with various strengths and weaknesses.

Dollars: Since discovering the actual amount paid for each spot would be virtually impossible, the CMAG figures represent the average cost of a particular time slot across stations in a particular market (i.e. the average cost of a 30 second spot in Phoenix on Tuesday from 7:30 to 8:00 p.m.). The u e of averages, however, suggests that the price of some spo are overstated or, much more likely, understated. Special . cumstances may cause a particular slot to be more expens or cheaper) than usual.[1] A more systematic problem, howev is that, in contrast to commercial advertisers, political adver ers are completely hostage to the calendar. Although Coca C la may prefer that a set of ads appear in the first week of November than in the second, all soft drink sales count as revenue. But political ads run after election day are a complete waste. The result is anecdotal data suggesting that many political advertisers pay top dollar in the final weeks of the campaign for some ads.[2] The CMAG data miss this last-minute inflation.

In addition, the price of the air time does not include the cost of producing and placing commercials, or the cost of advertising on local cable stations. All that said, the spending estimates for candidates—the only advertisers about whom much is known—in *Buying Time* are roughly in line other studies of political advertising. The most exhaustive studies of candidate expenditures come from the 1990 and 1992 editions of *The Handbook of Campaign Spending*.[3] The authors found that electronic (television and radio) advertising in the 1992 elections accounted for an average of 42 percent of money spent by Senate candidates and 27 percent of the money spent by House candidates.[4] How do the CMAG figures compare?

	House	Senate
FEC estimate[5] (in millions of dollars)	385.88	248.51
20% (of nation not covered by CMAG)	308.70	198.81
Estimated spending on TV (27% House, 42% Senate)	83.35	83.50
15% radio and local cable[6]	70.85	70.98
15% production and placement costs[7]	60.22	60.33
CMAG	47.33	85.23

[1] For example, a 30-second ad run by Wisconsin Senator Russell Feingold on Sunday, November 1 at 2:55 p.m. in the Green Bay media market was reported to have cost $37. While Sunday afternoon advertising on Green Bay stations may usually be this inexpensive, it is likely that this ad, run two days before the election and during a Green Bay Packers football game, cost much more than $37.

[2] See David Magleby, *Outside Money: Soft Money and Issue Ads in Competitive 1998 Congressional Elections* (A Report to the Pew Charitable Trusts, 1998).

[3] Dwight Morris and Murielle E. Gamache, *Handbook of Campaign Spending: Money in the 1992 Congressional Races* (Washington, D.C.: CQ Press, 1994) 10.

[4] Critics might object that the *Handbook's* results are out of date, but there is no reason to expect that campaigns became much more dependent on television advertising in the six years between 1992 and 1998. If they have, it is probable that much of the new expense may have been devoted to buying time on local cable stations, an activity not monitored by CMAG.

[5] General election candidates: See http://www.fec.gov/press/hgnlng98.htm and http://www.fec.gov/press/sgnlng98.htm.

[6] We assume that 85 percent of the average media budget is spent on broadcast television.

[7] The cost of producing an ad and buying the advertising time to air it. Ad agencies frequently charge a fee for purchasing air time, reflecting their supposed expertise in designing a strategy to reach the most appropriate audience. These fees, once 15 percent of the cost of the buy, have declined considerably in recent years as advertisers have objected. The combined 25 percent estimate for production and placement is a fairly conservative guess.

AN ANALYSIS OF THE 1998 AND 2000 BUYING TIME REPORTS

Dr. James L. Gibson

September 30, 2002

INTRODUCTION

I was retained by counsel for the AFL-CIO, the National Association of Broadcasters, and Senator Mitch McConnell to examine two reports produced by the Brennan Center on the subject of "issue advocacy" or political communication. These reports, and their abbreviated titles, are:

Buying Time 1998:

Krasno, Jonathan S., and Daniel E. Seltz. 2000. *Buying Time: Television Advertising in the 1998 Congressional Elections*. New York: Brennan Center for Justice.

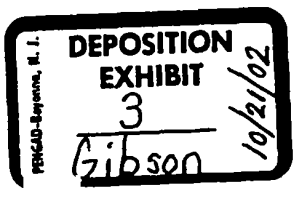
Buying Time 2000:

Holman, Craig B., and Luke P. McLoughlin. 2001. *Buying Time 2000: Television Advertising in the 2000 Federal Elections*. New York: Brennan Center for Justice.

In addition, I have reviewed the depositions of three of the authors of these reports (Seltz, Holman, and McLoughlin). Finally, I directly examined and analyzed the raw data upon which these two reports were based. The analytical and statistical methods and techniques I have used in preparing this report are those customarily employed in quantitative analyses of social scientific data.[1]

---

[1]A complete listing of the materials upon which I relied in preparing this report is attached as Exhibit 1.



QUALIFICATIONS

Evaluating these data and reports is entirely within my expertise and competence, as documented in my vita.[2] A few highlights are worth considering:

- I have taught statistics and data analysis for nearly thirty years, including regularly teaching a three-semester sequence on advanced statistics and data analysis for graduate students in political science and other social sciences.

- I have published widely in the leading political science journals, worldwide. I regularly contribute to what are considered to be the three most prestigious journals in our discipline: the *American Political Science Review*, the *Journal of Politics*, and the *American Journal of Political Science.*

- My research has been supported by millions of dollars of grants from the National Science Foundation.

- My published research has won several awards, including an award in August, 2002, for the best paper presented at the Annual Meeting of the American Political Science Association.

- I serve as a peer reviewer for many different publications and agencies, including approximately two dozen academic journals, several publishing houses,

---

[2]A copy of my curriculum vitae is attached as Exhibit 2 to this report. In my vita, I list all of my articles and books published in the last 10 years.

I also note that, in the last four years, I have been deposed or testified at trial only a single time. My deposition was taken in *Warren vs. St. Louis Police Officers Association*, on February 22, 2002. My role may only have been that of fact witness, rather than an expert witness, but I note this deposition nonetheless.

My consulting fee on this case is $250 per hour (plus expenses). To date, I have worked approximately 194.25 hours, for a total fee of $48,562.50.

and public and private research foundations ranging from the National Science Foundation and the Australian Research Council to the Russell Sage Foundation. I have also served as Associate Editor of the journal *Law and Society Review*.

Thus, I have had vast experience with data collection and analysis in projects precisely comparable to that reported in the Buying Time reports.

Because the 1998 and 2000 reports differ in many important respects, I report the results of my analyses separately.

BUYING TIME 1998

Summary of Conclusions about *Buying Time 1998*

- *Buying Time 1998* should not be accepted as the product of scientific inquiry, but is instead policy advocacy written by people with a strong ideological commitment to a particular position on campaign finance reform.[3]

---

[3]Nowhere are the intentions and motives of the investigators better documented than in the research proposal "Issue Advocacy: Amassing the Case for Reform." See Holman Deposition, Exhibit #3 (copy attached as Exhibit 3). In that proposal, conclusions about the extent of illicit issue advocacy and the dangers it poses are clearly asserted, prior to any data collection or analysis. Furthermore, the authors clearly state their policy objectives: "....the purpose of our acquiring the data set is not simply to advance knowledge for its own sake, but to fuel a continuous and multi-faceted campaign to propel reform forward" (p. 2). Further: "The questions we expect to answer are not part of some abstract inquiry. Rather, they are linked directly to a strategy for promoting reform" (p. 3). One of the authors of the 1998 Report notes that one of the goals of the study was explicitly to provide empirical data that would help the campaign finance reform effort. See Seltz Deposition, p. 23, lines 4-5. (Referenced excerpts from the Seltz Deposition are included as Exhibit 4.) Further, he acknowledges that "we were constantly consulting with attorneys at the Brennan Center" (p. 68, lines 12-13) in making decisions about the data base. The strong policy and ideological commitments of the investigators are not compatible with the conventional canons of scientific objectivity and may have undermined the integrity of the data collection and analysis.

- In general, the conclusions drawn in *Buying Time 1998* should not be accepted because the data upon which the report is based are fundamentally flawed.

- Indeed, to the extent that any reliable inferences can be drawn from the data, the conclusion would be that a very large percentage of the advertisements whose purpose was coded as seeking to influence an election in fact had policy issues as their primary focus.

General Comments About the Report and the Investigators

At the outset, several characteristics of *Buying Time 1998* should be noted:

- This Report is not the product of any peer-review process. The Report was published by the Brennan Center, not by a commercial or academic press. Thus, the Report was not vetted in any way whatsoever prior to its publication, and consequently the normal process of explication of the project methodology, error correction, and review of substantive conclusions prior to publication did not take place. This seriously limits the confidence one can place in the Report.

- The Report largely consists of a series of tables filled with quantitative information. The tables are rarely explained, analyzed, or interpreted. For instance, Chapter 4 consists of one page of text (p. 87), then 21 pages of tables (pp. 88 - 108), a half page of text (p. 109), and another page with tables (p. 110). Such a chapter cannot be taken as serious scholarship. Serious scholarship would carefully explain the objectives of each of the analyses, the details of how the data

were manipulated, and would carefully and clearly link any substantive conclusions with specific empirical evidence. None of this takes place in this chapter, or in this report more generally. This report was surely published "in-house," by the Brennan Center itself, because, based on my experience as a reviewer for publishing houses (and the sort of review I would submit were I called upon for publication advice), I doubt whether any academic publisher would publish a document like this.

- No data base has been (nor can be, it appears) produced that will generate the specific numbers found in this Report. This is in part because the data set is continuously being manipulated and changed. But it is also a function of the lack of transparency in the statistical analysis that underlies these tables. I will give specific illustrations of this throughout my report, but most tables do not provide sufficient information for an analyst to replicate the findings. In the social sciences, we demand that statistical analysis be replicable (that another investigator using the same data be able to reproduce exactly the same findings). This report is not replicable, and that undermines tremendously any confidence one should place in the findings produced.

- The Report is filled with questionable statistical techniques and applications.

- Finally, I must note that the authors of the Report apparently were not involved in collecting either the data about the airings (from CMAG) or the data describing the ads (from Professor Goldstein). As with any secondary analysis of

data collected by others, this most likely limits the authors' understanding of the nuances and peculiarities of the data base. This is particularly so since this is a demanding data base involving large numbers of cases, multiple units of analysis, various data infirmities (e.g., missing data), and variables based on highly subjective judgments made by undergraduate students.[4] Moreover, I note that one of the authors (Mr. Seltz) seems to have little if any training in statistical analysis, apparently learning whatever skills of data analysis he possesses on-the-job, in the course of preparing this report. Given that this data base is large, complicated, and difficult to analyze, it is extremely worrisome that the results are so heavily dependent upon the limited skills of an author who is a novice analyst.

The Sources of the Data for *Buying Time 1998*

The data for *Buying Time 1998* are drawn from two sources. First, information about the broadcasting of these advertisements was collected by Campaign Media Analysis Group (CMAG). Second, information about the content of the ads was coded by undergraduate students at Arizona State University under the direction of Professor Ken Goldstein. Both of these data sources require considerable scrutiny.

The first source of data analyzed by the authors of *Buying Time 1998* is the information

---

[4]Professors Krasno and Goldstein acknowledge that the coding of some of the variables in the data set — explicitly including the so-called purpose of the ad — was a subjective enterprise. See Krasno, Jonathan, and Kenneth Goldstein. 2002. The Facts about Television Advertising and the McCain-Feingold Bill." *PS: Political Science and Politics* 35 (#2, June): 207-212, at page 207 and at page 209. A copy of this report is attached as Exhibit 5.

about ad broadcasts. These data were purchased from CMAG, which relies upon a technology that purports to recognize "each separate commercial run, sending the storyboard (full audio and every four seconds of video) of every commercial to CMAG headquarters" *(Buying Time 1998*, 7). There are of course many limitations to the CMAG data, including:

- CMAG does not monitor all broadcasts; instead, its coverage is limited to only 75 media markets (out of 210 such markets in the country).[5] Given this limitation, one must be particularly careful about generalizing the findings of this study to all political communications.

- Media markets are not coterminous with electoral districts (e.g., the New York City media market), meaning that ads broadcast throughout a media market are reaching citizens with different relationships to politicians, candidates, and elections.

- The CMAG technology was not able to distinguish ads differing in only a few words, treating, for instance, "cookie cutter" ads mentioning Senator Kennedy as identical to ads mentioning Senator Boxer. Since ads apparently cannot be uniquely identified by the technology, the information in the data base about when any given specific ad was aired most likely contains errors.

- Complete visuals are not provided, meaning that some information depicted in the ad may be excluded and not captured, thus compromising the

---

[5]Krasno, Jonathan, and Kenneth Goldstein. 2002. The Facts about Television Advertising and the McCain-Feingold Bill." *PS: Political Science and Politics* 35 (#2, June): 207-212, at page 207. See Exhibit 5.

coders' ability to analyze the complete content of the ads.

- The technology is of course dependent upon the nature of the ads themselves. If ads display information in a way that is illegible, then the CMAG images will themselves be illegible.

- Not all methods of political advertising and communication are captured in the CMAG data, as for instance communication via radio and local cable ads.

- Crucial information was missing in the CMAG data, as for example, in more than a quarter of the broadcast data, the identity of the sponsor was not included in the data *(Buying Time 1998*, 8). Without accurate information about the sponsor of any given ad, the fundamental distinction drawn in *Buying Time 1998* among ads sponsored by candidates, parties, and groups becomes suspect. It appears that in some instances the ad sponsor was identified through methods independent of the CMAG data collection and analysis, although these methods are not thoroughly documented in the Report (e.g., *Buying Time 1998*, p. 8).

- The CMAG data apparently include subjective variables coded by CMAG staff, including judgments of the content of each ad and estimates of the cost of each ad. For example, it cannot be determined whether CMAG captured every ad that included the name of a public official or just those which CMAG deemed to be "political" in nature.

- Finally, no evidence has ever been adduced documenting the accuracy of the CMAG data. Indeed, insofar as the 1998 data were collected and assembled through methods similar to those apparently used in 2000, the lack of accuracy of

the data has been documented.[6]

A second source of data comes from the review of the ads by student coders at Arizona State University. The students were asked to respond to 26 questions about the ads, using printed copies of the storyboards (see *Buying Time 1998*, pp. 193-194, for a copy of the questions asked). The coding sheet has some obvious and important errors (as I discuss in more detail below).

A series of significant questions emerges about the process by which these ads were coded. For instance, it is unclear how the students were recruited, what expertise they had prior to being employed for the project, whether the students had been exposed to Professor Goldstein's classes, whether the students had ideological and/or policy commitments to a particular outcome in the project, etc. Insofar as the 1998 data collection process is concerned, none of these details is presented in the published reports or in other material made available to me. The absence of answers to these questions raises questions about the overall accuracy of the data collection process.

The question of how the students were trained to make these difficult judgments is of considerable relevance to the quality of the data they recorded. *Buying Time 1998* discloses nothing about coder training. However, in a 2002 report, Krasno and Goldstein indicate that the student coders received *no training whatsoever*, although I acknowledge that there is some ambiguity about whether this practice (or lack thereof) pertains to both the 1998 and 2000 data

---

[6]No detailed description of the 1998 CMAG data has been published or otherwise made available to me. However, in Appendix E to Professor Goldstein's Report for this litigation, he discusses various problems in 2000 with the CMAG data and technology. The data difficulties seem to be numerous and formidable. Given the general trend for technology to improve over time, it is difficult for me to imagine that the 1998 CMAG data are any less problematic than the 2000 data.

sets.[7] They assert: "This consensus [referring to the students reaching the same conclusions about the "tone" of the ads] is perhaps more remarkable because we chose not to embark on any training program for coders, preferring instead that they use their common sense to give us a better feel of how the average viewer would characterize these ads."[8] If I am correct that the student coders were not trained, then this is a flaw of considerable proportion. Not only are undergraduates at Arizona State University (or, in the case of the 2000 study, the University of Wisconsin) indisputably not a representative sample of the "average viewer," in the absence of training, the students were apparently free to exercise unstructured discretion in coding the ads. Without instruction and guidelines for what constitutes the difference between "providing information or urging action" and "generate support/opposition for candidate" — without training, practice coding, and discussion of coding rules based on the results of the practice coding — I do not believe that undergraduate students coders can make accurate assessments on highly subjective characteristics of these ads. This is, in my judgment, an extremely important limitation to the data generated by the coding process, leading me to have little confidence in the results.

A Data Set in Constant Flux

Based on the data sets on the CDs provided by the Brennan Center and Professor

---

[7]My reading of this report leads me to conclude that Krasno and Goldstein were describing their decision not to train the coders in 2000, but I have seen no documentation to suggest that their coders were trained in some way or another in 1998.

[8]Krasno, Jonathan, and Kenneth Goldstein. 2002. The Facts about Television Advertising and the McCain-Feingold Bill." *PS: Political Science and Politics* 35 (#2, June): 207-212, at page 211. See Exhibit 5.

Goldstein, it is apparent to me that no single *Buying Time 1998* Data Set exists. This is in part due to the fact that Professor Goldstein was (and may still be) continuously making changes in the codes assigned to individual ads and airings. Of course, one would always correct obvious factual mistakes in objective variables, but to alter subjective variables of crucial importance requires a great deal more discussion, transparency, and justification.

Moreover, the motives for making such changes are important. To the extent that one only examines codings that undermine the preferred conclusions, and one does not examine codings supporting the preferred conclusion, asymmetrical bias is introduced in the data set. Under such conditions, confidence in the ability of the data set to produce useful results and conclusions plummets.

The Reliability and Validity of Coding the Ads

Since the accuracy of the coding of each individual ad is unusually important for the results drawn from this data base, it is crucial that the coding procedures and processes receive additional close scrutiny. Social scientists use the term "reliability" to refer to the accuracy of data collection. I will therefore adopt this convention by considering the degree to which the coding of the ads is reliable.

Some of the characteristics of the ads are no doubt easy to code with considerable reliability. Entirely objective characteristics of the ads (e.g., whether a telephone number is mentioned in the text of the ad[9]) present few threats to reliability (although clerical errors are

[9]This comment only refers to the text of the ad. Since the storyboards are only captured in 4-5 second intervals, a telephone number not mentioned in the text but presented visually for a period of less than 4 seconds would not be properly coded with this methodology.

always possible, if not probable, in a large data base such as this).

However, other crucial attributes are far from being objective characteristics of the ads; instead, they are highly subjective and judgmental. For instance, consider Question 6 (from page 193, *Buying Time 1998*, emphasis in the original):

> 6. In your opinion, is the purpose of this ad to provide information about or urge action on a bill or issue, or is it to generate support or opposition for a **particular candidate**?
>
> 1. Provide information or urge action
>
> (If so, skip to Question #19)
>
> 2. Generate support/opposition for candidate
>
> 3. Unsure/unclear

To answer this question requires a large number of subjective assessments. First, some judgment must be made about *whose* purpose is under consideration here. For some ads, the sponsor is readily apparent, but for others it is not, so the coders were often faced with the difficult and ambiguous task of considering to whom or what to attribute "purpose." The unreliability of the coding process is reinforced by the absence of any explicit guidelines for how to ascertain an ad's "purpose."

Second, the decisions of the coders on these ads were often overruled by Professor Goldstein. For example, consider the ad entitled "NPLA/Call Feingold and Kohl":

> [Announcer]: America was outraged when two New Jersey teenagers checked into a Delaware hotel and delivered and disposed of their newborn baby in the dumpster. Most Americans couldn't believe that defenseless human life could be

so coldly snuffed out. But incredibly, if a doctor had been present that day in Delaware and delivered the infant, all but one inch from full birth and then killed him, it would have been perfectly legal. Instead of murder or manslaughter it would have been called a partial-birth abortion. Killing late in the third trimester. Killing just one inch from full birth. Partial birth abortion inflicts a violent death on thousands of babies every year. Your Senators Russ Feingold and Herbert Kohl voted to continue this grisly procedure. Contact Senators Feingold and Kohl today. Their number in Washington is 202-224-3121. [Announcer]: To join the fight against partial-birth abortion, contact the National Pro-Life Alliance.

The visuals in frames 11 through 14 of the ad read:

CALL SENS. FEINGOLD AND KOHL

AT (202) 224-3121

TELL THEM TO VOTE FOR THE

PARTIAL BIRTH ABORTION BAN

This ad is especially interesting for several reasons. First, it seems obvious that the central focus of the ad is on the policy issue of whether to ban partial birth abortions. If one had to speculate about the motives behind the ad, one might reasonably judge that the ad sought to capitalize on the widely publicized incident in Delaware as a means of generating support for a congressional ban on partial birth abortions.[10] One might also reasonably conclude that one purpose of the ad was to elicit support for the National Pro-Life Alliance. The most reasonable

[10]Note that the number depicted in the ad is the telephone number of Congress.

overall assessment of this ad is that it is an example of issue advocacy by an interest group.

The student coders at Arizona State University recognized this as an issue ad, originally coding it in Question 6 as: "1. Provide information or urge action". However, subsequent to that coding, the data were apparently changed so that the value in the data set of variable Q6 (representing the coding of Question 6) is: "2. Generate support/opposition for candidate." After the coding process was completed, Professor Goldstein apparently decided to overrule the judgment of the coders and substitute his own view that this ad represents electioneering.[11] Though this decision was apparently controversial — indeed, it is striking that one of the authors of *Buying Time 2000* concluded in an e-mail to Professor Rick Hasen: "It reads to me like a genuine issue ad."[12] — since the final version of the data base supplied to me records this as an electioneering ad, Professor Goldstein apparently thought otherwise, and thus the Buying Time studies (in both 1998 and 2000) treated this as a "sham ad."[13] This process hardly contributes to the reliability of the variables in the data base.

---

[11]One of the authors of *Buying Time 1998* reports that the data set was constantly evolving over time. Worse, it is not possible to "track the evolution of all the changes we made." See Seltz Deposition, p. 52, lines 20-21. (Exhibit 4.) Moreover, my comparison of the student coding sheets for several of the 1998 storyboards with the scores in the "final" version of the data reveals several instances of inconsistency, which apparently were brought about by the post-coding manipulations of Professor Goldstein.

[12]Exhibit #14, Holman Deposition. E-mail on 01/18/2001 from Luke McLoughlin (one of the authors of *Buying Time 2000*) to Rick Hasen. (A copy of the E-mail is attached as Exhibit 6.)

[13]I understand the Professor Goldstein has indicated in his expert report of September 23, 2002, that he now views this ad (when broadcast in 2000) as a "genuine issue ad." The data set provided to me does not reflect any re-examination of this ad (which also aired in 1998). Rather the data set indicates that it was recorded by Professor Goldstein as a "sham ad."

I have examined the actual student coding sheets for 25 of the 1998 storyboards[14] and have compared them to the "final" version of the 1998 data set. Focusing on the coding of the extremely important question requiring the coders to ascertain the "purpose" of the ad, I discovered that changes were apparently made in at least eight ads between the student coding and this version of the data base.[15] On its face, eight may not seem like an especially large number, but these ads are responsible *for over 2400 airings* in the 1998 data base. Furthermore, considering the 20,432 airings in 1998 coded as sponsored by groups and as independent expenditures or issue ads, making Professor Goldstein's changes on just these 8 ads results in the percentages of ads said to be providing information or urging action decreasing from 70.2 % (based on the original student coding) to 58.4 % (based on the Goldstein recoding). Moreover, the changes in the data base are entirely asymmetrical: In not a single instance in these storyboards was a change made on an ad originally coded as having candidate support or opposition as its "purpose." Since no documentation of how individual ads were selected for reconsideration by Professor Goldstein has apparently been produced, one is left wondering why all of these changes could have had the same effect: issue ads being converted to candidate support ads. My most important observation is thus that the changes apparently made by Professor Goldstein have a rather dramatic impact on whether these ads are considered to be

---

[14]The id numbers for these coding sheets are: 2, 7, 10, 11, 12, 15, 16, 21, 22, 605, 797, 1043, 1045, 1355, 1374, 1411, 2406, 2512, 2521, 2531, 2532, 2533, 2637, 2640, and 3024. I also examined the coding sheets for ad #137, but since there are two separate sheets for this ad, and since the sheets do not agree between themselves on some of the codes, I have excluded that ad from this analysis.

[15]The coding sheets and the storyboards for these eight ads are attached as Exhibit 7 to this report.

issue- or candidate-oriented.

There can be little question that the undergraduates were asked to make difficult subjective assessments of ads that are often ambiguous and ambivalent. How accurate are these subjective judgments? Social scientists do indeed often attempt to quantify subjective phenomena. But in doing so, certain procedures are essential so that the reliability of the data collected can be assessed. The methodology that is necessary involves the assessment of "inter-coder reliability."

One common meaning of reliability is the ability to replicate or reproduce results. Consequently, a proper methodology for assessing inter-coder reliability would involve the following procedures.

1. The ads are coded.

2. A sample (or the population) of the ads is coded for a second time by an independent coder, using exactly the same procedures as the initial coding. The sample is typically randomly selected, although stratified sampling is sometimes used (e.g., to assure the inclusion of infrequently occurring types of cases). It is necessary that all procedures on the subsequent coding be identical to those of the initial coding, therefore precluding, for example, the use of "expert" or highly experienced coders for the subsequent coding.

3. A variable-by-variable assessment of the two codings is then conducted. An inter-coder reliability coefficient is calculated for each variable that indicates the degree of consensus between the codings. Thus, coding cases should not be said to be reliable; reliability is an attribute attaching to individual variables.

Using these procedures, one can assess the degree to which subjective phenomena have been coded reliably.

Social scientists care about more than reliability; they also care dearly about validity. One's bathroom scale can produce reliable results day after day, but if the scale is not properly calibrated, the results are not valid in the sense that the weight the scale reports is not an accurate representation of one's true weight (in this instance, many of us prefer results that are not valid). It is possible, for instance, that coders could reliably but invalidly code an attribute of the ads. For instance, assume that the coders are asked to judge whether the "purpose" of an ad is to "generate support/opposition for candidate" [sic — from the 1998 coding form]. Assume further that this is a highly subjective judgment.[16] Consequently, coders must seek easily discernable "cues" in the advertisements as a means of making the required judgment. Since the presence of a political figure who seems to be a candidate is a readily accessible cue, the coders then develop an implicit decision rule that says: "when a political figure is depicted in the ad, the ad involves electioneering." Using this rule, the variable might be reliably coded. But this does not mean that the data are *valid*, since political figures appearing in ads could well be doing something other than electioneering. Reliability is the *sine qua non* of useful analysis of subjective data, but reliability must also be considered in the context of validity. Without validity, the meaning of the

---

[16]There is apparently no dispute among any of those involved in producing the Buying Time reports that the coding of these attributes of the ads is highly subjective — e.g., (as I have previously cited), see Krasno, Jonathan, and Kenneth Goldstein. 2002. The Facts about Television Advertising and the McCain-Feingold Bill." *PS: Political Science and Politics* 35 (#2, June): 207-212, at page 207 and at page 209 (see Exhibit 5). See also the Holman Deposition, pages 68, 72-73, (referenced excerpts from the Holman Deposition are attached as Exhibit 8), and the McLoughlin Deposition, page 38 (referenced excerpts from the McLoughlin Deposition are attached as Exhibit 9).

variables is conflated and confused. Thus, it is crucial to consider both the reliability and validity of the data produced by Professor Goldstein.

Unfortunately, the 1998 study apparently involved no assessment whatsoever of inter-coder reliability. Thus, unlike academic research based on subjective coding, no empirical evidence exists to indicate that the coders' subjective assessments of these ads were accurate. This is a very serious flaw in the methodology of the study, especially since the crucial variables under consideration require highly subjective judgments. Furthermore, as I demonstrate below, the consequences of error can be enormous in the sense that any coder errors are reproduced in the data base for every single airing of the advertisement.

The problems of unreliability and invalidity are exacerbated when non-expert data collectors are employed to code the data, as in this study, which employed undergraduate students at Arizona State University. At present, few details about the ways these students were recruited and trained have been reported, and we therefore do not know whether the students were competent to make the subjective judgments required by the coding process. (As noted, we do know that the students were *not* subject to any formal training process.) With student coders, however, it is highly probable that implicit coding rules evolved to aid the coders in making these subjective assessments of the advertisements.

Moreover, coding these advertisements is often simply difficult, irrespective of one's training and experience. For instance, the attributes of an individual airing are specific as to time. Consider a hypothetical ad depicting Senator Ted Kennedy. That ad, if shown in Boston in a year in which Senator Kennedy is a candidate for re-election, would be characterized as depicting a candidate for office. If shown two years later (when Kennedy is not a candidate for re-election),

the identical ad should not be scored as depicting a candidate for office.[17] One must bring considerable knowledge to bear in making the various assessments of these ads. Thus, the lack of information about the reliability and validity of the coding is a very serious shortcoming of the 1998 study. In the absence of evidence that the coding of the ads is reliable and valid, one must approach these data with great caution.

Analysis of the 1998 Data

For the purposes of this report, I have relied on what purports to be the "final" version of the 1998 data set.[18] This file was provided to me as an SPSS[19] file named: 'Final_1998_Cmag.sav'.[20] I understand that this file was produced by Professor Goldstein for the purposes of this litigation. Upon reviewing this data set, I discovered that the original raw data files — the files produced by CMAG for the airings and the data file Professor Goldstein produced from coding the ads themselves — had already been integrated (matched and merged), and the data had been documented (with labels, to the limited extent to which the file is

---

[17]*Buying Time 1998* provides no information concerning how information about who was and who was not a candidate in any given airing of an ad was made available to the student coders.

[18]I put the word final in quotations marks because, as I understand the practices of Professor Goldstein, the data base is continuously subject to change and revision (as I have noted earlier).

[19]SPSS is statistical software that is widely used to analyze data of this sort.

[20]Frankly, the statistical analysis reported in *Buying Time 1998* is not at all transparent. In my analysis in this report, I have tried to provide sufficient details so that other analysts can reproduce my results. For the not-too-statistically inclined, the detail may seem excessive, even if it is recognized as necessary. For that, I apologize in advance.

documented) in a single file. Though this process of merging data is fraught with difficulty and potential for error, I have not been given any material that would allow me to judge the accuracy of the data merging. Indeed, one sees telltale signs in the data set that this match/merge process was not executed flawlessly (e.g., the assignment of the same identification number to ads with vastly different titles).

The total data set has 307,208 airings. The distribution of sponsors of these ads is as follows:

	Number of Airings
missing, unknown/not coded[21]	4,745
Candidate	235,613
Party	44,924
Group	21,926
Total	307,208

Since my report addresses only group-sponsored issue advertisements, I confine all of my analysis to the 21, 926 airings coded in the data set as ads having been sponsored by groups.

Analysis of these data can be conducted using different "units of analysis." For instance, one approach would be to focus on the ads themselves. In the 1998 data base, there are over 100

[21]When no code is provided for an airing, the computer program inserts a "missing value" code. In SPSS, this code is printed in the output as a period (i.e., ".") and is referred to as a system-missing code (assigned by SPSS as missing because the program had no other way to interpret the input data). In this data set, missing data (of this and other types) are extremely poorly documented, leaving one with little ability to discern the exact cause of missing codes.

discrete ads.[22] Such an approach is generally misleading because different ads are aired at vastly different rates. For example, ad number 12 (AFL/HMO Said No) was aired 2,808 times, according to the data base, while ad number 3026 (SIERRA/World is) was aired only a single time. To treat these two ads as equivalent makes little sense if one is interested in potential influence on the electorate.

The second "unit of analysis," the one preferred by me and generally by the authors of *Buying Time 1998*, is the "airing."[23] This is the individual showing, which is of course specific as to time and place. For instance, precisely the same ad may be aired in more than one media market and at more than a single point in time.

To reiterate, in the 1998 "final" data set, there are 21,926 airings sponsored by groups. However, for purposes of this analysis I will focus on the 20,432 group-sponsored airings in the data base coded as issue ads.

---

[22]There are errors in the data base in how the ads are identified. There are two possible unique identifiers for each ad: the variables IDNUMBER and CUSTITLE. As it turns out, however, neither is unique. In most instances, the errors seem to be harmless, resulting from data entry mistakes. For example, I assume that the following two descriptions refer to the same ad:
AFJS/Stabeno and AFJS/Stabenow Tu
Four discrete identification numbers suffer from this sort of malady.

However, ads with the id number 3431 seem to have a more serious mistake. The two titles that appear with this id number are:
NRA/Sorvino your and NV/CFCD/Not Abov
It is difficult to reconcile how these two ad titles could both be given the ad identification number of 3431.

I recognize that an ad may be coded differently when it is broadcast in different media markets. The process by which the identification numbers are assigned to the ads and airings is nowhere discussed in *Buying Time 1998*. A proper unique identifier would be one that is a combination of the ad identification number and the specific context within which it was aired. Such an identifier does not exist in the data base.

[23]One might quite reasonably use other units for conducting useful analysis of this data base, as, for instance, in measuring ads in terms of the number of minutes they were aired.

I consider the airings to be the appropriate unit of analysis for a study of this sort, and in fact *Buying Time 1998* relies on this unit for most of its analysis. However, it is important to note at the outset that a single discrepancy in coding an ad can have considerable consequences for the analysis of the airings data base. Let me give an example.

In the 1998 data base, Question 6, assessing the "purpose" of the ad, reveals that a majority of the airings had as their "purpose" providing information or urging action. The variable is distributed as follows.

	Percentage of Cases
Missing, not coded	3.3
Provide information or urge action	58.4
Generate support/opposition for a candidate	37.2
Unclear	1.0
Not coded	0.1
Total	100.0 %
N	20,432

Let us assume for a moment that the coding of a single ad was erroneous; in particular, assume that ad #11 was coded as "promoting issues" (provide information or urge action) rather than "promoting a candidate" (generate support/opposition for a candidate).[24] If the coding of *only* this single ad changed, then the distribution of the "purpose" variable becomes:

---

[24]I do not here make the allegation that the coding of this ad was in error. I am merely using this example to illustrate the consequences for the analysis of the coders making a single mistake in coding one of the attributes of an advertisement.

	Percentage of Cases
Missing, not coded	3.3
Provide information or urge action	64.6
Generate support/opposition for a candidate	31.0
Unclear	1.0
Not coded	0.1
Total	100.0 %
N	20,432

By changing the coding of only a single *ad*, the percentage of *airings* promoting issues rises from 58.4 % to 64.6 %, or a change of over 6 percentage points. This example illustrates just how volatile this data base is, and how very small errors in coding can be amplified substantially (with significant implications for substantive conclusions). Obviously, the consequence of an error in coding the attributes of an individual advertisement varies according to the frequency with which that ad was aired. I provide this example merely to illustrate that for some ads, a single small error can have quite large consequences for the statistical results. This means the statistical findings of the study are not robust; they are instead highly sensitive to error.

General Comments on the Accuracy of the 1998 Data

The 1998 data provided by the Brennan Center/Professor Goldstein are riddled with internal inconsistencies and errors.[25] One important consequence flowing from this problem is

---

[25]The report also seems to have engaged in the practice of rounding all percentages less than 1.0 up to 1. Thus, according to my calculations, the 1 % reported in the first portion of Figure 1.3 on page 15 is actually 0.2 %. This rounding convention is not part of accepted social

that virtually none of the results in the 1998 report could be reproduced with any of the data sets made available to me.[26] This seriously undermines my confidence in the substantive conclusions drawn in the report.

The discrepancies are numerous. A few examples can illustrate the inconsistencies in the data set.

Number of media markets represented:

Buying Time 1998 (p. 78):	75
"Final" data set:	76

Number airings analyzed in the report:

Buying Time 1998 (p. 32):	302,860
"Final" data set:	307,208

Number of group-sponsored airings:

Buying Time 1998 (p. 32):	22,151
"Final" data set:	21,926

Many additional inconsistencies could be noted; because they are not documented in *Buying Time 1998* or any other material I have reviewed, one wonders how these discrepancies arose.[27]

---

science practice.

[26]Even more confusion is added if one tries to use the data available on the Buying Time web site to reproduce the results. See http://www.buyingtime.org/

[27]I should perhaps note that the technical difficulty of merging data sets drawn from different sources and collected under different circumstances is substantial, and likely requires someone with considerable experience with statistical and data management software.

Many variables in the data set are not properly documented; indeed, many variables have no documentation whatsoever. For instance, one of the most important variables in the data is that which describes the ad type. Two variables seem to represent this information: WHAT and WHAT4. The former is entirely undocumented while the latter's variable label in the SPSS data set read "four category ad type" and the values of the codes are listed as "1 cand camp" "2 coord exp" "3 ind exp" "4 issue ad". Thus, the implication seems to be that WHAT4 is the preferred depository of this information. However, when one tries to reproduce the figures in Figure 1.3 (for example) on page 15 of *Buying Time 1998*, the figures are more consonant with WHAT than with WHAT4, which, to say the least, is puzzling and confusing.

What is the difference in the coding of the two variables? Two differences occur. First, some of the airings coded as missing by WHAT4 in fact have valid values on WHAT. Second, a portion of the airings coded at 3 on WHAT are coded as "3 ind exp" on WHAT4, while another portion is coded as "4 issue ad." The distribution of the two variables is as follows.

	Frequencies on WHAT[28]	Frequencies on WHAT4
missing	97	195
cand camp	342	342
ind exp	1,055	1,281
issue ad	20,432	20,108
Total	21,926	21,926

Thus, there is a great deal of confusion here. The most important implication is simple: The data provided cannot be used to replicate the findings of *Buying Time 1998.*

For purposes of my own analysis, I will focus on group sponsored ads that are coded as issue ads on the variable WHAT. There are 20,432 such airings sponsored by groups. These 20,432 airings represent approximately 117 individual advertisements. Exhibit 10 to this report lists the titles of these ads and includes each storyboard made available to me.

The Characteristics of Group Sponsored Issue Ads Aired Within 60 Days of the Election, 1998

Of considerable importance in analyzing the *Buying Time 1998* data are group-sponsored ads having the following characteristics:

---

[28]The categories are not labeled on this variable so I have had to make some assumptions about the meaning of the individual codes. For instance, when I see that 20,108 airings appear in the data base in an unlabelled category, I make the assumption from the size of the group in comparison to the frequencies on the labeled variable that the meaning of this category is "issue ads." This is course an inference that may or may not be correct, but I have found no other source of documentation that would allow me to make more certain judgments about the meaning of the variables in the data set.

1. Aired within 60 days of the election

2. Do not contain "magic words"

3. Mention or picture one or more candidates

The final version of the 1998 data set provided by Professor Goldstein reflects a total of 6,640 group-sponsored airings have these characteristics.[29] Of the 6,640 airings with these attributes,

57.7 % were coded as promoting candidates

42.3 % were coded as promoting issues

In this "final" version of the data base, these 42.3 % airings stem from a single ad –

12. AFL/HMO Said No

The 57.7 % airings stem from 26 ads. The most common such ad is

11. AFL/CIOStrengthe

The full listing of the ads coded as "promoting candidates" follows:

ID	Title	Number of Airings
2	AAHP/Look Out Fo	77
7	AFJS/Stabeno	129
9	AFL/Chavez His L	11
10	AFLCIO/Call ES	17
11	AFLCIO/Strengthe	1265
15	AFL/SS Trust Fun	575
16	AFLT/Stand Up Fo	142

[29]Note that these airings represent a tiny fraction of all airings in 1998.

17	AFLT/OR People S	45
20	AFLT/Call WU	5
21	AFTL/KY Lucas	1
22	AFTL/WI Tell Rya	213
137	AVO/Ashcroft Val	55
1043	MI/MICA Bonior L	108
1045	MICF/Fieger Stab	167
1355	NJ/AFJS FPallone	72
1411	NPLA/Call Feingo	3
2406	WI/ LCWAF Neuman	71
2512	AFL/PBR Call Tod	10
2521	AFLT/Call WU Rev	39
2531	AFTL/KY Stop Wil	76
2532	AFTL/UT Cook Sig	35
2533	AFTL/Willams is	210
2637	IA/Bosewell Edu	129
2860	NM/DCCC Redman D	281
3024	SIERRA/NC Faircl	4
3431	NV/CFCD/Not Abov	92

The testimony in this case reflects that the authors of *Buying Time 1998* regard *two* ads as genuine issue ads aired within 60 days of the 1998 election: ad #12 and ad #318. See Seltz

Deposition, pages 100-102 (Exhibit 4); Holman Deposition, pages 50-53 (Exhibit 8). For reasons that are unclear to me, in this "final" version of the 1998 data set, ad #318 is coded as an issue ad, aired within 60 days of the election, but *not* as depicting a candidate.[30] On the basis of the depositions, I have accepted the author's judgment that the ad should be coded as depicting a candidate and have therefore treated it accordingly. As a result, the distribution of the "purpose" variable, based on 6,896 airings, then becomes:

55.6 % were coded as promoting candidates

44.4 % were coded as promoting issue

Of course, the source of the 44.4 % of the airings is two ads: #12 and #318.

It is important to take note of a highly significant attribute of the coding sheets for the 1998 ads. As reported on page 193 of *Buying Time 1998*, ads coded in Question 6 as "provide information or urge action" should *not* have been coded on *any of the variables indicating whether candidates were depicted or not.* When coders scored an ad as "providing information" on Question 6, the coding form says: "(If so, skip to question #19)." That is, every single ad that was coded in 1998 as providing information or urging action should have been coded as "not applicable" on questions 7 through 18, *since the coders were not asked to make a judgment about these questions*. Thus, if the coders performed as instructed by the coding sheet, there would never be in this data base an ad that is coded as both providing information and depicting a candidate, since that information is captured in Questions 7 and 8. Thus, if one is interested in ads that provide information while depicting candidates, this data set, by the design of the coding

[30]Note that in some versions of the 1998 data base made available to me by the Brennan Center this ad is in fact coded as depicting candidates. I have no information about how, why, and when the codes for this ad were altered.

form, can say absolutely nothing about these ads. To the extent that the data set includes information about the depiction of candidates for these ads, such information must stem from either (a) coders who ignored the instructions on the coding form, or (b) changes made to the data base after the students coded the data. It is unclear why an investigator interested in ads that both provided information and depicted candidates would design such a coding form. More important, the confusion in the instructions regarding Questions 7 through 18 may have introduced a degree of bias into how the students coded Question 6 by suggesting that any advertisement that included the name of a candidate should be coded as having a purpose of promoting or opposing a candidate.

The data base includes 7,606 group-sponsored airings coded as having a "purpose" of generating support or opposition to a candidate. My analysis reveals that *fully 97.7 % of these airings were also coded as mentioning candidates.* The most important conclusion I draw from this analysis is that mentioning a candidate and promoting a candidate are virtually the same thing, as these data were coded by the undergraduate students (and/or Professor Goldstein). It seems highly likely to me that the student coders coded these three questions (6, 7, and 8) virtually simultaneously: A candidate (or what the coder thought was a candidate) was observed in the ad, and then Question 6 was coded as electioneering (in part because the coders knew that the presence of a candidate was not coded if Question 6 was coded as providing information), and then the student made the determination of whether the candidate was "the favored candidate" (Question 7) or the "favored candidate's opponent" (Question 8). Thus, the entire relationship — empirical and logical — between Questions 6 and Questions 7, and 8 renders the data set of little utility for answering important questions about these ads and airings.

The Focus of the Ads

As I have noted, *Buying Time 1998* places considerable emphasis on ads with three characteristics: those aired within 60 days of the election, which mention one candidate or the other, and which do not use so-called magic words.[31] As noted, this group represents 6,896 airings.

The authors of *Buying Time 1998* apparently believe it appropriate to consider prohibiting or regulating ads depicting candidates aired within 60 days of the election if in fact the content of these ads is nothing more than what they characterize as electioneering (the promotion of candidates, not issues). As I have noted, according to the coding in Q6, the bulk of these ads (55.6 %) was judged to be promoting candidates. I will refer to these ads as possible "candidate-promoting issue ads," by which I mean that they appear to be issue ads (sponsored by groups, not using "magic words"), but that they were coded by the undergraduate students as promoting candidates rather than promoting issues. It is instructive to consider these ads in greater detail.

The undergraduate students were also asked to code Question 22. It is useful to reproduce the exact content of Question 22 (from page 194 of the Report):

> 22. In your judgement, is the primary focus of this ad on the personal characteristics of either candidate or on policy matters?
>
> 1. Personal characteristics
> 2. Policy matters
> 3. Both

---

[31]Of course, were the ads to use so-called magic words, they would be characterized as express advocacy under *Buckley v. Valeo*, 424 U.S. 1 (1976).

4. Neither

This variable is also given special attention in the Report on page 196 (see the variable "Focus"). One might predict that these possible "candidate promoting issue ads" ads would *not* be coded as primarily addressing policy matters; instead, they should be overwhelmingly coded as "personal characteristics" or "neither" personal characteristics nor policy.

In fact, that expectation is *not* supported by the data; instead, exactly the contrary conclusion is revealed by the data analysis. The coding on Question 22 for the 6,896 airings aired within 60 days of the election, the results are:

Question 22 Code	Percentage
missing	0.1
Personal characteristics	1.2
Policy matters	98.1
Both	0.7
Neither	0.0
Total	100.0 %
N	6,896

Thus, *virtually all* of the ads aired within 60 days of the election and depicting a candidate were coded as having a "**primary focus**" on policy matters. Thus, virtually of these ads have clear and obvious policy content. This is a very important conclusion that bears considerable emphasis:

More than 98 % of the group-sponsored ads broadcast within 60 days of the 1998

election and depicting one candidate or the other were coded by the coders (or Professor Goldstein) as having policy matters as their "primary focus."

How should this result be understood within the context of the coding of Question 6? It is obvious that many ads were coded in Question 6 as promoting candidates but also as being "primarily" focused on policy matters in Question 22. Is it reasonable that coders would have coded the ads on these variables in this fashion?

I believe so, since Questions 6 and 22 differ in several obvious and extremely important respects. Question 6 requires that the undergraduate coder make a rigid distinction between two "purposes" of the ad. The specific question answered by the student is (from page 193, emphasis in the original):

> 6. In your opinion, is the purpose of this ad to provide information about or urge action on a bill or issue, or is it to generate support or opposition for a **particular candidate**?
>
> 1. Provide information or urge action
>
> (If so, skip to Question #19)
>
> 2. Generate support/opposition for candidate
>
> 3. Unsure/unclear

Note that in Question 6, the coders are not given the option of responding that the ad *both* provides information or urges action *and* generates support or opposition for a particular candidate. In the event of mixed content, the undergraduate coders were forced to make a

dichotomous judgment about the ad's "purpose."[32] Question 6 does not ask the coder to discern the "primary" purpose of the ad. Indeed, the question provides no guidance whatsoever as to how to code mixed-content ads.

In contrast, the construction and coding of Question 22 is an improvement over Question 6 in a number of respects. First, the question allows the coders options of "3. Both" and "4. Neither". Thus, the problem of forcing a choice between different parts of the manifest content of the ad is resolved by allowing a coding of "mixed" content. One wonders why such a strategy was not used on Question 6, especially given its use by Professor Goldstein in Question 22. Second, the question provides at least some guidance for how to make the judgment required, telling the coder to consider the "primary focus" of the ad. This is still a difficult coding decision for the student coders — and there is no evidence that coders were trained or given formal coding rules for how to resolve ambiguity and ambivalence — but Question 22 is structured in such a way as to provide more reliable information than Question 6.

It is also revealing to note some of the characteristics of the ads coded as "candidate-promoting issue ads." According to the coders, the 6,896 airings in 1998 have the following characteristics:

95.6 % of the airings urged the viewer to take some action

74.3 % of these were coded as providing a toll-free telephone number in the ad

Another 18.4 % of the airings provided a toll number

4.4 % were coded as missing data

---

[32]As I have discussed earlier, it is impossible for the coders to have known anything about the ad-maker's"purpose" or intention in broadcasting the ad, since such information was entirely unavailable to the coders.

And only 2.9 % were coded as providing no telephone number

45.7 % were coded as addressing health care issues

30.1 % were said to address the issue of taxes, and

27.8 % addressed the Social Security issue.

And, of course, as already mentioned

98.1 % were coded as having policy matters as their primary focus.

Thus, these ads certainly have quite a number of characteristics of what the authors of *Buying Time 1998* refer to as "genuine issue ads."

Thus, several conclusions emerge from analysis of these data:

1) The coding in Question 6 is deeply flawed.

2) When Question 6 and Question 22 clash (i.e., the coded attributes differ), the coding of Question 22 should be considered more valid and reliable.

3) According to the coding, the vast and overwhelming majority of ads said to be examples of illegitimate electioneering (by virtue of promoting candidates) in fact were judged by their own coders to have "policy matters" as their "primary focus."[33]

The 7 % Figure in *Buying Time 1998* — Deconstructing Figure 4.22 and Page 109

For the purposes of my analysis, perhaps the most relevant section of *Buying Time 1998*

---

[33]It is perhaps obvious but nonetheless important to note that in many if not most instances, the person who is a candidate for public office is also an *office-holder*. Thus, whatever else they might be, many of these ads are direct efforts to influence the actions of sitting elected representatives.

is Figure 4.22 (page 110) and the text purporting to interpret this figure (page 109).[34] It is crucial to understand the analysis that underlies this text (and the depositions have spent a great deal of time on this text and figure). With only a single paragraph of *Buying Time 1998* devoted to this crucial issue (the second paragraph on page 109), one can surmise that the Report left a great deal unsaid.

The conclusion drawn on page 109 is as follows:

> The results show that while 41 percent of issue ads that provide information or urge action appeared within 60 days of the fall election, just 7 percent of those ads (consisting of just two spots) appeared within 60 days and referred to a candidate.[35]

In his deposition, Mr. Seltz affirmed that, even now, he stands behind this conclusion as accurate (page 77).

Focusing on the 20,432 group-sponsored airings in the 1998 data base, I find that 38.5 % of the airings promoting issues were aired within 60 days of the election. Thus, a significant majority of the airings (61.3 %) were aired earlier than two months before the election.[36]

The authors of *Buying Time 1998* also claim that:

---

[34]Note that Figure 4.21, page 108, has a fairly obvious error in it. The figure for 30 second ads in September is reported as 97%. In fact, the figure seems to be 87 %. This can be readily seen because the point on the line depicted at 97 % is in fact lower (has a smaller score on vertical axis) than the adjacent point of 91 % (for the first half of October).

[35]Note that in the terminology that has emerged, and in the terminology of this report, the "7 percent of those ads" phrase in the quoted text should be understood as "7 percent of those airings" and the "consisting of just two spots" actually means "consisting of just two ads."

[36]These percentages do not total to 100 % because the data base includes 16 airings (0.1 %) for which the date of the airing is unknown.

.... just 7 percent of those ads (consisting of just two spots) appeared within 60 days and referred to a candidate.

The "those ads" phrase clearly refers to "issue ads" – the 11,939 airings in the data base. The question as the authors have framed it then becomes: Of the 11,939 airings, what percentage have *both* the characteristics of having been aired within 60 days of the election *and* referring to a candidate?

According to the version of the data base designated as the "final" version, the answer to this question is 25.7 % (3,064 airings divided by 11,939 airings = 25.7 %). Because this figure is so widely contested — and because so many different percentages have been proposed as being correct — it is useful to examine the data with great care. The following table does so:

Group-Sponsored Airings Having as Their "Purpose" Providing Information or Urging Action (N = 11,939)

	Date of Airing		
Was a candidate was mentioned or pictured in the ad?	More than 60 days before the election	60 days or closer to the election	Missing
Yes	0.0 % (0)	25.7 % (3,064)	0.0 % (0)
No	61.3 % (7,324)	12.9 % (1,535)	0.1 % (16)
Column total – Number of airings	7,324	4,599	16

Note: Total number of airings = 11,939. Cell entries are percentages of the total (11,939) and the number of airings in the cell. Thus, the 3,064 airings aired within 60 days of the election and depicting a candidate represent 25.7 % of the 11,939 group-sponsored issue airings in the data base.

This table is based on a denominator of 11,939, which represents *all* issue ads aired in 1998. All percentages in the table reflect that denominator (the numbers shown in parentheses are the actual numbers of airings). Thus, the first cell in the table indicates that 0 % of the 11,939 issue airings (which is 0 actual airings) were shown more than 60 days prior to the election and depicted a candidate. The next figure in the table, 25.7 %, indicates that of the 11,939 issue airings in the data (the total 1998 issue airings), 25.7 % (or 3,064 actual airings) mentioned a candidate and were shown within 60 days of the election. Thus, according to this data base, were one to assume that all ads shown within 60 days of the election and depicting a candidate were electioneering ads, one would be wrong 25.7 % of the time. Or to put this in terms of actual airings, one would make 3,064 mistakes by assuming these airings to be electioneering ads.

Of course, using a denominator of all issue ads broadcast in 1998 for these calculations is arbitrary and makes little sense. Why use January 1, 1998, as the starting date for the total pool of issue ads (i.e., the denominator)? Why not include ads from December 1997, or even the entire election cycle beginning in November 1996? Why not limit the denominator to ads shown in the last half of 1998? The authors of *Buying Time 1998* selected a denominator for their calculations that has no theoretical meaning. Consequently, their choice is arbitrary. Unfortunately, their conclusions depend heavily on this choice.

What choice for a denominator would not be arbitrary? The pool of cases one might want to examine would be those aired within 60 days of the election. Suppose the question is:

> I will assume that all ads aired within 60 days of an election and which depict a candidate for public office have a particular characteristic (i.e., they are engaged

in electioneering). What percentage of the time would this assumption result in an error in the sense that the assumed characteristic is not the same as the actual characteristic?

This question demands that one use a denominator of all airings shown within 60 days of an election and which depict a candidate. This is not an arbitrary choice of denominators. Instead, this is the only reasonable denominator that answers this question.[37]

The denominator chosen by the authors of *Buying Time 1998* relates to an entirely different, and virtually meaningless question: What percentage of the total ads run throughout the year that mentioned a candidate by name and were coded as providing information or urging action appeared *within* 60 days of the election, rather than *earlier than* 60 days before the election? Since, as noted above, the coders coded virtually all of the advertisements that mentioned a candidate by name as promoting a candidate, the percentage of such ads must be very small. However, this tells us nothing about the error rate over the course of the entire year or even in the 60 day period.

Using the 1998 data, I have run the analysis necessary to answer the more meaningful question. There are 6,896 airings in the denominator. To repeat, these are airings that were shown within 60 days of the election and which depicted a candidate (and, of course, which did not use so-called magic words). Of these airings 44.4 % were coded as intended to promote issues. Thus, to answer the question posed above, *the assumption that these airings are electioneering ads*

---

[37]As an analogy, when we ask: What percentage of jury trials falsely convict the innocent, we use as a denominator those defendants who are at risk (those who go to trial) and do not include those defendants who have their charges dismissed prior to trial (because they are not at risk at a trial).

would be wrong (empirically disproved) for 44.4 % of the airings (3064 actual airings).

There is quite some difference between a figure of 7 %, reported in *Buying Time 1998*, and my calculations of 44.4 %. How could there be such a large discrepancy?

The discrepancy is due in part to the use of different denominators (although using a conceptually equivalent denominator — all issue ads shown in 1998 — my estimate is 25.7 %, not 7 %). But the discrepancy is also very much a function of the after-the-fact recoding of airings by Professor Goldstein. As it turns out, the student coders may not have done a valid job of coding the questions about whether candidates were depicted because they did not have the information about where the ad was broadcast. Thus, Questions 6, 7, and 8 could not be reliably and validly answered by the coders.

Through various unpublished and undocumented machinations, it appears that only a portion of the airings of ad #12 and #318 were eventually counted as "genuine issue ads" by the authors of the 1998 report These are airings depicting a political figure that aired in areas in which that political figure was not a candidate for office. There is some confusion about the exact locations in which such ads were broadcast (e.g., it seems that Pittsburgh was often referred to as Philadelphia).[38] But for the correct figure to be 7 %, as claimed by *Buying Time 1998*, the numerator must be approximately 836 airings. That is,

836 / 11,939 = 7 %

---

[38]I believe this confusion stems mainly from the fact that neither Professor Goldstein nor the authors of *Buying Time 1998* nor CMAG has storyboards for each unique ad. Since the CMAG software cannot distinguish between ads that are generally the same but that substitute the name of different political figures in different airings (so-called cookie cutter ads), CMAG treats the airings as identical and does not capture a storyboard for each unique ad. Consequently, it cannot be known with certainty whether a particular ad mentions a candidate or not.

Without conceding the numerator (because the codings of airings have changed so frequently and are not necessarily represented in the "final" version of the data set), it is important to understand exactly what this fraction depicts. The numerator is limited to ads appearing within 60 days of the election. The denominator refers to *all airings* of issue ads shown in 1998. As I noted above, however, a considerable proportion of these ads represented in the denominator were broadcast well before the election season even began. I can see no justification for making the denominator equal to all issue ads aired in 1998.

Returning to the data base, let us reconsider the Report's finding using issue ads aired within 60 days of the election. There are 3,064 issue airings shown within 60 days of the election and depicting candidates. If we assume that just 836 of these are valid (and this assumption should not be accepted as an accurate depiction of reality), then the percentage becomes:

836 / 3,064 = 27.3 %

It is important *not* to accept the numerator of this quantity as 836 since it represents so many data manipulations after the fact (i.e., after the student coding). But if we assume that the authors of *Buying Time 1998* have reduced this figure to its smallest possible quantity, then the proper conclusion from their calculations, as measured against this data base, is that at least 27 percent of the ads would be improperly restricted by a rule banning or regulating ads depicting candidates within 60 days of the election. From my calculations above, this figure may range to as high as approximately 45 % of the airings in 1998.

From page 109 and the accompanying data in *Buying Time 1998*, as well as from the data base itself, one can surmise that the report considers about 5,064 issue airings as having been

shown within 60 days of the election.[39] With this denominator one can deduce the percentage of ads aired within 60 days of the election, depicting candidates, without using "magic words," but which had as their "purpose" providing information or urging action. If we accept the Buying Time numerator of 836 (and I repeat that I do not accept that figure), then according to the Brennan Center calculations, this percentage would be

836 / 5,064 = 16.5 %

Thus, even accepting the flawed numerator and using the Brennan Center's own figures for calculating the proper denominator (airings within 60 days), 16.5 % of the group ads were "genuine issue ads" (as defined by the Brennan Center) that would be, in the words of the Brennan Center authors, "unfairly caught by" application of the three criteria that I understand were adopted by the Bipartisan Campaign Reform Act (BCRA).[40]

The primary reason why I do not accept the numerator in this quantity is that it represents the post-coding manipulations of the variables by Professor Goldstein. If we return to the eight ads that the students determined were intended to provide information before they were overruled by Professor Goldstein (see my discussion of these ads above), and undo Professor Goldstein's changes by reverting back to the original coding of the students as revealed in the actual coding sheets, we find that these eight ads were aired a total of 2,405 times in the last 60 days of the

---

[39]This estimate is the quantity: .408 • 12,411. The 12,411 is number of issue ads in the data base (irrespective of how the ad type is coded). The .408 is my calculation from the data base of the percentage of such ads that were aired within 60 days of the 1998 election. Note that the figure .408 is entirely compatible with the claim on page 109 of *Buying Time 1998* that "41 percent of issue ads that provide information or urge action appeared within 60 days of the fall election."

[40]See Holman Deposition, page 89, line 15, through page 91, line 17.

1998 election. According to the data, none of these ads used "magic words" and all mentioned candidates. Therefore, they are appropriate to add to the numerator of the above quantity, as in:

(836 + 2,405) / 5,064 = 64.0 %

Thus, by this calculation, nearly two-thirds of the group ads that aired within 60 days of 1998 election were coded by the students as "genuine issue ads"; all of these (again in the words of the Brennan Center authors) would be "unfairly caught" by application of the criteria now set forth in BCRA.

CONCLUSIONS — *BUYING TIME 1998*

1. Careful examination of the Brennan Center/Professor Goldstein 1998 data provides little confidence that the data were collected and assembled accurately and reliably. For instance, the data provided for this analysis cannot reproduce the findings reported in the 1998 Report, and several obvious errors exist in the data base.

2. The methodology of coding the advertisement is deeply flawed.

3. The practice of engaging in idiosyncratic, standardless, and highly subjective post-hoc alteration of the data base by partisans severely undermines the credibility of the data set and the conclusions drawn therefrom.

4. The 1998 data provided by the Brennan Center/Professor Goldstein not only do not support the conclusions announced in the Report, but they in fact lead to exactly the opposite conclusion: The vast majority of ads broadcast within 60 days of the 1998 election, which depicted one candidate or the other, had as their primary focus issue advocacy.

Buying Time 2000

Summary of Conclusions about *Buying Time 2000*

- The data collected for *Buying Time 2000* are based on an improved methodology in which the authors clearly attempted to correct the considerable deficiencies of *Buying Time 1998*. For example, it appears that, in contrast to the 1998 study, at least some systematic effort was made to assess the reliability of the 2000 data. Many methodological flaws were not corrected, however, rendering the 2000 data base susceptible to many if not most of the same criticisms that have been applied to the 1998 data base.

- No evidence is adduced in the 2000 project that

  — the coding of the key variables is any less subjective than in the earlier study

  — the process of post-hoc data alteration has been abandoned or constrained

  — the biases so pervasive in *Buying Time 1998* have been made subservient to proper scientific procedures.

- Consequently, it should not be surprising that none of the multiple data bases provided by the Brennan Center and Professor Goldstein allow the figures and findings reported in *Buying Time 2000* to be reproduced.

- I adduce evidence that the *Buying Time 2000* data base produces such wildly divergent estimates of the number of airings with certain characteristics

(such as issue ads that are aired within 60 days of the election and which depict a candidate), that the data cannot be used to provide useful conclusions about the characteristics of such ads and airings.

- As with the 1998 data, the most reliable finding that emerges from my analysis of the 2000 data is that most of the ads coded as promoting candidates in fact had "policy matters" as their "primary focus."

General Caveats and Reservations about *Buying Time 2000*

Many of the problems I have already identified in *Buying Time 1998* also plague the 2000 edition of the data base, so I will mention them only in passing:

- This is not a peer reviewed study and has not been, and, in my professional judgment based on being a peer reviewer for publishers, journals, and research foundations, could not be, published as an academic report. Since *Buying Time 2000* has not been through a process of peer review, the study has not been subject to the close scrutiny of objective reviewers and its methodology has not been judged to be acceptable by the social scientific community.

- The authors appear to be proponents of a particular policy point-of-view rather than disinterested scientists.

- One of the authors of this report (McLoughlin) appears to have little if any training in the methods of quantitative analysis and the skill level of the other

author (Holman), when it comes to analyzing data bases of this sort, is unclear.[41]

- The data base has numerous errors and inconsistencies in it.

- The same ailments affecting the CMAG data base in 1998 continue to affect it in 2000, as for example the presence of missing data and the inability to capture each unique ad in its own storyboard.

- The analysis relies heavily on extremely subjective coding by students (at the University of Wisconsin), and the student coders were not trained in how to assess the ads.

- The data set is subject to continuous alteration by Professor Goldstein, in consultation with Brennan Center staff. No documents have been produced that indicate how Professor Goldstein has or should exercise his enormous discretionary powers to change or recode the data. On the basis of the data sets I have been provided by Professor Goldstein and the Brennan Center, I believe it likely that the data set as it exists today does not bear a strong relationship to the data coded by the students.

In sum, I conclude that one should not place any confidence in the data set or the conclusions drawn therefrom in *Buying Time 2000.*

---

[41]There is the intimation in an e-mail on 09 Feb 2001 from Craig Holman to Ken Goldstein that Goldstein accused Holman of not understanding the database. Holman's e-mail reply to Goldstein defends against the charges laid by Goldstein. See 09 Feb 2001 E-mail from Craig Holman to Ken Goldstein (KG 00012853). A copy fo the E-mail is attached as Exhibit 11 to this report.

The Data Provided for This Study

For the purposes of this report, I analyzed the following data set: cmag_2000_labeled_data_may02_feb01_w_compet.sav .[42] This data set contains information on 970,428 airings. In general, the 2000 data are cleaner (fewer undocumented codes, internal consistencies, etc.) than the 1998 data, although (as documented later in this report), numerous errors still plague the data set.[43] Moreover, the processes by which errors in the data were identified and corrected are suspect and worrisome.[44] In the final analysis, one cannot have much confidence in the validity and reliability of this version of the data base.

Replicating the Findings of *Buying Time 2000*

The numerous changes in the data base mean that many specific findings from Buying

---

[42]The Brennan Center/Professor Goldstein provided multiple versions of the 2000 data base, including a file named: Final CMAG data - merged.sav . Despite its name, that file seems to be an earlier iteration of the 2000 data base. The file I am analyzing for this report has a creation date of 8/1/2002. I understand that it was produced by Professor Goldstein in this litigation. I have not examined the "new" 2000 data set included in Professor Goldstein's expert report of September 23, 2002, except to establish that it differs from the data set previously produced.

[43]In an e-mail on 09 Feb 2001 from Craig Holman to Ken Goldstein, Holman identifies a significant error in a version of the data set transmitted from Goldstein to Holman. The error has to do with the transformation of a crucial variable. Holman asserts: "Q8New has been offered as a correction to Q8, but it is really a correction to Q11, which is fine. However, the corrected version of Q11 by Q8New provides radically different results from the 1998 database." See 09 Feb 2001 E-mail from Craig Holman to Ken Goldstein (KG 00012853). This is one example of the sloppiness of the data base. A copy of this E-mail is attached as Exhibit 11 to this report.

[44]A variety of data correction processes were apparently used, but none is detailed in the report, and at least some of the description of what happened raises serious methodological questions (e.g., missing values were determined by "extrapolating from other data" — *Buying Time 2000*, 20).

Time 2000 cannot be replicated. This stems in part from a basic discrepancy in the number of airings considered to be group sponsored. According to my analysis, there are 133,956 such airings. According to the table reported on page 72 of *Buying Time 2000*, however, there are at least 142,421 group-sponsored airings. I simply do not understand this discrepancy. Without basic agreement on the number of ads and airings, it is impossible to reproduce the findings of that report. Indeed, given the additional issue of *post hoc* alterations applied to the data base, it is unclear that the authors of the study themselves could stand by any given figure reported in *Buying Time 2000.*

Group Sponsored Ads in 2000

My focus in this report is on group-sponsored advertisements. Two variables in the data set apparently address the identity of the ad sponsor: SPONSOR1 and SPONSOR. Inconsistent codings exist in the data base on 125,084 airings, with the variable SPONSOR reporting these as "1. Candidate Sponsored Airings Including Primary" but SPONSOR1 coding them as "6. Unclear Primary Sponsor Airings." This discrepancy, however, does not seem to implicate group-sponsored airings and therefore can be ignored for present purposes.

For the purposes of this analysis, I focus on the 133,956 airings that are represented in the data base as being sponsored by interest groups in federal elections.[45] These airings were identified by a code of "1" on a variable named IGDUM. According to the ad identification

---

[45]These airings are identified in the data base by a code of "1. IG sponsored ads" on the variable IGDUM. Using the variable FEDERAL, I exclude 132 airings coded as "non-federal races." The titles of these ads are: MDCLT/Glendening Bell Overcharges, SC/NA Hodges Broken Promises, WVBC/Governor Underwood. Note that I included the 54,661 airings coded (implicitly) as missing data on this variable (FEDERAL).

variable (ADCODE), there are approximately 339 unique ads in this airings data base. These ads are listed in Exhibit 12 to this report.[46]

The Reliability of the Coding

Unlike the 1998 data collection, it appears that the 2000 project made some serious attempt to assess the reliability of the data collected by the student coders. A great deal of information is still not available (e.g., how were the students recruited, to what degree had they been exposed to Professor Goldstein's courses and lectures), but *Buying Time 2000* does at least present some discussion of inter-coder reliability (e.g., *Buying Time 2000*, 19).

The most detailed explanation I have seen of the procedures undertaken to assess inter-coder reliability is presented not in *Buying Time 2000* but in a 2002 report by Krasno and Goldstein.[47] Here we are told that a sample of 250 ads was recoded, but no details are provided on how these ads were sampled (e.g., random versus non-random selection), who the coders were (e.g., expert versus novice), and variable-by-variable reliability results for the full ad coding data set are not presented. Before the inter-coder reliability methods can be accepted as producing any useful information, a great deal more information must be revealed.

In footnote 4 of this publication, Professors Krasno and Goldstein discuss the reliability of the coding of the "purpose" of the ad, asserting that the coders differed on the "purpose" of the ad in only a single instance. Although the report does not so indicate, the implication is that this

---

[46]All of the available storyboards for the ads listed are also included in Exhibit 12.

[47]Krasno, Jonathan, and Kenneth Goldstein. 2002. The Facts about Television Advertising and the McCain-Feingold Bill." *PS: Political Science and Politics* 35 (#2, June): 207-212, in footnote 4, page 211. See Exhibit 5.

variable was reliably recoded in 249 instances and disagreement appeared over only 1 single ad. The footnote also discusses the reliability of the question assessing the "tone" of the ad. No other variables are discussed. It appears from this footnote, however, that Professors Krasno and Goldstein place considerable confidence in the coding judgments of their student coders.

Another Data Set in Constant Flux

As I noted at the outset of this section, this data set, like the 1998 data, is continuously being changed. Specifically, I mean that the original codings assigned to the ads and airings by the students are altered with some regularity by Professor Goldstein.

Over the life of the data set, practices and policies in this regard have varied. For example, in an e-mail from Craig Holman (Brennan Center) to Ken Goldstein dated 03/10/2001, Dr. Holman makes the following assertion:

> At our last conference call, the staff decided that we are on more solid ground not substituting our own judgment of the ads for the undergrad coder's judgments, except in the case of clearly factual errors, so I am reconverting a couple of key variables back to the original codes on q11. [the "purpose" of the ad variable][48]

Dr. Holman refers to ADCODE 1367 as an example of this decision.

The data base also provides some information about the alterations of the original data set. Consider the coding of Question 11 (the so-called purpose of the ad). In the 8/01/2002 version of the data base, there is a variable named Q11. This, I initially assumed, is the original

---

[48]See 03/10/2001 E-Mail from Craig Holman to Ken Goldstein, Holman Deposition Exhibit #23. A copy of the E-mail is attached as Exhibit 13 to this report.

coding of the students for the crucially important indicator of the purpose of the ad. I assumed this in part because the variable is in sequence in the data set with the other variables derived from the original coding of the ads. However, three additional variables exist, named OLDQ11, Q11_1, and NEWQ11. The variable OLDQ11 has different numerical codes for the categories, but is identical to Q11_1. Similarly, Q11 and NEWQ11 are identical.[49]

My analysis suggests that the original coding of the ads on this attribute is most likely to be found in Q11_1.[50] I identify this variable because it contains the full range of codes, including unsure and missing data codes. The variable Q11_1 was apparently first transformed into the Q11 (which was then inserted in the sequential order that Q11_1 once occupied). One difference between the two variables is that Q11 has no missing data. How the problem of no information for the ads/airings was resolved is not apparent from the data set.

The data base contains 338 ads.[51] Of these, 276 are coded consistently throughout the various versions of this variable. That is 181 ads (and these are ads, not airings) were coded as designed to generate support or opposition to candidates and were not later altered by Professor

---

[49]There are actually a number of variables in the data set characterizing the ad. Though differences may exist outside the group-sponsored portion of the total data base, in the 133,956 airings considered in this report, the following variables are equivalent:

Q11 = NEWISSUE
Q11 = NEWELEC
Q11_1 = OLDISSUE
Q11_1 = NEWISSUE

[50]There is yet another representation of the ad's purpose in the data set in the variable ADCODE. The differences between this variable and Q11 are on four ads that are coded as electioneering on Q11 but as genuine issue ads on ADTYPE. Being unable to discern why these changes were made, I will use Q11 as the indicator of the ad's "purpose" throughout this report.

[51]Actually, there are 339 ads, but one ad is not consistently coded on Q11 so I excluded it from this analysis.

Goldstein. Similarly, 95 ads were coded by the students as advocating issue positions and these codes were then apparently accepted by Professor Goldstein.

The codes for the remaining 62 ads appear from the data base to have had their codes changed on the "purpose of the ad." These findings thus bring an entirely new perspective to the 2000 Buying Time data base. It is certainly true that the original ad data were coded by the student coders. But it is apparent that a substantial number of these codes — 62 ads out of a total of 338 — apparently were not the product of the student coders but were instead based on the judgment of Professor Goldstein (often with the consultation of the staff of the Brennan Center). Unfortunately, very little information has been reported regarding what limits, if any, were placed on Professor Goldstein's discretion in making these changes, information which is highly important for assessing the findings of the 2000 Report.

Let us return to the inter-coder reliability results reported for this variable by Professors Krasno and Goldstein. It now appears that changes in the codes for Question 11 were fairly common. One wonders therefore whether the inter-coder reliability results discussed in the report in *PS* reflect a comparison between the subsequent codings for the reliability test and (a) the original coding by the students or (b) the Goldstein-adjusted version of the variable. These data present a puzzle: Given that the student coding was deemed after the inter-coder reliability test to be so reliable for Question 11, why were so many changes subsequently made by Professor Goldstein in the codings of these variables? Were *reliable* student codings altered by Professor Goldstein? If so, on what basis and with what justification? Since it is difficult for me to imagine that one would conduct an inter-coder reliability test on data that had subsequently been altered by the investigator — such a test would obviously not provide evidence of the reliability of the

student coding — I can only imagine that Professor Goldstein decided to alter the data base insofar as Question 11 is concerned *even after drawing the conclusion that the coding of the ads on Question 11 was reliable.* Whatever specifically happened with these data, the conclusion I am compelled to draw is that the data as presented in the data set cmag_2000_labeled_data_may02_feb01_w_compet.sav are of highly questionable reliability.

Two important concerns arise from this examination of the volatility in the coding of the ads.

1) What motives were involved in making the various changes to the codes for these ads? Were the changes solely a function of scientific concern for accuracy in the data, or were they instead related to disappointment that the initial findings did not support the expectations and preferences of the investigators?[52]

2) What specific standards did Professor Goldstein employ in judging the ads? Apparently, no documents exist that formalize and make explicit his decision rules and heuristics. Indeed, it is unclear to me that a specific formal definition has ever been provided for an ad that has as its purpose the generation of "support or opposition for a candidate" or for an advertisement that "provides information or urges action." Further, how should an ad that undeniably does both be coded? The investigators have asserted numerous times that subjective

---

[52]One sees evidence throughout the e-mail exchanges of the investigators of a clear preference that the data support a particular conclusion. For instance, in an e-mail of 11 Jan 2001 from Dr. Holman to various NYU people, Dr. Holman (referring to the 1998 data) asserts that "the numbers crunched for Rick do not paint as great a picture as previously thought." After summarizing his view that the number of issue ads in 1998 that would be deemed "electioneering within a 60-day regulatory period" is more likely 40 % than the originally claimed 7 %, Dr. Holman asserts: "Hopefully, we will get more positive results from the 2000 database." A copy of this E-mail is attached as Exhibit 14 to this report.

judgments are involved in coding Question 11. Presumably, some specific decisional heuristics exist to guide these judgments. Unfortunately, none has been produced, insofar as I am aware.

Additional Considerations Regarding the Coding of Question 11

One of the most important variables in the 2000 study is Question 11. The student coders were asked to respond to the following query (*Buying Time 2000*, 99):

> 11. In your opinion, is the purpose of the ad to provide information about or urge action on a bill or issue, or is it to generate support or opposition for a particular candidate?
>
> 1. Generate support or opposition for a candidate
> 2. Provide information or urge action
> 3. Unsure/Unclear

Thus, the wording of the question in the 1998 and 2000 projects is identical.[53]

Consequently, the coding of this variable contains all of the frailties of the 1998 coding. Without repeating my conclusions from my review of the 1998 data, I will simply summarize my concerns about the data collected with this question:

- The coders are asked to make a judgment about the "purpose" of the ads from the manifest content of the ad. Of course, no data are available on the "purpose" of the ads' sponsors so the student coders are obviously therefore not coding anything about the intention of the ad sponsor.

---

[53]In 1998, the words "particular candidate" were printed in bold-face type; in *Buying Time 2000* the words were printed in ordinary type.

- The question is highly subjective, and it is not clear that any rules or guidelines were provided to the coders to guide the exercise of discretion. What constitutes "support" or "opposition" is nowhere detailed for the coders, requiring them to rely on their own subjective values in the exercise of their discretion in answering this question.

- The question is difficult to answer, requiring a great deal of information, including knowing whether someone depicted in the ad is in fact a candidate in the geographical area where the ad was aired.

- The question requires a forced choice of either providing information/urging action or generating support/opposition for a particular candidate, thus failing to recognize that many ads may have mixed content.

- In the event of mixed content ads, the question provides no instructions for how to weight the various parts of the ad in drawing a conclusion about its "purpose."

Thus, this analysis reinforces my deep and serious reservations about the quality of the data collected with this question.

The Issue Content of Candidate Support Ads

I have adduced in this report indirect logical and empirical evidence indicating that the student coders over-reacted to the appearance of a politician in these ads, often falsely coding

issue advocacy ads as electioneering ads.[54] But it is unnecessary to rely on indirect evidence or conjecture in investigating these ads; instead, direct evidence as to their issue content is available from the data base itself.

In 2000, the student coders were required to answer the following question:

27. In your judgement, is the primary focus of this ad on the personal characteristics of either candidate or on policy matters?

1. Policy matters
2. Personal characteristics
3. Both
4. Neither

The question is superior (i.e., more likely to generate valid data) to Question 11 in at least two important respects. First, it allows ads to be coded as having mixed content. Second, it provides guidance to the coders, however vague, by asking them to judge the "primary focus" of the ad. This renders the coding of mixed content ads more reliable and valid.

Recall that 79,648 of the group-sponsored airings (59.5 % of the total) were coded on Question 11 as having as their purpose the generation of support or opposition to a particular candidate.[55] The following table reports the coding on Question 27 of the 79,648 airings said to be electioneering ads in Question 11.

---

[54]As I have suggested earlier, many of the alterations in the data base were necessary because the student coders over-reacted to the presence of a political figure in an ad, often wrongly assuming that if a political figure appeared, he or she was a candidate for public office, and that the ad were therefore highly likely to be engaged in electioneering.

[55]For the purposes of this analysis, I use variable Q11, which I believe to be the Goldstein adjusted version of the original codings.

Coding on Question 27 – Primary Focus of the Ad	Airings Coded as Having the Purpose of *Generating Support or Opposition for a Candidate*	
	Number of Airings	Percentage
Policy Matters	62,764	[redacted]
Personal Characteristics	1,956	2.5
Both	14,035	[redacted]
Neither	379	0.5
Missing/not coded	514	0.6
Total	79,648	100.0 %

Thus — and this is a very important conclusion — 78.8 % of the ads coded as having "electioneering" purpose in Question 11 were simultaneously coded as having as their "primary focus" advocacy on "policy matters." Another 17.6 % of the airings had both policy and personal characteristics as their foci, *resulting in a total of 96.4 % of the group-sponsored ads in 2000 having policy matters as their primary focus, or at least as part of their primary focus.*

As with the 1998 data, it is obvious that the coding of Question 27 highlights the deficiencies of the coding of Question 11. Indeed, the results I produce are entirely logical given the artificial nature of the coding decision required by Question 11. Question 11 produces the results it produces because (a) the coders failed to code mixed content ads, since Professor Goldstein provided them no coding category with which they could record mixed-content ads, and (b) they also over-reacted to the presence of politicians in the ads when they coded Question 11. Thus, their coding on Question 11 vastly underestimated the number of ads with policy

content. When given a chance on Question 27 to acknowledge that ads might have mixed content, the coders overwhelmingly judged that these ads focused primarily on issues. To reiterate, due to the defects in the structure of Question 11, and the much sounder structure of the question asked in Question 27, the results of Question 27 should be accepted because the coding is both more reliable and more valid.

Exactly what is the issue content of these ads? In Questions 32 - 35, the student coders identified the specific policy issues addressed in the ads.[56] Fully 37.8 % of the ads coded as electioneering addressed issues of "Health Care" and another 25.8 % considered "Medicare."[57] A panoply of additional issues is addressed in these ads, clearly attesting to their issue content.

The Characteristics of Group-sponsored Issue Ads Aired Within 60 Days of the Election, 2000

Two variables become relevant when we consider ads and airings within 60 days of the 2000 election[58]: (1) The presence of "magic words" — the indicator of the presence of "magic words" is derived from the responses to Question 8[59], and (2) whether a candidate is mentioned

---

[56]There is an apparent error in some of the coding, since the undocumented code "9" appears in 514 of the airings.

[57]Note that "Health Care" and "Medicare" are independent and separate categories in the data base.

[58]There are 225 airings with a missing value on the variable SPOTDATE (date on which the airing was aired). It is not clear from the available documentation how these missing values were corrected to produce scores on the variable DAY60.

[59]Note that, according to Dr. Holman, this is not an objective variable; instead, the coding of this variable is subjective. Holman Deposition, page 72, lines 4-6. See Exhibit 8.

in the advertisement. Because it appears that this information was subject to considerable *post hoc* coding by Professor Goldstein, I merely accept for the purposes of this analysis the coding in the data set for the variable MENTION[60]:

Some 60,000 airings in 2000 satisfy these three criteria (aired within 60 days of the election, no "magic words," and a candidate for office was apparently mentioned or depicted). According to the data base (and without accepting the results), in the overwhelming majority of these airings — indeed, 59,997 out of 60,006 — the "purpose" of the ad was to generate support or opposition for a candidate.

On the other hand, consider how these ads were coded on Question 27:

Q27	Number of Airings	Percentage
Primary focus on policy matters	47,356	78.9
Primary focus on personal characteristics	541	0.9
Primary focus on both	11,884	19.8
Primary focus on neither	225	0.4
missing/not coded	0	0.0
Total	60,006	100.0 %

Thus, the overwhelming conclusion is that, of the airings with these three characteristics, 78.9 % were coded by the students or Professor Goldstein as having policy matters as their "primary focus" and another 19.8 % were said to have *both* policy matters and personal characteristics as

---

[60]There is some inconsistency between this variable and the coding on Questions 12 and 13. The causes of this inconsistency are not apparent to me, suggesting that it is due to Professor Goldstein's adjustments on the data base.

their primary foci. *Thus, 98.7 % of these ads in fact have a clear policy focus.*

An example of the coding of one of the ads might be useful to document how this coding works in actual practice. Consider ad #1544, which reads:

> [Woman]: "Many people who have cancer are looking for miracles. At this point, it's my faith, my support from my family and my friends. And then there's the medicine." [Announcer]: Congresswoman Anne Northup is working to add a prescription drug benefit to Medicare and make sure medicines are available for every senior who need them. Call Congresswoman Anne Northup. [Woman]: "Without the medicine, I would not be where I am." [PFB Citizens for Better Medicare Inc.]

Frame #7 reads:

> Call Cong. Northup
>
> Support her Rx Plan for Seniors
>
> (502) 582-5129
>
> PAID FOR BY CITIZENS FOR BETTER MEDICARE INC.

This ad was coded as promoting a candidate (although this is itself a highly questionable assessment). But the ad was also coded as having policy matters as its primary focus, and was also properly coded as to the policy substance ("Medicare"). Surely the coders judged the "purpose" of this ad to be candidate promotion because they over-reacted to the mentioning of a candidate and because they were not allowed to code such ads as having mixed content. In this and many other instances, the coding of Question 27 regarding the primary focus of the ad is more reliable and valid than the coding of Question 11 (the "purpose" of the ad).

According to the data base, nearly all of these ads in 2000 were coded as "electioneering issue ads." But according to my analysis nearly all are primarily focused on issues. Which of the two conclusions should be accepted and believed?

To answer this question requires that we return to the relative value of the coding of Questions 11 and 27. I need not repeat the arguments nor reproduce the data analysis I presented above. The simple conclusion is that when the codings on Questions 11 and 27 seem to indicate different content in an ad, the results from Question 27 ought to be accepted due to the greater reliability and validity of the coding on Question 27.

Ads Aired Within 60 Days of the 2000 Election

Buying Time 2000 asserts: "Of all group-sponsored issue ads that depicted a candidate within 60 days of the election, 99.4 % were found to be electioneering issue ads (see Figure 8-2). In absolute numbers, *only three genuine issue ads (which aired a total of 331 times in the 2000 elections) would have been defined as electioneering communications under the Snowe-Jeffords amendment*" *(Buying Time 2000*, page 73, emphasis in the original). From a statistical viewpoint, this conclusion ought to be fairly easy to reproduce from the version of the 2000 data base with a creation date of 8/1/2002 (the data base I have been analyzing for this report). In order to avoid any confusion, I will focus on a single number in this claim: 331, the number of airings represented by ads with these characteristics.

According to the data base, if this question is addressed with the coding on Question 11, *there are 9 airings with these characteristics*. Not even the authors of *Buying Time 2000* believe that this is a correct estimate of the number of airings.

According to the coding of Q11_1 (which I believe to be the original, student coded version of the "purpose" question), there are 1,082 airings with these characteristics. Thus, there are at least three possible answers to the question of how many ads were aired within 60 days of the election which depicted candidates and which had as their purpose providing information:

9 airings

331 airings, or

1,082 airings.

I have no doubt whatsoever that many additional estimates could be produced from the *Buying Time 2000* data base. In fact, however, I have confidence in none of these.

Just as important as the accuracy or inaccuracy of the data cited in the 2000 Report, however, is the extent to which changes in a relatively small number of the highly subjective codings can affect the results reported and the conclusions reached. To make this point, let us consider 30 specific ads from the 2000 data base. These ads are:
676, 704, 1270, 1367, 1381, 1422, 1552, 1648, 1757, 1761, 1844 2065, 2088, 2089, 2107, 2158, 2301, 2344, 2502, 2558, 2588, 2695, 2712, 2810, 2900, 2935, 3158, 3170, 3220, 4001.[61]
Let us assume for the purposes of this analysis that each of these ads could fairly be coded as "providing information."

When I examine the number of airings of these ads within 60 days of the election[62], I discover that they were aired 14,133 times. The total number of genuine issue airings under this

---

[61]This list of 30 ads was provided to me by counsel. The storyboards for these 30 ads are attached as Exhibit 15 to this report.

[62]Note that none of these ads uses "magic words;" all mention candidates.

hypothesis is therefore the 331 acknowledged by the Brennan Center report and these 14,133, for a total of 14,464 airings. Since there were 60,006 issue airings (without "magic words") within 60 days of the 2000 election, this figure of 14,464 constitutes 24.1 % of these airings. Thus, if one assumed that airings presented within 60 days of the 2000 election, which mentioned candidates, but which did not mention "magic words," were intended to promote candidates (or were, in the words of the Brennan Center authors, "electioneering ads"), one would be wrong, under this scenario, approximately 24 % of the time.

Several conclusions should be drawn from this analysis. First, changes in the coding of very small numbers of ads can change the results dramatically. Second, the current version of the 2000 data base supports many possible estimates of the number of ads with these characteristics. And finally, given all of the deficiencies of the data base and the coding on which it is based, the wisest course is to draw no conclusions whatsoever about these ads on the basis of the empirical evidence in the data base.

Summary and Conclusions

One important conclusion from this analysis is that there is a great deal of subjectivity, ambiguity, and ambivalence in the coding of the content of these advertisements. The student coders often disagree with Professor Goldstein, Professor Goldstein occasionally disagrees with the staff of the Brennan Center, and indeed, Professor Goldstein himself seems to change his mind over time about the nature of some of these ads. Whether an ad is issue based or candidate focused is not an objective, clearly discernible attribute of the ads. Reasonable people disagree about the content of these ads.

Consequently, this data base is constantly in flux. The figures reported in *Buying Time 2000* have changed with alterations and manipulations of the data by Professor Goldstein and it is entirely possible they will change again in the future. Since no reports have been written about this process of data manipulation, there has been no transparency in this process. Since this research has not been subject to the peer review process — through which such procedures would have to be revealed, defended, and accepted by the academic community — one cannot help but worry quite a bit about the reliability and validity of the data collected. Without reliability and validity, no useful conclusions can be drawn from the data base.

This entire research process in *Buying Time 2000* thus deviates in many important ways from conventional scientific research. One cannot help but be impressed by the e-mail correspondence among the analysts. The conclusion one draws from a review of this e-mail is that the investigators were committed to drawing a particular set of substantive conclusions from the data. When the conclusions were not forthcoming, the data were scrutinized further and alterations were made in the data base. Indeed, at least with regard to some of the conclusions of the *Buying Time 1998*, internal disagreement among the investigators themselves about the veracity of the findings appears to have been substantial. Such strong *apriori* commitments to drawing certain substantive conclusions from the data base seriously undermines the credibility of the reports.

Overall Conclusions

The essence of science is replication. When scientists claimed to have discovered "cold fusion" excitement spread quickly throughout the word. That excitement dimmed and then was extinguished when it was discovered that crucial experiments could not be reproduced. Unless research procedures can be explained in such a way that others can repeat the procedures and produce the same results, then no value attaches to the information produced. We do not believe that cold fusion can be accomplished today because the initial claims of the investigators cannot be reproduced by others.

The Buying Time reports are not of the potential significance of cold fusion for the advancement of human welfare, but the reports do share the fatal flaw of relying on data that are not capable of generating reproducible findings and conclusions. For example, what proportion of issue ads in 1998 would have been improperly regulated or banned were a bright-line test imposed? *Buying Time 1998* says 7 percent; the depositions point to several figures between 13 and 40 percent; and my own analysis of the "final" 1998 data produces a figure of between 82 and 89 % (based on the coding in Question 22). Similar ambiguity and ambivalence characterize the 2000 data.

Why is there so much disagreement about the correct statistics? Part of the answer resides in factors such as the sloppiness of the data base and the fact that published reports did not undergo a review process prior to publication. But the larger portion of the answer has to do with the inherently subjective nature of the coding process. Certainly neither the coders nor Professor Goldstein could have accurately coded the "purpose" of any of the ads since no effort whatsoever was made to investigate the intent of the ad sponsors. The coding of the content of the ads is

inherently a subjective process — and proper scientific procedures for dealing with subjective phenomena were not followed — but subjectivity and coder discretion were exacerbated by failure to formally constrain and guide the exercise of discretion (including the exercise of discretion by Professor Goldstein) and by the inadequacies of the coding form itself (e.g., the failure to allow coders to acknowledge that ads typically have more than a single focus). And to allow coding decisions to be reversed on an idiosyncratic basis by those with a strong policy interest in the outcome of the research violates all canons of scientific objectivity. It is little wonder that the findings of *Buying Time 1998* and *Buying Time 2000* concerning so-called sham issue advocacy were not published in peer-reviewed social scientific journals or monographs.

After reviewing the reports, the data, and the depositions, I conclude that the findings of these reports should not be accepted as accurate and valid descriptions of the nature of political advertising in the 1998 and 2000 federal elections. The Buying Time reports have not taught us much about sham issue advocacy but they have been instructive on the nature of sham science.

Dr. James L. Gibson

September 30, 2002

Rebuttal to the Expert Reports of

Kenneth M. Goldstein and

Jonathan S. Krasno and Frank J. Sorauf

Dr. James L. Gibson

October 7, 2002

In my review of the original Buying Time reports (1998 and 2000), I drew several major conclusions about the reports, their conclusions, and the data upon which the studies were based. It is perhaps useful to begin this rebuttal with a brief restatement of these conclusions.

- As provided by Professor Goldstein, the 1998 and 2000 data bases are riddled with errors and inconsistencies and the findings of the reports cannot be replicated from the data.

- The key measures that are the focus of this litigation require highly subjective assessments and the student coders making those judgments never received any instruction on how to exercise their enormous discretion.

- The data bases are constantly being manipulated by Professor Goldstein, with many iterations of new scores being produced for the variables of primary interest in this litigation.

- Thus, the conclusions from those reports should not be accepted.

- Indeed, to the extent that the data bases are able to support any

conclusions, they are that the ads under consideration in this litigation have policy matters as their primary focus and therefore to ban or regulate them would be a substantial intrusion upon the freedom of interest groups to communicate with citizens and to attempt to shape public policy.

I will focus this rebuttal to the expert reports around the question of whether my conclusions are altered by the reports or the new data analysis that underlies them. I conclude that nothing in these reports has caused me to change the conclusions I expressed in my report of 9/30/2002.[1] Since nearly all of these issues are common to both the 1998 and 2000 studies, I will generally integrate my discussion of these studies in the analysis that follows.

Repudiating the Results from *Buying Time 1998* and *Buying Time 2000*

These new reports make plain that the authors and I agree on one basic point: The statistical analyses presented in *Buying Time 1998* and *Buying Time 2000* should not be accepted. Professor Goldstein now reports entirely new estimates of the effect of the Bipartisan Campaign Reform Act (BCRA) on "genuine issue speech,"[2] thereby repudiating the figures reported in *Buying Time 2000*, and Dr. Krasno and Professor Sorauf report new estimates for 1998 that differ from the numbers reported in *Buying Time 1998*.

The same conclusion applies to the data bases. Professor Goldstein has provided with his

---

[1]A listing of additional materials upon which I have relied is attached as Exhibit 1.

[2]To assist readers in comparing my analysis with that of Goldstein/Krasno/Sorauf, I adopt some of their language (e.g., "genuine issue ads"). The use of these terms should not be taken to represent agreement with either the conceptualization or operationalization of these concepts and issues, but is instead simply a convention I adopt so as to facilitate communication.

expert report yet another version of the 2000 data base, so presumably the earlier versions are defunct and should not be relied upon. Dr. Krasno and Professor Sorauf do not provide a newer version of the 1998 data base, but do offer calculations in the appendix that apparently render the old data base obsolete. I assume that one day these data bases will be archived and made available for use by other interested scholars, as most data bases of this sort are, and that at that point the authors will create clean, properly documented, and final versions of the data bases.[3] That day has not yet arrived.

As I have noted, the Buying Time data and studies are constantly evolving (with changes apparently being made in the codes of airings even in connection with this litigation), and with the conclusions of the authors representing targets that are always on the move. But at the same time, details about the methods employed by the studies are slowly beginning to emerge. Let us consider first the basic question of the quality of the data collected by CMAG.

Errors in the CMAG Data

The data collected by Campaign Media Analysis Group (CMAG) deserve closer scrutiny than they have received to date. Professor Goldstein professes confidence in these data[4], but considerable empirical evidence exists to suggest that this confidence is misplaced.

The CMAG data base should be scrutinized from several vantage points:

---

[3]While Appendix F of Professor Goldstein's report provides a document entitled "2000 Coding Sheet," this document provides information on only a tiny fraction of the variables included in the data set.

[4]See Ridout, Travis N., Michael Franz, Kenneth Goldstein, and Paul Freedman. 2002. "Measuring Exposure to Campaign Advertising." Unpublished paper, June 19, 2002, at page 17.

1) Were all of the airings represented in the data base in fact aired?

2) Does the data base represent all ads aired or instead were some airings not captured by the CMAG technology?

3) Did the CMAG "analysts" accurately code the content of the ads?

4) Was CMAG in fact able to capture storyboards for every unique ad?

Unfortunately, the answers to some of these questions cast very serious doubt on the CMAG data.

I have access to no independent information with which I can verify that the airings represented in the CMAG data base were actually broadcast, or that the information captured is in fact accurate.[5] However, Professor Goldstein has presented some evidence in his paper entitled "Measuring Exposure to Campaign Advertising" indicating that the accuracy of the information collected is certainly less than perfect (see Ridout et al. 2002).[6]

An equally serious issue is whether the data base fails to capture all ads that are broadcast. The available evidence suggests that there are some very serious omissions.

---

[5]In the appendix to their expert report, Dr. Krasno and Professor Sorauf claim that the CMAG data base is incorrect in the information it depicts about the ad with the id number 13. They provide no explanation for the error in the data base. Unfortunately, the appendix is not paginated, but this seems to be the 15th page of text in the appendix.

[6]Counsel for the AFL-CIO has provided me a copy of the Declaration of Denise Mitchell, 10/4/2002. In her Declaration, Mitchell asserts that she cannot reproduce from AFL-CIO records Professor Goldstein's calculations from the data base on the airings sponsored by the AFL-CIO. She notes that the CMAG data base makes two types of errors: errors of omission (failure to include ads aired by the AFL-CIO) and errors of commission (the inclusion of ads said to be aired by the AFL-CIO but that were in fact not aired). Since the AFL-CIO is the group with the second largest number of airings in the data base (well behind the Citizens for Better Medicare), any such failure to track accurately the airings of this organization would be a major indictment of the quality of the data base. See Declaration of Denise Mitchell, 10/4/2002, paragraph 64.

Professor Goldstein has reported that he compared the information in the CMAG data base with the actual billing records of eight television stations in five different markets (see Ridout et al. 2002). He discovered that the error rate of the CMAG data varied from 0 to 20 % (see Table 2, page 28).[7] From the figures reported in Table 2, we can derive an estimate of the overall degree of accuracy of the CMAG data base.

Note first that station billing invoices appear to document more ads than found in the data base in every one of the stations examined. Thus, the CMAG data always under-represent the number of airings by each of these television stations.

By how much does the data base under-represent actual broadcasts in 2000? We can derive an estimate quite simply. For instance, the 8,526 airings in the data base for WJRT apparently represent 99 percent[8] of the actual airings broadcast by WJRT. Applying the arithmetic, this means that WJRT probably broadcast 8,612 spots, and that the CMAG methodology therefore failed to capture 86 of these.[9] Applying this methodology to the figures for each of the stations, we discover that the CMAG data for these 8 stations likely missed 1,764 ads.[10] This represents a small percentage of the overall airings of these eight stations (5.04 %),

---

[7]Working through the example Professor Goldstein provides in the paper is useful for understanding the meaning of the error rate. The error rate for station WJRT is .01. This means that of the 618 airings for which the station has billing records, the CMAG data base includes only 612. Thus the error rate is .01: ((618 - 612) / 618) = .01.

[8]This figure is equal to 100 % minus the error rate.

[9] (8,526/8,612) = .99.

[10]Were I to use the more demanding standard of "Percent correct within 1 sec." (the first data column in the table), the number of ads estimated to be missing or that contain serious errors would be signficantly higher than the numbers I report here.

but when this proportion is applied to the data base as a whole, it translates into an estimate of 48,864 airings that in fact were broadcast but that were not captured by the CMAG methodology. This is a sizable number indeed. Moreover, the number is of sufficient magnitude that failure to include these airings in the analysis most likely biases the substantive conclusions drawn from the data base.

We do not know any of the characteristics of these 48,864 missing airings. But one might reasonably hypothesize that the clearer and more obvious the political content of an airing, the more likely it is to have been included by the CMAG analyst. Conversely, to the extent that ads did not have a clear "political purpose" that could be discerned by the CMAG analysts, the more likely they were to have been excluded from the data base by the analysts. Thus, it seems reasonable to assume that these 48,864 missing airings are not simply a random sub-set of the data, but instead tend to include a disproportionate number of "genuine issue ads."

Having addressed the "errors of omission" in the CMAG data base, I now consider the evidence on "errors of commission." In footnote 4 (page 17) of the Ridout et al. paper, the authors note an important anomaly in the CMAG data base: ".... some advertisements .... were never coded in the original data set [the student codings of the ads]. This is because these ads were not political, but were nonetheless included in the data provided by CMAG." This sentence raises a host of questions about the data base, including: What criteria did CMAG employ for selecting ads to be included in the data base? How was the content of the ads (e.g., "political" versus not) coded, and by whom, according to what criteria, and with what degree of reliability? Perhaps this problem with the CMAG data accounts for some basic puzzles and discrepancies in the data base.

Finally, I return to the inability of the CMAG technology to capture text and storyboards for each unique ad. Because the technology cannot distinguish airings that are very similar but not identical, so-called cookie-cutter ads are not properly represented in the data base. The failure to capture all airings of all unique ads is a very serious limitation of the data base.

In conclusion, the CMAG data are apparently plagued by numerous errors, making reliance on these data bases for drawing conclusions about the nature of political communications in the United States extremely precarious, hazardous, and subject to error.

The Evolution of Question 11

The second portion of the data bases used for these two reports was created by the student coders at Arizona State University and the University of Wisconsin. In the 2000 data set, Professor Goldstein provides yet another incarnation of the ad's scores on the all-important Question 11.[11] In the 2000 data set appended as Appendix L to his expert report, an entirely new variable can be found. This variable was not present in any of the earlier versions of the data base previously provided by either Professor Goldstein or the Brennan Center. The name and label of this variable are:

BT00_Q11 oldest available q11 codes (older than feb version)

This variable is of considerable interest given the importance of Question 11 to the findings and conclusions of Professor Goldstein's expert report. Where it came from, what it means (the

---

[11]I will not repeat all of the objections I laid out in my report of 9/30/2002, but I do incorporate those objections into this analysis since I have not changed my views on any of the issues addressed in that report.

individual response codes on the variable, ranging from 0 to 99, are unlabelled[12]), how it was match/merged to this data set, and indeed whether it is in fact the "oldest available" incarnation of Q11 all remain open questions. I assume that this variable gets closer to representing the original codes assigned to the ads by the student coders at the University of Wisconsin. I do not know if in fact some earlier version of this variable exists.[13]

Many issues surround the coding of Question 11 by the students and by Professor Goldstein. Since new information about this variable has come to light in Professor Goldstein's expert report, how this information about the ads was created deserves some additional scrutiny.

Inter-coder Reliability. In my report of 9/30/2002, I offered an analysis of both the reliability and validity of the information collected by the students. I surmised in that report that no formal test of inter-coder reliability was conducted in 1998. Nothing in the new expert reports leads me to a different conclusion about the 1998 data.

However, new information does indeed emerge about the 2000 data. In Appendix I to his expert report, Professor Goldstein offers a copy of a paper currently under review at *Political Communication*. In this paper, consideration of the reliability of the variables is presented, beginning on page 17 and extending through page 21. Here, the author reports an analysis of "the

---

[12]Data sets are typically documented with a "codebook" that details the meaning of each variable and each code on each variable in the data set. Moreover, to facilitate understanding and avoid analytical errors, variables are typically fully documented with labels. This variable is labeled as to its possible content (e.g., "oldest available q11 codes (older than feb version)"), but the individual scores or codes (what we would call the "response set") are not documented. No codebook was provided with this data set, and the norm in the data set is for the variables to be entirely undocumented.

[13] I have confirmed that this variable differs from the variables Q11 and Q11_1 as represented in the data set: cmag_2000_labeled_data_may02_feb01_w_compet.sav

accuracy of the human coding" (page 17), based on a randomly selected sample of "150 unique advertisements from the over 4,000 advertisements airing in 2000" (page 17).[14] Goldstein asserts: "As part of a larger project, we asked five undergraduate students to code the ads using our coding sheet, thus allowing us to compare our original coding with five independent assessments. We then compared each coder's results with those of the CMAG data base" (pages 17-18). Several questions about this process emerge.

One first wonders about how Professor Goldstein compared the new codings to the "original coding." Is the original coding represented in the variable "BT00_Q11 oldest available q11 codes (older than feb version)"? If so, why is this variable referred to as the "oldest available" incarnation of the coding rather than the "original version" of the variable? This a most troubling discrepancy, unless BT00_Q11 purports to be the scores assigned to the ads by the students and before the various post-coding manipulations and changes by Professor Goldstein.

This "test" of the accuracy of the coding relies upon 143 advertisements (since some of the 150 selected had to be discarded due to the discrepancies in the CMAG data base). In order to establish the relevance of this test to the issues in this litigation some considerable attention must be given to this sample of 143 airings.

Consider the 2000 data base and the coding of Question 11 in the version represented by BT00_Q11. Once four pesky airings are removed from the data base, 970,424 airings remain.[15]

---

[14]It is not clear how this sample of 150 ads relates to the 250 ads discussed in footnote 4 of the Krasno/Goldstein PS report (published in June 2002).

[15]Professor Goldstein asserts on page 8 of his report that "the CMAG database collected 970,424 political television ad broadcasts ....". Yet the data base he attaches as an appendix to his

When aggregated by the ad identification number, this produces 3,111 unique ads (or ads found to be unique by virtue of their ad identification number, to be absolutely precise).[16] Of these airings, 2 seemed to have been coded differently at different times (i.e., for different airings) on the "earliest known" Question 11. Furthermore, the student coders assigned what are apparently missing data scores to 19 of the ads.[17] Of the remaining ads, 2,953 were said to be "electioneering" ads (promoting support or opposition for a candidate) and only 127 ads were scored as providing information or urging actions ("genuine issue ads").

This distribution on the "purpose" variable has two very important consequences. First, in statistics, we call this a "skewed" distribution since almost all of the cases are scored in a particular way (and in the sense that deviation from a 50/50 split is extreme). When a variable is highly skewed, as this one, sampling cases on this variable becomes very difficult. A randomly selected sample of 100 ads from this "population" of 3,111 would be expected to find only 4 ads (actually 4.1 ads) that were coded as issue ads. A sample of 150 ads would be expected to generate 6 issue ads; and a sample of 250 would be expected to include no more than 10 issue ads. Thus, if one took a sample of 150 ads as Professor Goldstein asserts in his paper, one would find only a tiny number of ads that were coded as having a "purpose" of providing information or

---

report includes 970,428 airings. I removed the four airings with a "system missing" score on the variable named COUNT. This produces a data base with a number of airings consistent with the number identified in Professor Goldstein's report.

[16]It is unclear to me why Professor Goldstein refers to "4,000 advertisements airing in 2000." See Ridout, Travis N., Michael Franz, Kenneth Goldstein, and Paul Freedman. 2002. "Measuring Exposure to Campaign Advertising." Unpublished paper, June 19, 2002, at page 17.

[17]I say "apparently" since the codes "98" and "99" are conventionally used as "missing data" codes. However, the scores are undocumented in the data set.

urging action.[18] Since group-sponsored ads constitute less than 14 % of the data base from which the sample was apparently drawn, one would expect that 21 group ads were included in the inter-coder reliability data base (and of course that a tiny number of ads would be group-sponsored issue ads).

Consequently, any conclusions about whether this sort of ad was in fact reliably coded cannot be accepted on the basis of an examination of such a small number of ads. Perhaps this reliability test provides some information about candidate and party ads, but it provides virtually no useful information about group-sponsored ads and especially the degree to which group-sponsored ads were reliably and validly coded as to the "purpose" of the ad.[19]

A second consequence of this distribution is that, for whatever reasons, the overwhelming decision of the coders was that these ads represent electioneering. Issue ads were not quite the equivalent of needles in a haystack, but they were exceedingly rare. Thus, it seems quite likely that even after coding only a few ads, the coders developed a strong expectation, implicit or explicit, that the next ad they coded would be an electioneering ad. It is very difficult to make subjective assessments of infrequently occurring events. Once a coder discerns a pattern in the responses to a subjective variable, it becomes difficult indeed for the coder to "break the habit."

---

[18]The situation is little improved with a sample of 250 ads, the number said by Dr. Krasno and Professor Goldstein in their PS report to be subject to inter-coder reliability checks.

[19]The solution to this problem is stratified random sampling. That is, the proper methodology is one in which the original ads are segregated by type on this variable — "strata" — and then sampling takes place within stratum. Of course, with so few ads coded as issue ads in the first place, one would most likely have to adopt a strategy of recoding all ads coded as having an issue purpose and a random sample of those ads coded as having an electioneering purpose. There is no indication in any of the Buying Time documents, published or unpublished, that this was the strategy employed in assessing the reliability of this variable.

Especially in a circumstance in which there is an ideologically charged bias in favor of finding a particular outcome, "true" issue ads had a very heavy burden to overcome among the coders.

Substantive Consideration of the Results From Question 11

As I have noted, both Professor Goldstein and Dr. Krasno/Professor Sorauf have provided in their expert reports new estimates of the percentage of "genuine" ads/airings that mention or depict candidates and that were aired within the last 60 days of an election. Notwithstanding the flaws in their data bases, it is important that we give their reports close scrutiny.

Professor Goldstein offers Table 7 (immediately following page 24) as his calculations on the 2000 data base. He concludes:

> .... there is persuasive evidence that the BCRA definition of Electioneering Communication is not overbroad in the sense of applying to any meaningful number of advertisements that are perceived as Genuine Issue Ads. Using the coder's classifications, only 3.1 percent of the 45,001 Genuine Issue Ads broadcast by interest groups during the 2000 election cycle would have been covered by BCRA. (Pp. 26-27)

The denominator he uses for these calculations — 45,001 — represents the number of "genuine issue ads" broadcast in 2000. In the use of such a denominator, Professor Goldstein is at odds with both the authors of *Buying Time 2000* and with me. But in order to make some progress, let us put aside the denominator issue for a moment and focus on the numerator.

In his expert report, Professor Goldstein analyzes only 6 ads as "genuine issue ads" that would be prohibited under the criteria adopted in BCRA. These ads are:

627 KY/CFAW Call Northrup

1389 FAIR/IA Latham Foreign Worker Bill

2862 UT/COC Matheson Cant Decide Rx

1269 CBM/Rx Plan for Seniors 60

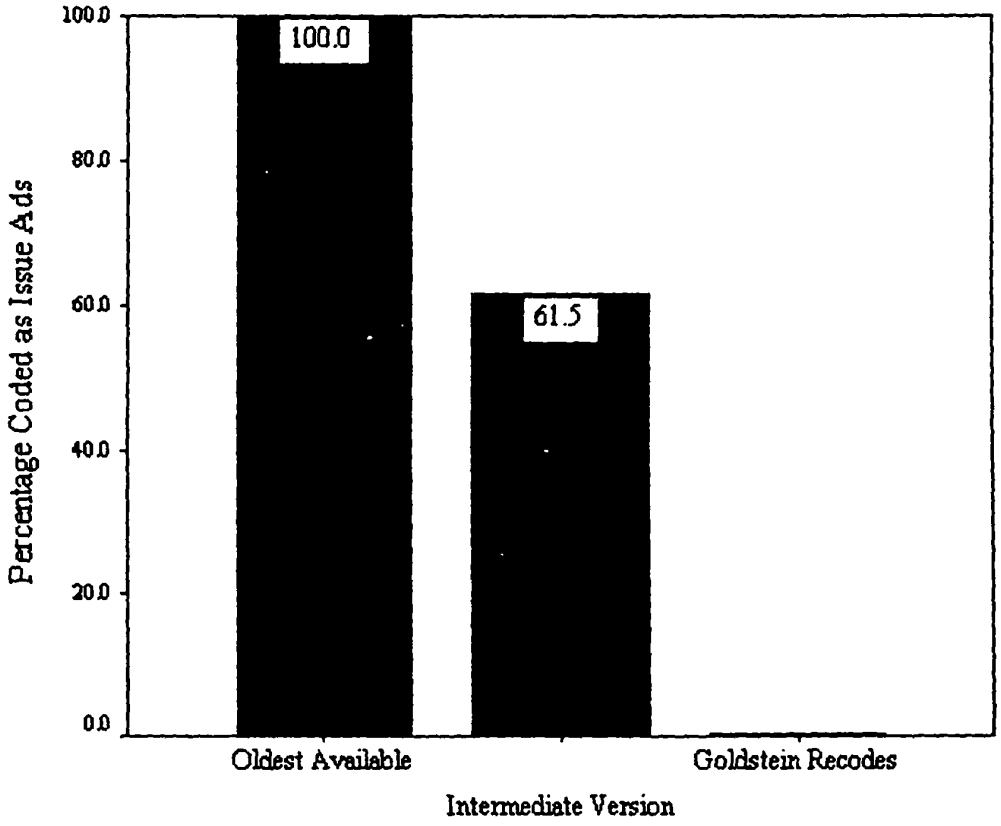
1367 RI/RIWV Langevin Abortion

2107 WI/NPLA Feingold Kohl Abortion 60

These six advertisements were aired 2,123 times in 2000. It is instructive to look at the results of the three codings on Question 11 as they have evolved over the course of the different manipulations of the data set. Figure 1 reports the transformation of these variables as represented in the most recent version of the 2000 data provided by Professor Goldstein as an appendix to his expert report.

The figure makes plain that these ads/airings were transformed from being coded entirely as issue ads to being almost entirely represented as electioneering ads! I recognize that Professor Goldstein, in his expert report, has converted these six ads yet again, changing them back to issue ads as originally coded by the students. But the treatment of these ads demonstrates two important points: Coding these ads is highly subjective and reasonable people disagree about their content, and Professor Goldstein seems to have claimed the authority to manipulate the data base at will.

Figure 1. The Transformation of the Coding of Question 11



Note: The percentage for the Goldstein transformation of the variable is 0.4 %.

Number of airings = 2,123.

The three variables represented here are:

BT00_Q11 Q11_1 Q11

As noted above, Professor Goldstein identifies 6 ads in 2000 that would be adversely and unfairly affected because they are "genuine issue ads," broadcast in the last 60 days of the election, and mentioning candidates. Two of these ads (#1367 and #2107) are included in the 30 ads I focused on in my discussion of the 2000 data (see pages 62-63 in my report of 9/30/2002). It is instructive to revisit this analysis based on the new claim by Professor Goldstein that these six ads should now be identified as "genuine issue ads" in the 2000 data base.

Putting these two samples together, we find that these 34 ads (Professor Goldstein's six plus my non-duplicative ads) were aired 15,347 times during the last 60 days of the 2000 election, according to the data base. Using the denominator from my 9/30/2002 report, these airings therefore represent 25.6 % of the 60,006 issue airings (without "magic words") within 60 days of the 2000 election and which depicted or mentioned candidates. What this demonstrates again is that Professor Goldstein's estimates of the impact of the three criteria are wholly dependent on the subjective assessments of the "purpose" of individual ads, assessments that are reasonably subject to debate.

As I noted in my report of 9/30/2002, a more reliable variable exists characterizing the content of the ads. The students were asked to code Question 27:

> 27. In your judgement, is the primary focus of this ad on the personal characteristics of either candidate or on policy matters?
>
> 1. Policy matters
> 2. Personal characteristics
> 3. Both
> 4. Neither

In my opening report of 9/30/2002, I set forth the view that the data produced by this question are of superior reliability and validity.

The coding on Question 27 is more reliable in part because it does not seem to have been subject to the post-coding manipulations inflicted on Question 11. If we consider once more the ads with the three important characteristics, we discover that the "primary focus" of 51,144 airings was on policy matters. Applying the Goldstein denominator:

51,144 / 60,623 = 84.4 %

This means that any policy that banned ads with these three characteristics would in fact pertain to political communications that overwhelmingly (84.4 %) had policy matters as their "primary" focus. Or to restate Professor Goldstein's conclusions 7 and 8 (page 3):

7. An extremely large percentage (84.4 %) of the political television ads aired during 2000 that would have been covered by BCRA were perceived by project coders as having policy matters as their primary focus.

8. BCRA requirements would have applied to 42.9 % of ads broadcast by interest groups during 2000 which were perceived by the coders as having policy matters as their "primary focus."[20]

Comparing the Characteristics of Ads

In Table 9 of his expert report, Professor Goldstein reports data apparently designed to demonstrate that a) "Interest Group Ads with No Candidate" differ from "Interest Group Ads

---

[20]This figure represents 48,370 divided by 112,844: ads with policy matters as their primary focus, mentioning candidates, and airing within the last 60 days of the election.

Mentioning Candidate 60 Days Before General Election," and b) that the latter share some characteristics with "Ads run by Candidates" and "Political Party Ads." The first part of this analysis is indeed interesting, although to the extent that the two types of issue ads do *not* exhibit different characteristics, the second part of the inquiry is not particularly revealing or relevant. Therefore, I set out to test the hypothesis that the two types of issue ads identified by Professor Goldstein in his analysis in Table 9 of his expert report are cut from different cloth.

In column 3 of this table, Professor Goldstein reports data concerning "Interest Group Ads with No Candidate" and in column 4, data for "Interest Group Ads Mentioning Candidate 60 Days Before General Election." The new data set provided as Appendix L to his expert report allows me to approximate his analysis, but also to consider a broader range of characteristics of the airings.[21] Professor Goldstein concludes that these two types of ads are fundamentally different in character. Unfortunately, his analysis fails to report the many ways in which these two groups of ads are similar.

In Table 1, I report my analysis of the characteristics of these two types of ads. The table demonstrates that these ads are similar in a number of important respects. First, large majorities of both types of ads urge the viewer to take some action; for a majority of these ads, that action involves writing, calling, or telling someone to do something, or sending a message or calling someone to express yourself. Approximately none of either type of ad asks the viewer to vote for someone. A majority of both types of ads give a telephone number for viewers to call. Virtually

---

[21]I cannot replicate the findings in Professor Goldstein's Table 7. Although I can identify in the new data set 55,648 ads sponsored by groups and not depicting a candiate, 2.4 % of these (not 0.0 %) in fact mention "magic words." Moreover, I find from the data set that the size of the second group of ads is 60,087, not 60,623, although my analysis does indicate that 0.0 % of these 60,087 airings mention "magic words."

all of both types of ads (93.7 % and 98.6%) have policy matters as either their primary focus or have a primary focus on both policy matters and personal characteristics. Interest group ads with no candidate mentioned or depicted tend to stress issues of health care more often than interest group ads mentioning a candidate and airing within 60 days of a general election, but the reverse is true in terms of the issue of Medicare. The most important conclusion I draw from analysis is that these two types of ads have quite similar attributes. This is not surprising since virtually all of them focus on policy matters.

Table 1. Comparing Group-Sponsored Ads in the 2000 Election

Ad Characteristic	Interest Group Ads with No Candidate (N = 55,648)	Interest Group Ads Mentioning Candidate 60 Days Before General Election (N = 60,087)
Does the ad urge action?		
% urging some action	72.9 %	65.8 %
Type of actions urged		
% to vote for someone	1.2 %	0.0 %
% to write, call, or tell someone to do something, or to send a message or call someone to express yourself	54.1 %	53.0 %
Information Provided		
% giving telephone number (toll-free or toll)	76.6 %	56.8 %
Primary Focus of the Ad		
% with a primary focus on policy matters	93.3 %	78.8 %
% with either a primary focus on policy matters or a primary focus on both policy matters and personal characteristics	93.7 %	98.6 %
Issue Content		
% addressing Health Care	28.8 %	39.2 %
% addressing Medicare	34.9 %	24.3 %

Summary of Conclusions about Professor Goldstein's Expert Report

Professor Goldstein's expert report continues to focus on the highly subjective coding of Question 11, ignoring his own evidence on the policy focus of the ads. As I have demonstrated in this rebuttal, ads that have policy matters as their primary focus predominate in the data set in general, as well as comprising a large majority of those ads depicting candidates and broadcast in the last 60 days of the 2000 election (and not mentioning "magic words"). Indeed, the key distinction Professor Goldstein seeks to make between ads with these characteristics and so-called genuine issue ads is not supported by the data, especially insofar as the policy foci of these ads is concerned.

The 1998 Conclusions

> In my report of 9/30/2002, I concluded the following about the 1998 data:
>
> Thus, by this calculation, nearly two-thirds [64.0 %] of the group ads that aired within 60 days of [the] 1998 election were coded by the students as "genuine issue ads"; all of these (again in the words of the Brennan Center authors) would be "unfairly caught" by application of the criteria now set forth in BCRA.

It is necessary to determine whether the report of Dr. Krasno and Professor Sorauf affects these calculations.[22]

My calculations were based on two groups of ads. First, I accepted the 2 ads that the

---

[22]For reasons I outline above (and in my report of 9/30/2002), I do not accept the method of calculation used in *Buying Time 1998*, and therefore I will not address the portion of the Krasno/Sorauf appendix entitled "SPREADSHEET: 1998 Formula" (pages 13-15, pagination I assigned to the pages). Instead, I will consider their calculations as reported in "SPREADSHEET: 2000 Formula" (pages 16-21, my pagination).

authors of *Buying Time 1998* treated as "genuine issue ads." To that list of 2, Dr. Krasno and Professor Sorauf have now added a third ad (named CENT/Breaux, aired two times — see page 60 and page 20 of the appendix). Second, I accepted the original coding by the students of 8 ads. These are all ads that Professor Goldstein transformed from issue ads to electioneering ads after the coders had completed their job. Were I to re-do my calculations on the basis of the 2 newly designated CENT/Breaux airings, the figures would change insignificantly (since so few airings are involved).

Dr. Krasno and Professor Sorauf make several other important "corrections" to the original 1998 calculations (see pages 3-4 of the appendix) and ultimately derive a numerator of 713 ads. The authors refer to these as "all pure issue ads (as rated by the coders) by groups that appeared within 60 days of the election and mentioned a candidate by name" (see Appendix page 3). Since they acknowledge in their description of these ads the importance of coder ratings — and incorporating the justification I presented in my report of 9/30/2002 — I revert to the original student codings of the 8 ads I identified in my report, and (accepting their identification of 713 airings only for the sake of argument), calculate a new percentage as follows:

(713 + 2,405) / 5,064 = 61.6 %

Again for the sake of argument, let me accept their adjusted 60-day denominator, and the percentage becomes

(713 +2,405) / 4,847 = 64.3 %

Thus, the substantive conclusions I drew in my report of 9/30/2002 remain unfazed by this new analysis.

Since my report was completed, new information has become available about the airings

of these 8 ads. In particular, the "Index of AFL-CIO Issue Advertising, 1995-2001" has been produced as an Exhibit to Ms. Mitchell's declaration. This document allows me to address the problem of so-called cookie cutter ads. Two of the eight ads appear to be cookie-cutter ads (#11 and #15). According to the data base, these ads were broadcast in several different locations throughout the country. Unfortunately, however, the CMAG technology does not capture the text or images of such ads, so one cannot confirm from the storyboards that a candidate was actually mentioned (because storyboards do not exist, the name of the individual mentioned or depicted in each ad is not known and must be inferred). In their expert report, Dr. Krasno and Professor Sorauf have offered a methodology for addressing this problem (see their Appendix).

Following the methodology of Dr. Krasno and Professor Sorauf, I have reexamined the airings of these 8 ads. In 8 instances, we do not have any verifiable information as to whether a candidate was identified. These are:

Ad number	Market Location
2	Washington DC
1411	Greenville
11	Albuquerque
11	Birmingham
11	Kansas City
15	Albuquerque
15	Grand Rapids
15	Kansas City

For the remainder of the airings of #11 and #15, the data base is compatible with the information

presented in the AFL-CIO document, and therefore I accept the data base's coding of these ads as mentioning candidates.

With this adjustment to the data, the figures are:

(836 + 1,736) / 5,064 = 50.8 %

Thus, using the methods adopted by Dr. Krasno and Professor Sorauf in the appendix to their expert report, about one-half of the airings in 1998 would be, in the words of the authors, "unfairly caught" by application of the BCRA criteria, if one were to accept (as Professor Goldstein does in his analysis of the 2000 data) the coders' determinations of the purposes of the ads prior to Professor Goldstein's manipulation of the data.

For the sake of argument, let me produce one additional calculation. Let me accept the new claim of Dr. Krasno and Professor Sorauf that their portion of the numerator is 713 airings. Let me accept their denominator as 4,847. And let me accept my new calculations on the 8 student-coded ads. The new percentage therefore becomes:

(713 + 1,736) / 4,847 = 50.5 %

Perhaps the only other adjustment that could be made to this figure would be to declare, as Dr. Krasno and Professor Sorauf do, contradictory evidence as missing data and exclude the airings from the calculations. Such a strategy would only serve to reduce the denominator in the above quantity, thereby increasing the percentage to some figure above 50.5 %.[23] To be clear, this analysis in effect assumes that none of the ads airing in the markets for which we have no

---

[23]For instance, the data base reports that ad #2 was aired in the Washington DC market, although the storyboard mentions North Carolina Senator Faircloth. According to the Dr. Krasno and Professor Sorauf methodology, these airings could be treated as missing data and therefore excluded from both the numerator and the denominator for purposes of these calculations.

independently verifiable information actually identified a candidate in that market. Consequently, this 50.5 % figure represents the statistical floor under the Krasno/Sorauf methodology; the 64 % figure cited in my report of 9/30/2002 (calculated from the data base itself) provides the ceiling.

The dispute over this percentage reflects two fundamental truths about this data set. First, the coding of the ads is highly subjective. It is difficult if not impossible to determine what the "purpose" of an ad is, and reasonable people will disagree even about categorizing the manifest content of an ad. Since errors in coding a single ad are amplified by the number of showings of that ad, the conclusions one draws from the data base are highly volatile and sensitive to errors.

Second, this data base has been subject to numerous post-coding manipulations. Not only does this undermine all of the conclusions and calculations reported in *Buying Time 1998*, but it also strongly undermines the claim that the student coding has much to do with the figures in the "final" analysis.[24] The figures Dr. Krasno and Professor Sorauf produce are no longer a function of the student coding, but instead reflect a series of ever-changing manipulations and recodings of the data base. Such a methodology sabotages any claim the study might make to scientific accuracy.

Groups Communicating With Citizens

Missing from the entire discussion of ads and airings in the expert reports submitted is any consideration of the people who consume these ads. After all, the purpose of all ads is to communicate with citizens. Fortunately, the Buying Time data sets provide some information

---

[24]Indeed, I should note that the report by Dr. Krasno and Professor Sorauf may not be the final analysis, since there may ultimately be some means of resolving various missing data problems that afflict their calculations.

that allows us some purchase on this important question.

The Buying Time data bases include a variable representing the "Gross Rating Points" of each airing. As described on page 6 of *Buying Time 1998:* "Gross ratings points are the sum of ratings for a particular time: if the local news is watched by ten percent of viewers with televisions, an ad run during the program represents ten gross ratings points." These points are of course idiosyncratic to each individual airing, and were apparently added to the data set by CMAG.

According to Dr. Krasno and Professor Sorauf, there were 713 airings in 1998 that the authors treat as "genuine issue ads," depicting a candidate and broadcast within 60 days of the election. These ads and the markets in which they were shown are listed on page 20 of their appendix. According to the data base, no Gross Ratings Points estimate is available for 6 of these airings.

Using the 1998-1999 Neilsen estimates of the number of households with television sets in the seven markets in which these ads were broadcast[25], and using the Gross Ratings Points to estimate the number of households viewing each airing, I find that these 707 airings represent communications with a staggering number of household — 30,108,857. Thus, were these ads (just the ads accepted by Dr. Krasno and Professor Goldstein as "genuine issue ads") prohibited, over 30 million group-citizen political communications would be affected (and this figure is based on the quite conservative assumption that each household only has a single person viewing television).

---

[25]A copy of these estimates is attached as Exhibit 2.

Why Would Interest Groups Attempt to Persuade Citizens on Policy Matters During the Election Season?

Professor Goldstein asserts in his expert report that a reasonable interest group would not air its issue ads during an electoral period. His argument is based on a couple of suppositions:

1) the ads are more expensive at this time of year

2) citizens are overwhelmed with information and will not pay attention to the ads, and

3) partisan attachments harden as the election draws near.

He cites the work of William McGuire and John Zaller. In fact the theory he cites argues in favor of exactly the opposite conclusion.

Issue advertising is an exercise in persuading citizens. Political Psychologists have long known (McGuire) that to persuade someone involves two steps. First, one must get the attention of the person one is attempting to persuade. Second, one must overcome the strength of existing attitudes if the attempt at persuasive communication is to result in attitude change. Unfortunately, these stages in the process are related to each other.

As it turns out, those with strong attitudes tend to pay attention to political communications while those with weak political attitudes tend to ignore them. This results in a paradox. Those most easily reached are least easily changed; those most easily changed are those most difficult to reach. This paradox has long bedeviled those researching processes of persuasion.

Those with strong political attitudes pay attention to politics year round, in part because the reason why they have strong attitudes in the first place is because political affairs are

interesting to them. Those with weak attitudes, on the other hand, only pay attention to political affairs under the most extreme circumstances (indeed, perhaps only when there is little choice but to pay attention because the airwaves are saturated with political information). Thus, election periods open a window of opportunity to reach those citizens who are most likely to be influenced by policy advertising. During the election, those with weak and potentially alterable attitudes are paying attention to political affairs. It therefore makes sense that an interest group would attempt to influence people when in fact they are most "influencable." Just as one does not attempt to sell Christmas trees in July (because people will screen out such attempts and will not pay attention to them), interest groups will try to take advantage of the window of opportunity opened by the electoral season. Consequently, it is entirely reasonable that an interest group would attempt to reach citizens on public policy issues during any given election period.

The last point Professor Goldstein makes is that partisan attachments harden. By this logic, candidates should abandon advertising as the election approaches since these hardened attitudes are difficult to convert. In fact, that does not happen, since, as the election approaches, candidates try to reach an even greater percentage of marginal voters, who have little interest in politics, and relatively pliable issue views.

This point about susceptibility to persuasion within 60 days of an election turns out to be intimately related to the controversy concerning the most appropriate denominator to use in our calculations of the proportion of "genuine issue ads" affected by BCRA. Those who would use a denominator based on all issue ads aired in the year are implicitly making the assumption that ads aired anytime throughout the year are equally as valuable as ads aired in proximity to the election. As I have indicated, this is not so. The damage of prohibiting an ad within 60 days of an

election cannot be ameliorated by allowing that ad to be broadcast at some other point throughout the year.

Thus, I conclude that there is nothing peculiar or nefarious about interest groups seeking to influence citizens on policy matters during the last 60 days of an election period.

Overall Summary of This Rebuttal

The various reports based on the 1998 and 2000 data are built upon a house of cards. As more information about the methodology becomes available — information that would ordinarily have become public were these reports subject to peer review — one sees the numerous infirmities in the data base. From the inability of CMAG to capture each unique ad (even in the limited part of the country it purports to monitor), to the subjectivity of the student coding, to the unrestrained and undocumented recoding of the data by Professor Goldstein, one comes to see that this data set and the analyses based upon it cannot be trusted.

Dr. James L. Gibson

October 7, 2002