**UNITED STATES DISTRICT COURT**
**FOR THE DISTRICT OF COLUMBIA**

SENATOR MITCH McCONNELL; REPRESENTATIVE BOB BARR; REPRESENTATIVE MIKE PENCE; ALABAMA ATTORNEY GENERAL BILL PRYOR; LIBERTARIAN NATIONAL COMMITTEE, INC.; ALABAMA REPUBLICAN EXECUTIVE COMMITTEE, AS GOVERNING PARTY FOR THE ALABAMA REPUBLICAN PARTY; LIBERTARIAN PARTY OF ILLINOIS; DuPAGE POLITICAL ACTION COUNCIL; JEFFERSON COUNTY REPUBLICAN EXECUTIVE COMMITTEE; AMERICAN CIVIL LIBERTIES UNION; ASSOCIATED BUILDERS AND CONTRACTORS, INC.; ASSOCIATED BUILDERS AND CONTRACTORS POLITICAL ACTION COMMITTEE; CENTER FOR INDIVIDUAL FREEDOM; CHRISTIAN COALITION OF AMERICA, INC.; CLUB FOR GROWTH; INDIANA FAMILY INSTITUTE; NATIONAL RIGHT TO LIFE COMMITTEE, INC.; NATIONAL RIGHT TO LIFE EDUCATIONAL TRUST FUND; NATIONAL RIGHT TO LIFE POLITICAL ACTION COMMITTEE; THE NATIONAL RIGHT TO WORK COMMITTEE; 60 PLUS ASSOCIATION, INC.; SOUTHEASTERN LEGAL FOUNDATION, INC.; U.S. d/b/a ProENGLISH; MARTIN CONNORS; THOMAS E. McINERY; BARRET AUSTIN O'BROCK; TREVOR M. SOUTHERLAND,

Plaintiffs,

v.

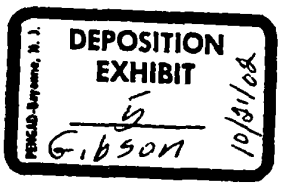
FEDERAL ELECTION COMMISSION; FEDERAL COMMUNICATIONS COMMISSION,

Defendants.

Civ. No. 02-CV-582
(CKK, KLH, RJL)

All related actions.

**REBUTTAL OF KENNETH M. GOLDSTEIN**
**TO EXPERT REPORT OF WILLIAM L. GIBSON**



# Table of Contents

# REBUTTAL TO PROFESSOR WILLIAM L. GIBSON

KENNETH M. GOLDSTEIN
UNIVERSITY OF WISCONSIN

In this rebuttal, I address issues raised by Professor James L. Gibson in his report of September 30, 2002 entitled "An Analysis of the 1998 and 2000 *Buying Time* Reports." I draw on my experience in assembling the databases that were the basis of *Buying Time 1998* and *Buying Time 2000*, as well as my extensive experience in the fields of political science, political advertising, and research methodology. I set out my qualifications more fully on pages 1-2 of my expert report of September 23, 2002.

I focus my discussion on Professor Gibson's criticisms of *Buying Time 2000*, although many of my points are also relevant to his critique of *Buying Time 1998*. I understand that Jonathan Krasno extensively addresses Professor Gibson's related criticisms of *Buying Time 1998* in his rebuttal report.

This rebuttal has five parts. In the first section, I address Professor Gibson's claim that *Buying Time 2000* was the product of a "data set in constant flux" and his related allegations that: (1) the findings of *Buying Time 2000* cannot be replicated; and (2) re-coding of data biased the study's results. I find that neither of these allegations is true. Indeed, many of Professor Gibson's claims stem from his own elementary error of analyzing the wrong data set.

In the second section, I address Professor Gibson's criticisms of the raw data underlying the *Buying Time* reports—the storyboards and frequency data provided by the Campaign Media Analysis Group (CMAG). Contrary to Professor Gibson's suggestion, none of the alleged defects he identifies in the

CMAG data undermine the conclusions of the *Buying Time* studies. Indeed, as Professor Gibson does not dispute, the CMAG data are the most reliable and comprehensive source of information on the content and targeting of political advertising in existence.

In section three, I turn to Professor Gibson's criticisms of the coding instrument. Specifically, I respond to Professor Gibson's suggestion that the best way to identify genuine issue ads is to ask whether an ad focuses on policy issues or on the personal characteristics of a candidate. As I explain below, Professor Gibson's hypothesis would lead to absurd results, and require that most explicit electioneering ads—including a great majority of ads run by candidates themselves—be classified as genuine issue ads.

Section four addresses Professor Gibson's criticisms of the selection and training of the coders. I conclude that none of Professor Gibson's concerns suggest that any systematic bias affected the coding process; indeed, Professor Gibson's suggestion that the coders be specifically "trained" on how to answer the key question on ad purpose would have introduced more, not less, bias into the project.

Finally, section five considers Professor Gibson's critiques of the *Buying Time* databases themselves. Although Professor Gibson claims to have found many (largely unspecified) "errors" in the databases, some random error is to be expected in any database containing more than 40 million data points. Professor Gibson fails to demonstrate that the presence of a small amount of random error in any way undermines the conclusions set forth in the *Buying Time*

reports. And, contrary to Professor Gibson's claim that the databases have never faced academic scrutiny, CMAG data have in fact served as the basis for several peer-reviewed articles, and have been widely available to scholars and researchers.

## I. REPLICABILITY AND RE-CODING

Two serious, and related, accusations are at the heart of Professor Gibson's report. First, he claims that the conclusions of the *Buying Time* studies cannot be replicated—and are not to be trusted—because they were the result of data sets "in constant flux". Second, and more specifically, he charges that I—apparently in collusion with the staff of the Brennan Center—routinely re-coded the responses provided by the student coders in order to bias the results of the studies. In particular, he claims that, in numerous cases, I changed the student's coding of Question 11, which asked the students to characterize the "purpose of the ad" as either "to provide information on a bill or issue" or "to generate support or opposition for a particular candidate".

Neither of these charges is true. I believe that they stem from, first, Professor Gibson's misunderstanding of my role in the coding process; second, Professor Gibson's misunderstanding of my relationship with the Brennan Center; and third—and most importantly—Professor Gibson's misunderstanding of the significance of the various databases produced in this litigation.

As an initial matter, Professor Gibson mischaracterizes the role I played in connection with the *Buying Time* studies. I played no part in writing

*Buying Time 1998* or *Buying Time 2000* or in selecting the conclusions that the authors of these reports chose to draw from the database. On one occasion during the drafting of *Buying Time 2000*, I was called upon by the Brennan Center to advise them of my view with respect to the coding of Question 11 for three ads, and I complied with this request. (See pages 16–17, below.)

The databases that I assembled were subject to routine cleaning, involving the correction of apparent errors and inconsistencies in the coding that inevitably occur in data sets of this size. This is a recognized, routine practice in social science research and bears no resemblance to the result-oriented manipulation that Professor Gibson insinuates. (See pages 10–11, below.)

The vast majority of "re-codings" and "changes" to the database that Professor Gibson identifies do not in fact represent changes to the Question 11 coding reflected in *Buying Time 2000*. Rather, they result from Professor Gibson's failure to consult the appropriate database and his misunderstanding of the database he chose to use. For example, most of the 62 "changes" he identifies in the 2000 database are not changes at all, but rather original student coding of additional CMAG storyboards that had not previously been coded at all, and were not part of the database used by the authors of *Buying Time 2000*. (See pages 14–15, below.)

To the extent that Professor Gibson alleges that I deliberately re-coded data for the purpose of decreasing the number of ads treated as genuine issue ads, I categorically deny his baseless charge and deeply resent the implied challenge to my integrity as a scholar. To prove the point, I note that in pursuing

my own scholarly work after the *Buying Time* studies were drafted, I did re-evaluate the purpose coding of the 2000 database in some 30 instances. In 26 of these instances, I changed the coding from electioneering to genuine issue. To reiterate, changes were made for my own scholarly work, having nothing to do with campaign finance, *after Buying Time 2000* was completed; and 87 percent of these changes involved classifying ads as issue ads, not electioneering. These changes did not and could not have anything to do with the *Buying Time* studies because they were not reflected in those studies. (See pages 14–15, below.)

Professor Gibson's use of the wrong database has also frustrated his attempts to replicate the findings of *Buying Time 2000*. As I demonstrate below, had he used the correct data set, he would have had little trouble in replicating these figures. (See pages 18–21, below.)

A. My Involvement In Preparing the CMAG Databases.

Because Professor Gibson repeatedly insinuates that I somehow colluded with Brennan Center staff to improperly manipulate the data contained in the two *Buying Time* studies, I begin by setting forth exactly what my role was in each of those projects.

My involvement with *Buying Time 1998* and *Buying Time 2000* consisted principally of assembling and providing the Brennan Center with copies of databases containing both data from CMAG and data from student coders. I did not author any part of the two *Buying Time* reports, nor did I create any of the tables or charts or calculate any of the numbers contained in the reports.

The Pew Charitable Trusts funded the work on the *Buying Time* studies.[1] I was technically a subcontractor of the Brennan Center, which forwarded me the portion of Pew funds dedicated to assembling the CMAG databases. The Brennan Center had no control over my use of my portion of the Pew funds, or over my work on the database.

My work on the 1998 database was essentially limited to collecting the CMAG data, supervising the coding by students at Arizona State University, merging the two sets of data (coding and targeting), and forwarding the resulting database to the Brennan Center. I did some routine cleaning of the database, in the sense of correcting errors and logical inconsistencies in the data. I was not asked by the Brennan Center to review any of the responses to Question 6 of the 1998 survey. Jonathan Krasno, one of the authors of *Buying Time 1998*, subsequently sent me a revised version of the 1998 database that incorporated some of the corrections made by the Brennan Center analysts. I have not used my original database or the Brennan Center's revised database in any of my subsequent scholarly work.

My involvement with the 2000 database was somewhat more extensive. Students at the University of Wisconsin coded the ads contained in the 2000 database under my supervision. I cleaned the 2000 database by checking the student coding for errors and inconsistencies and correcting them. I then provided a copy of this cleaned database to the Brennan Center in January 2001.

---

[1] I understand that the Brennan Center submitted a grant proposal to Pew that Professor Gibson references on page 3 of his report. I had no part in preparing this proposal, and, indeed, have never even seen it (apart from the except contained in Professor Gibson's report).

I understand that the authors of *Buying Time 2000* made some additional updates to the database, including some re-coding of the data. There was one instance in which the authors of *Buying Time 2000* consulted me about re-coding the Question 11 variable for certain ads, which I discuss in more detail at pages 16–17, below. The database I provided to the Brennan Center in January 2001, as revised by the Brennan Center analysts, formed the basis of *Buying Time 2000*. I do not have—and did not produce in this litigation—a copy of this "final" iteration of the *Buying Time 2000* database.

While the Brennan Center was cleaning up its version of the 2000 database, I was also updating and cleaning up *my* version of the database, as described below and, in further detail, in Appendix E to my expert report of September 23, 2002. In particular, in the summer of 2001, CMAG sent me about 100 new storyboards, which were coded by students and then incorporated into my version of the database. These storyboards were received too late to be included in the published version of *Buying Time 2000*. Although I subsequently sent the Brennan Center an updated data set containing these additional storyboards, I understand that this updated database was *not* used to calculate the numbers contained in *Buying Time 2000*. Rather, the authors relied on my January 2001 database, as further cleaned by Brennan Center analysts. Consequently, the changes I made in my version of the database subsequent to January 2001 are *not* reflected in the *Buying Time 2000* report.

Since the fall of 2001, I have continued to use my own version of the CMAG database as the basis of scholarly projects and articles having nothing

to do with campaign finance reform. In the process of reviewing and analyzing the database during these ongoing projects, I have continued to clean up my copy of the database and to correct coding errors such as those described in Appendix E to my expert report. Again, these updates and re-codings were *not* part of *Buying Time 2000.*

I will comment briefly on Professor Gibson's assertion that the *Buying Time* studies "should not be accepted as the product of scientific inquiry", but are rather "policy advocacy written by people with a strong ideological commitment to a particular position on campaign finance reform." (Gibson p. 3.) To the extent that Professor Gibson insinuates that the work done by me or by persons working under my supervision in creating the databases perverted their results, his insinuation is both baseless and irresponsible.

I approached the task of constructing the *Buying Time* databases as I would approach any other scholarly project: in the spirit of scientific inquiry and objectivity. I had no preconceived notion about what results would emerge from the CMAG data. To the extent that Professor Gibson relies on the Brennan Center's grant proposal to the Pew Charitable Trusts as evidence of bias, I did not take part in drafting that proposal, and, indeed, never even saw it until it appeared in Professor Gibson's report.

My interest in creating a scientifically valid and reliable database was based on more than just abstract notions of professionalism and objectivity, as important as they were. I always intended to use—and, in fact, have used—CMAG databases in a wide variety of other scholarly studies having nothing to do

with campaign finance reform. The peer-reviewed articles resulting from these studies include:

- Ken Goldstein & Paul Freedman. 2002. "Campaign Advertising and Voter Turnout: New Evidence for the Stimulation Effect." *Journal of Politics*. 64(3).

- Ken Goldstein & Paul Freedman. 2002. "Lessons Learned: Campaign Advertising in the 200 Elections." *Political Communication*. 19(1).

- Ken Goldstein & Paul Freedman. 2000. "New Evidence for New Arguments. Money and Advertising in the 1996 Senate Elections." *Journal of Politics*. 62(4).

- Paul Freedman & Ken Goldstein. 1999. "Measuring Media Exposure and the Effects of Negative Campaign Ads." *American Journal of Political Science*. 43(October).[2]

Like any conscientious researcher, I was careful not to compromise my academic reputation by involving myself in any result-oriented research project. Professor Gibson has cited absolutely no facts to suggest otherwise and I deeply resent his baseless attack.

B. Comparing Different Versions of the Databases.

Professor Gibson correctly notes that several versions of both the 1998 and 2000 databases have been produced in this litigation by both myself and the Brennan Center. The reason for this is no mystery: social science researchers, like all prudent people, periodically back up their work when using computers. When preserved, these backups give us a "snapshot" of the database at a certain

[2] Several of these articles use CMAG data from the 1996 and 1997 election cycles, which I assembled and coded in a manner similar to the *Buying Time* databases.

point in time. One does not simply push a button and instantaneously generate a single, definitive version of a 40-million point data set.

Professor Gibson also notes, correctly, that certain versions of the database contain more than one variable corresponding to a particular question. For example, my database titled "cmag_2000_labeled_data_may02_feb01_w_compet.sav", on which Professor Gibson relies for his analysis of the *Buying Time 2000* data, contains two different variables coding responses to Question 11: "OLDQ11" and "NEWQ11".[3] This, too, presents no mystery. A researcher might want to include different iterations of the same variable in one database for many reasons: to compare changes in the variable over time, for example. A researcher would *not* follow the practice of retaining prior codes if he were trying to manipulate data in a sinister fashion.

Finally, Professor Gibson correctly notes that the data varies somewhat between different versions of the database and between different versions of a variable within the same database. Again, there is nothing mysterious about this fact. There are two principal reasons for these differences, neither of which in any way undermines the reliability of the databases.

*First*, some of the differences Professor Gibson describes may have resulted from routine "cleaning" of the data sets. It is standard social science practice to clean a data set by correcting apparent errors after the codes have been entered into the database. See, for example, Herbert F. Weisberg, Jon

---

[3] Actually, as Professor Gibson notes, there are four Question 11variables. Of the two other variables, one, labeled "Q11", is coded identically to "NEWQ11", and the other, labeled "Q11_1", is coded identically to "OLDQ11". For simplicity's sake, I confine my discussion to the "OLDQ11" and "NEWQ11" variables.

A. Krosnick & Bruce D. Bowen. 1996. *An Introduction to Survey Research, Polling, and Data Analysis*. 3rd ed. Thousand Oaks: SAGE Publications. For example, after merging the 2000 coding into the database, my assistants and I examined the data for the presence of "wild codes"—that is, entries for which no corresponding code existed in our codebook. When we discovered such an error, we looked at the ad's storyboard and inputted the correct information into the database. In other cases, we reviewed and corrected logically inconsistent answers. My assistants and I also filled in missing data for some ads on a number of objective questions, such as candidate mention, in both the 1998 and 2000 databases.[4]

Because this cleaning process involves the correction of errors in the database, one would expect to see slightly different results when comparing an uncleaned database to a cleaned database. Thus, one would expect the cleaned versions of the databases to differ slightly from the original student coding. Moreover, because database cleaning for such an enormous data set can be a lengthy process, one would expect to find small discrepancies between databases saved at different points during the cleaning process.

*Second*, the more substantial differences between the various versions of the 2000 database produced in this litigation can be explained by the fact that, after I transmitted my cleaned version of the 2000 database to the Brennan Center, I continued updating and revising my own copy of the database

---

[4] For example, student coders who coded an ad as intended to promote an issue were directed not to fill in certain objective information regarding candidate mention. This created gaps in the data set which the Brennan Center analysts and I reviewed and filled in as part of the cleaning process.

as part of my continuing scholarly work. These continuing revisions are explained in detail in Appendix E of my expert report of September 23, 2001.

I continued updating the data set because I have continued to use the CMAG database in a number of other academic projects that are entirely unrelated to the subject of the *Buying Time* reports or to the topic of campaign finance. For example, my research assistant Patty Strach discovered a number of miscodes in the course of her work on the Gender in Political Advertising Project. We corrected these mistakes, which were mostly the result of errors in assigning contextual information (for example, coding an ad as relating to the New Jersey 2nd Congressional District race rather than the New Jersey 3rd Congressional District Race).[5]

Similarly, during validity checks with a number of local television stations after we had transmitted the data set to the Brennan Center, we discovered that CMAG had not transferred to us information on approximately 100 unique ads, representing about 20,000 airings. These new storyboards were coded by student coders, and the data added to my version of the database.

While these continuing updates are reflected in some of the databases I have produced in this litigation, they were *not* included in the database used to produce *Buying Time 2000.*

C. The Alleged Re-Coding of Ads in the 2000 Database.

Professor Gibson's misunderstanding of my role in the coding process and of the significance of the various versions of the CMAG database

[5] There was no systematic pattern to these errors.

produced in this litigation leads him to exaggerate greatly the amount of re-coding that occurred in the *Buying Time* project.

Professor Gibson claims that there are substantial differences between the data coded by the student coders and the data that is reflected in *Buying Time 2000*—particularly in regard to Question 11. The best way of judging the accuracy of this accusation would be to compare a database containing the original student's coding with a database reflecting the numbers used in the preparation of the *Buying Time* studies. While Professor Gibson claims to have done this, in fact he actually used the wrong database, rendering his conclusions meaningless.

Professor Gibson asserts that 62 group-sponsored ads in the 2000 database "appear . . . to have their codes changed on the 'purpose of the ad.'" (Gibson p. 52.) He bases his accusations of re-coding on his analysis of a database produced by me entitled "cmag_2000_labeled_data_may02_feb01_w_compet.sav". (Gibson p. 47.) Professor Gibson compares the data recorded in that database for the variable OLDQ11—which he assumes reflects "the original coding of the ads on this attribute"—to data recorded for the variable NEWQ11—which he apparently assumes reflects the coding used in the preparation of *Buying Time 2000*. (Gibson p. 51.)

The problem with Professor Gibson's analysis is that both of these assumptions are wrong. The data set he analyzes is *not* the version of the database used by the Brennan Center to produce *Buying Time 2000*. Rather, it is a later iteration of my own version of the database containing my own after-the-fact

updates and re-codes, and including additional ads later received from CMAG. The OLDQ11 variable in this database reflects, not the original student coding of the ads, but rather a later iteration of my version of the database that incorporates both the cleaning performed before transmitting the data set to the Brennan Center, and some additional re-coding performed in February 2001. Even more critically, the NEWQ11 variable reflects, not the data used in *Buying Time 2000*, but rather a much later version of the data contained in my version of the database that incorporates rather extensive updating and re-coding, including information on 100 ads that we had not initially received. While both of these variables include re-coding, none of this re-coding ever made its way into the *Buying Time 2000* report.

As a result, Professor Gibson is not measuring the re-codes that were reflected in *Buying Time 2000*, but rather the updates that I made *after the publication of Buying Time*. These post hoc updates were *not* reflected in the database used by the Brennan Center to produce *Buying Time 2000*. Consequently, Professor Gibson's comparison of OLDQ11 to NEWQ11 tells us nothing about the extent to which the data reflected in *Buying Time 2000* differed from the original student coding.

Indeed, Professor Gibson's analysis misrepresents even the extent and nature of my own post-publication revisions. Professor Gibson states that there were "changes" in the Question 11 coding of 62 group-sponsored ads. My own analysis shows that there were only 31 actual changes. For 26 of these ads, the coding was changed from electioneering to genuine issue; one other ad was

changed from unclear to genuine issue. Only four ads were re-coded from genuine issue ads to electioneering. In addition, 61 additional storyboards were coded as "missing data" in OLDQ11, but re-coded as either "providing information" or "promoting a candidate" in NEWQ1.[6] These 61 cases do not represent coding changes, but rather represent the initial student coding of the new storyboards first provided to me by CMAG in the summer of 2001. In other words, the "change" was from a missing data field to a coder assessment of Question 11.[7]

Putting aside Professor Gibson's unsuccessful exercise, can we more accurately measure the changes in Question 11 coding that were actually incorporated into *Buying Time 2000*? As explained above, to perform this calculation, we would need to compare the original student coding of the advertisements with the coding used to produce the figures in *Buying Time 2000*. Based on my communications with Craig Holman, the primary author of *Buying Time 2000*, and on my independent review of the data set, I have identified a file entitled "federal.sav", produced by the Brennan Center, as closely tracking the database on which *Buying Time 2000* was based.

---

[6] I list the ads affected by these changes and additions in Appendix A to this rebuttal.

[7] One of the more disturbing aspects of Professor Gibson's report is that the information available to him in the database clearly put him on notice of the fact that most of the "changes" he describes were, in fact, original coding of storyboards for which no coding had earlier been supplied. And, of the actual changes made, the great majority were from electioneering to genuine issue. Professor Gibson chose to ignore this information and instead insinuate that these "changes" were made to bias the result of the survey.

I have been unable to identify any 2000 database containing only the unmodified input of the student coders on all questions. However, based on my communications with Dr. Holman and on my independent review of the data set, I have identified a file entitled "cmagdata_1", produced by the Brennan Center, as the cleaned version of the 2000 database that I transmitted to the Brennan Center in January 2001. Because I did not change any of the student coders' responses to Question 11 in the data cleaning process, I believe that this database accurately reflects the original student coding on Question 11.

I have compared these two databases to determine whether any advertisements were re-coded on Question 11 from "promoting a candidate" to "providing information or urging action on an issue". My analysis reveals that such changes were made in three cases:

- #1269 (CBM/PA Sherwood If You Don't Have Health);
- #1367 (RI/RIWV Langevin Abortion); and
- #2107 (WI/NPLA Feingold Kohl Abortion 60).

I decided to re-code these three ads as electioneering following a conference call in March of 2001. Ad #1269 was a cookie-cutter ad sponsored by "Citizens for Better Medicare" (a front group representing the pharmaceutical industry). The authors of *Buying Time 2000* had noted that a number of other, extremely similar, CBM-sponsored ads had all been identified by the student coders as electioneering, while #1269 had been coded as a genuine issue ad.[8] After reviewing the storyboards of these ads, I concluded that #1269 was not

[8] The CBM ads coded as electioneering included #1544, #1650, #2163, #2424, #2926, and #3309.

meaningfully distinguishable from the other CBM ads, and re-coded it as electioneering.

Ad #2107, sponsored by the National Pro-Life Alliance, was directed at Wisconsin Senators Russell Feingold and Herb Kohl. The authors of *Buying Time 2000* had noted that a virtually identical version of this ad, targeting Senator Charles Robb of Virginia (#2089), had been classified by the student coders as electioneering. After reviewing the storyboards of the two ads, I concluded that the ads were not meaningfully distinguishable, and resolved the inconsistency by re-coding #2107 as electioneering.[9]

Finally, ad #1367, sponsored by "Rhode Island Women Voters" (a front group for Emily's List), was originally coded as a genuine issue ad. The authors of *Buying Time 2000* disagreed with this coding, and after consulting with them, I decided to re-code this ad as electioneering. I stand by this classification; in my opinion, this ad is clearly electioneering.

These are the *only* ads that were changed from "genuine issue ads" to "electioneering" in the 2000 data set. The logical next question is, did the re-coding of these ads undermine the basic conclusions of *Buying Time 2000*?

The answer is no. In my expert report of September 23, 2002, I decided to test *Buying Time 2000*'s conclusions about the impact of BCRA on genuine issue ads by treating all of these three ads as genuine issue ads. Treating these three ads as genuine issue ads, I found that genuine issue ads made up only

---

[9] I have subsequently come to a different conclusion as to the proper coding for this ad. In my own version of the database, I have re-coded this ad as a genuine issue ad. However, it was coded as electioneering for the *Buying Time 2000* report.

2.3 percent of all interest group ads that would be regulated under BCRA. In other words, even counting these three cases as genuine issue ads, the student coders classified 97.7 percent of all interest group ads that were broadcast within 60 days of the election and mentioned a candidate as electioneering.

D. Replicability of *Buying Time* 2000.

Professor Gibson also claims that he has been unable to replicate the results contained in *Buying Time 2000* by analyzing certain (but not all) of the databases produced in this litigation. Professor Gibson's inability to replicate some of these conclusions from a certain database does not prove that those conclusions are not replicable, nor does it undermine the validity of any of the major conclusions of *Buying Time*. Rather, Professor Gibson's inability to replicate *Buying Time*'s conclusions once again stems from his use of the wrong data set.[10]

In peer-reviewed academic work, social science scholars typically provide reviewers with the coded data necessary to replicate a study's analysis. The Brennan Center and I have gone far beyond this. In the course of this litigation, we have provided, not only the data sets necessary to replicate the

[10] Even putting aside Professor Gibson's use of the wrong database, in those rare instances where he discloses the actual quantitative content of his alleged inconsistencies, the triviality of his observations is striking. For example, on page 24 of his report, Professor Gibson announces that Buying Time 1998 reported 302,860 airings of political television advertisements, while the database he analyzed showed 307,308—a difference of 1.4 percent. Similarly, he complains that Buying Time reported 22,151 group-sponsored airings, while his database showed 21,151—a variance of 3.5 percent. These variances are so small as to suggest their own triviality; Professor Gibson makes no attempt to show how these modest variances could affect the general conclusions yielded by the database.

analysis in the *Buying Time* studies, but also the underlying storyboards and targeting data. Had Professor Gibson so chosen, he could have recreated not just particular tables, but the data underlying the entire study.

Notably, although Professor Gibson had access to the CMAG storyboards and targeting data, he has apparently never attempted to replicate the *Buying Time* studies in the most direct fashion—that is, by re-coding all (or even a sample) of the captured advertisements and comparing the results of his coding exercise with the results of mine. Because Professor Gibson never attempted to test the conclusions implicit in the database by replicating the coding exercise, most of his assertions about the reliability and validity of the conclusions drawn from the databases are necessarily speculative.

Rather than attempting to replicate the studies, Professor Gibson tried to replicate results from the already-coded databases produced to plaintiffs in this litigation. Of course, before a scholar can replicate the findings of a study, he must ensure that he is working off of the same data as the authors of the study. Although, as Professor Gibson notes, both the Brennan Center and I produced a number of iterations of both the 1998 and 2000 databases, Professor Gibson apparently confined his replication exercise to analysis of a single database for each year.[11] Unfortunately, he chose the wrong 2000 database. Because the 2000 database that Professor Gibson chose to analyze is *not* the same one that was used

[11] Although Professor Gibson identifies the databases he analyzed, he does not explain why he chose to analyze these particular databases and not others. Nor does he reveal whether he conducted any analysis to determine, as a preliminary matter, whether other databases produced in this litigation would be more likely to produce results more consistent with the published *Buying Time* studies.

to derive the figures reported in *Buying Time 2000*, it comes as no surprise that he was unable to replicate the results of the studies.

As noted above, Professor Gibson relies throughout his report on a data set created by me on August 1, 2002 entitled "cmag_2000_labeled_data_may02_feb01_w_compet.sav". (Gibson p. 47.) As the creation date of the file makes unmistakably clear, this is not the data set used by the authors of *Buying Time 2000*. Rather, as I explain above, it is a later iteration of my own version of the database containing updates incorporated in the course of other academic projects.

As noted above, I understand that the Brennan Center has produced a file entitled "federal.sav" that very closely tracks the "final" data set used in the production of *Buying Time 2000*. I believe that most, if not all of the "inconsistencies" reported by Professor Gibson are the result of his use of my updated data set rather than this "final" database.

Using the "federal.sav" data set, I have been able to replicate key findings of the *Buying Time* study. For example, my analysis of this database reveals 142,421 airings of ads by interest groups—exactly the same number reported on page 72 of *Buying Time 2000*. I find 845,923 total airings in federal elections—the same as reported on page 53 of *Buying Time 2000*. And I find 57,451 airings using *Buckley*'s "magic words"—the same number reported on page 72 of *Buying Time 2000*. Other results, while not matching exactly, correlate to within a fraction of a percentage point. For example, I find 230,147 airings by political parties, compared to 230,123 in *Buying Time 2000*. These

results confirm my belief that "federal.sav" provides a much more appropriate basis for replicating the conclusions of *Buying Time 2000* than the data set that Professor Gibson used.

Professor Gibson also specifically claims that he is unable to replicate one of the key findings of *Buying Time 2000*:

> Of all group-sponsored ads that depicted a candidate within 60 days of the election, 99.4% were found to be electioneering ads. 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*, p. 73; emphasis in original).

Once again, it appears that Professor Gibson's difficulties areise from his use of the wrong database. Using the "federal.sav" database, I was able to identify exactly three ads, comprising 331 airings, that satisfy *Buying Time 2000*'s criteria. They are:

- #627 (KY/CFAW Call Northup): 172 airings;
- #1389 (FAIR/IA Latham Foreign Worker Bill): 81 airings; and
- #2862 (UT/COC Matheson Can't Decide Rx): 78 airings.

To replicate the denominator used in *Buying Time 2000*, I used the "federal.sav" database to identify all ads run by interest groups that mentioned a candidate and aired within 60 days of the election. I arrived at a number of 53,840. Dividing these figures, I arrive at a percentage of 0.6 percent—the same as reported in *Buying Time 2000*:

$$331 / 53{,}840 = 0.6\%$$

In short, the results reported in *Buying Time 2000* are fully replicable using the materials produced to plaintiffs in this litigation.

Rather than performing this relatively simple replication, however, Professor Gibson instead chooses to engage in a remarkably bizarre manipulation of the data in order to artificially inflate *Buying Time 2000*'s "false positive" count. Specifically, he identifies a list of 30 specific ads "provided to [him] by counsel" that were coded as electioneering and that constitute approximately 24 percent of all airings by interest groups that mention a candidate within 60 days of the election. (Gibson p. 62.) Without discussing the content of any of those ads or elaborating any argument as to why we should ignore the student coding, Professor Gibson asks us to *assume* "that each of these ads could be fairly coded as 'providing information.'" (Gibson p. 62.) His conclusion: if we accept his baseless assumption and re-code these 24 percent of BCRA-regulated airings ads as genuine issue ads, the false positive percentage will rise to something around 24 percent.

I have no idea what Professor Gibson was trying to accomplish by this exercise. If we "assumed", based on the representations of plaintiffs' counsel, that 100 percent of the BCRA-regulated group ads were genuine issue ads, then we could arrive at a false positive percentage of 100 percent. But where—as here—such assumptions are based on nothing more than the self-serving conjecture of plaintiffs' counsel, any such percentage is meaningless. Because we can accurately replicate the results of *Buying Time 2000* from the

actual student coding data, there is no reason to accept Professor Gibson's exercise in unsupported assumption as social science.

## II. THE CMAG DATA

Professor Gibson also voices a number of concerns about the "raw materials" used in compiling the *Buying Time* databases—that is, the storyboard and frequency data collected by CMAG. However, while Professor Gibson asserts that there are "many limitations" to this data, he does not even attempt to explain how these alleged limitations undermine the validity of the conclusions set forth in *Buying Time*.

### A. The Number of Markets Monitored.

First, Professor Gibson points to CMAG's monitoring of "only 75 media markets (out of 210 such markets in the country)", and its lack of monitoring of local cable ads, as somehow limiting the ability to "generaliz[e] the findings of this study to all political communications". (Gibson, pp. 7, 8.) But Professor Gibson fails to note that the 75 largest media markets monitored by CMAG are geographically dispersed across the nation and cover over 80 percent of the television households in the United States.[12] Candidates and political parties interested in monitoring elections across the nation (including both the Democratic and Republican Executive Committees, not to mention several of the

[12] I detail the scope of the CMAG data in "Measuring Exposure to Campaign Advertising", an article I coauthored with Travis M. Ridout, Michael Franz and Paul Freedman, which is currently under review at the journal *Political Communication*. A copy of the manuscript is attached as Appendix I to my expert report of September 23, 2002.

plaintiffs in this litigation) regularly purchase and rely on CMAG's political advertising data—evidencing market acceptance of the data's comprehensiveness and reliability by those to whom it matters most. In short, the CMAG data includes the great majority of political advertisements appearing on American television screens in 1998 and 2000.

I have no reason to believe that CMAG's inability to capture ads run on local cable undermines the reliability of the data or the conclusions drawn from it in the *Buying Time* studies. Aggregate spending on political ads run on local cable is extremely small in relation to spending on political broadcast spots. Moreover, Professor Gibson does not offer any reason to believe that the ads run on local cable are significantly different than the broadcast ads captured by CMAG. And he does not even suggest that CMAG's inability to capture local cable spots introduced any systematic bias into the data.[13]

B. The Video Captures.

Professor Gibson also complains that CMAG does not capture the entire video portion of an advertisement, instead capturing only a still "snapshot" of every fourth second of video. In my extensive experience working with thousands of CMAG storyboards, coupled with my familiarity with many of the actual videos, I have found no reason to believe that this factor compromised

[13] Professor Gibson also notes that CMAG does not cover radio advertisements. This is true, but irrelevant—the Buying Time studies, by their own terms, only sought to analyze political television advertising. In any event, aggregate spending on political radio ads is only a fraction of what is spent on political television advertising. See, for example, Anna Nibley Baker & David M. Magleby. "Interest Groups in the 2000 Congressional Elections", in David B. Magleby, ed. 2002. *The Other Campaign*. Lanham: Rowman & Littlefield.

coders' ability to accurately analyze the content of ads, especially because CMAG provides a complete transcription of the audio portion of the ad along with the video captures.[14] Moreover, as is the case with many of Professor Gibson's criticisms of "incomplete" information, there is no reason to believe that there in any systematic bias associated with the CMAG technology capturing only one video frame every four seconds.

C. The CMAG Media Markets.

A third criticism offered by Professor Gibson is that the media markets analyzed by CMAG are not always co-extensive with electoral districts. However, the particular electoral district targeted by an ad was usually obvious—most typically, from the identity of the candidate or officeholder mentioned in the ad. Where no district was obvious from the ad itself (for example, in so-called "cookie cutter ads), we turned to contextual information, including information from media outlets in those markets, sponsor web sites, and consultations with congressional experts and campaign consultants, to place the ad in the appropriate district.

D. Cookie-Cutter Ads.

Finally, Professor Gibson observes that the CMAG technology is unable to distinguish between different versions of so-called "cookie cutter" ads—that is, ads run by interest groups in multiple election districts that are identical except for the candidate or officeholder they name. (Gibson p. 7.) It is

---

[14] Significantly, federal regulations require that ad sponsorship disclosures be displayed on the screen for at least four seconds, making it possible for CMAG to capture this information in its video snapshots in most cases. 47 C.F.R. § 73.1212(a)(2)(ii).

true that because the CMAG technology cannot identify these slightly different ads a unique, all cookie-cutter airing are initially assigned to the same storyboard by the CMAG tracking system. For example, CMAG reported that a Chamber of Commerce ad mentioning Arkansas congressional candidate Mike Ross aired in both Little Rock, Arkansas and Detroit, Michigan. In fact, the ad running in Detroit---while largely identical in content to the Arkansas ad—mentioned not Ross, but Michigan candidate Dianne Byrum.

Contrary to Professor Gibson's suggestion, however, the existence of such cookie-cutter ads did not result in errors. My assistants and I were able to adjust for the cookie-cutter phenomenon in the 2000 database by using simple, old-fashioned social science research. As explained in Appendix E to my expert report, we did not unquestioningly accept CMAG's classification of cookie-cutter ads. When we identified an ad as a cookie-cutter, we looked beyond the CMAG storyboard to determine which candidate or officeholder was actually mentioned by the ad in each market where it ran. We compared maps of media markets and congressional districts, looked at published information on competitive races, and contacted political consultants, party officials and local media outlets to determine the actual targeted district and individual. In this way we were able, for example, to assign the "Ross" ads running in Detroit to the Byrum race. After assigning a cookie-cutter ad to a new district, we also made any necessary changes to associated contextual information, such as the incumbency status of the mentioned candidate. I am confident that these procedures enabled me to

accurately adjust for the appearance of cookie-cutter advertisements in the 2000 database.[15]

As with Professor Gibson's other criticisms, he fails to provide any reason to believe that his impressionistic complaints about CMAG affected the general conclusions drawn from the databases by the authors of the *Buying Time* studies.

## III. THE CODING INSTRUMENT

I now turn to Professor Gibson's criticisms of the coding instrument—that is, the questionnaire the coders filled out while viewing the storyboards. I focus particularly on Professor Gibson's critiques of 1998's Question 6 and 2000's Question 11, which ask the coders to classify the purpose of the ad as either generating support or opposition to a candidate, or providing information or urging action on an issue.

### A. Alleged Bias in the Construction of Questions 6 and 11.

Professor Gibson notes that 97.7 percent of airings coded as having the purpose of generating support or opposition to a candidate were also coded as mentioning a candidate. While this is hardly unexpected—after all, it is difficult to imagine an ad intended to promote or oppose a candidate that did *not* mention that candidate—Professor Gibson tries to draw a different lesson from this entirely unremarkable statistic.

---

[15] I did not undertake to adjust for the presence of cookie-cutter ads in the 1998 database. However, I understand that Jonathan Krasno, one of the authors of *Buying Time 1998*, undertook a similar process of identifying and re-coding cookie-cutter ads appearing in the 1998 data set.

Specifically, Professor Gibson charges that, because the coding questionnaire instructed coders to fill in information regarding candidate mention only if the ad was coded as generating support or opposition to a candidate, coders may have erroneously assumed that "any advertisement that included the name of a candidate should be coded as having a purpose of promoting or opposing a candidate". (Gibson p. 30.) This argument is a non-sequitur.

The use of branching questions is a common and accepted survey technique. There is no logical reason to believe that the subsidiary questions about a candidate who was mentioned somehow predisposed coders to answer the earlier, primary question about ad purpose in a particular way. Professor Gibson provides no citation to any social science or survey design literature in support of his hypothesis. And the only empirical evidence he presents is the fact that most ads coded as intended to promote or opposing a candidate also mention a candidate—again, an entirely expected state of affairs.

Moreover, the coders' classification of ads as electioneering is fully supported by other objective characteristics of those ads. One would expect, for example, that ads designed to promote or oppose a candidate would air relatively close to Election Day. The data verifies this assumption. The "federal.sav" database reveals that during the 2000 election cycle, 79.8 percent of the group-sponsored ads classified as electioneering were coded as having run within 60 days of the election, compared to only 18.7 percent of non-electioneering ads. In light of this corroborating data, I am unpersuaded that the

coders were systematically biased toward coding ads mentioning candidates as electioneering as opposed to issue ads.

B. The Alleged Conflict Between Purpose of Ad and Policy/Personal Focus.

Professor Gibson further argues that there is a conflict between the coding of Questions 6 and 11 and the coding of Question 22 (in *Buying Time 1998*) and Question 27 (in *Buying Time 2000*), which ask coders to indicate whether the "primary focus of [the] ad [is] on the personal characteristics of either candidate or on policy matters". Professor Gibson speculates that "[o]ne might predict that . . . 'candidate promoting 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." (Gibson, p. 32.)

Of course, Professor Gibson's speculation is not borne out in the data, and no one expert in (or even familiar with) elections or political advertising would expect it to be. Contrary to Professor Gibson, there is no reason to believe his unstated (and unsupportable) assumption that an ad that focuses on policy matters cannot support or oppose a candidate. Indeed, one would expect the bulk of electioneering ads to address "policy matters".[16] The vast majority of

[16] The use of policy issues in electioneering ads is widely noted in political science literature. See, for example, David B. Magleby & Marianne Holt, eds. 1999. *Outside Money: Soft Money & Issue Ads in Competitive 1998 Congressional Elections*. Provo: Brigham Young University; Marion R. Just, Ann N. Crigler, Dean E. Alger, Timothy E. Cook, Montague Kern & Darrell M. West. 1996. *Crosstalk: Citizens, Candidates, and the Media in a Presidential Campaign*. Chicago: University of Chicago Press; Craig Leonard Brians & Martin P. Wattenberg. 1996. "Campaign Issue Knowledge and Salience: Comparing Reception from TV Commercials, TV News and Newspapers". *American Journal of Political Science* 40: 172-193; Darrell M. West. 1993. *Air*

electioneering ads do not repeat, ad nauseum, "Vote for Smith". Nor do most electioneering ads say, "Don't vote for Jones because he's a liar". Rather, most electioneering ads seek to influence votes by portraying the favored candidate as espousing reasonable policy positions on hot-button issues like taxes, Social Security or abortion, and the opponent as having unreasonable or even dangerous positions on the same issues.

To confirm this view, we need only turn on the radio or television and look at the torrent of candidate-sponsored ads dealing with their positions on policy issues. These advertisements, which by definition are electioneering, make great use of policy positions and policy differences.[17] The CMAG data confirms this point. The "federal.sav" databases show that in 2000, fully 53.3 percent of candidate-sponsored ads focused on policy issues rather than personal characteristics; an additional 35 percent focused on both policy and personal issues. Only 10.8 percent focused just on personal issues. Ads that use *Buckley*'s "magic words"—that is, which explicitly urge the viewer to vote for or against a candidate—also focus more heavily on policy issues rather than personal characteristics. In 2000, 47.4 percent of ads using magic words were coded as

---

*Wars: Television Advertising in Election Campaigns, 1952-1992.* Washington: Congressional Quarterly Press.

[17] In Appendix B, I provide two sample storyboards illustrating this phenomenon. The first, sponsored by Colorado congressional candidate Ken Toltz, focuses on the issue of gun control. The second, sponsored by Illinois congressional candidate Shawn Donnelly, focuses on the issues of taxes, Medicare and Social Security. Each of these ads—which were paid for by the campaigns themselves—is indisputably designed to promote the election of a candidate, and each just as indisputably focuses on policy issues, not personal characteristics.

having a policy focus; 34.5 percent focused on both policy and personal issues. Only 16.5 percent focused solely on personal issues.

In short, Professor Gibson's strange thesis—that ads coded as electioneering should disproportionately focus on personal characteristics instead of policy issues—is disproved by the fact that both candidate ads and express advocacy ads, which everyone agrees are electioneering, themselves focus primarily on policy issues.

## IV. THE CODERS

While Professor Gibson criticizes my use of undergraduate student coders, he is unable to undermine any of the conclusions reached by the coders. As noted above, Professor Gibson had full access to the same CMAG storyboards viewed by the student coders. Nevertheless, he chose not to conduct his own survey, using his own coders and his own training techniques, and compare it to the results reached by the undergraduate coders. As a result, his criticisms of the decision to use undergraduate coders are necessarily speculative.

### A. Lack of Coder Training.

Professor Gibson claims to have "significant questions" about the coding process. (Gibson p. 9.) For example, Professor Gibson claims that because the student coders received no formal training, he doubts their ability "to make accurate assessments on highly subjective characteristics of these ads". (Gibson p. 10.) In particular, he questions the student coders' ability to accurately answer 1998's Question 6 and 2000's Question 11, which ask:

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?

Professor Gibson's criticism misses the point of this question. Questions 6 and 11 are intended to record the perception of viewers as to whether they believe an ad was intended to promote an issue or elect a candidate. The reason for asking these questions was to determine the impression that an ad made on viewers. Failing to provide the coders with any additional training about how to answer Questions 6 and 11 was not a methodological oversight, but a deliberate choice that is well-supported by social science principles. It was aimed at getting the untutored common-sense impression of the coders while minimizing the possibility of biasing coders with any preconceived notions that might have been implicit in a set of instructions. After all, Questions 6 and 11, as Professor Gibson repeatedly acknowledges, are in a sense subjective questions. By not exposing the coders to formal training, we were able to ensure that the coders' responses were their own, and not influenced by any perceptions that could have been communicated to them by formal training. The kind of formal training which Professor Gibson seems to endorse would only undermine the independence of the coders' assessments and possibly introduce systematic bias into the survey.

Moreover, by not exposing the coders to formal training, we were better able to simulate, as much as possible within a controlled environment, the experience of a typical viewer watching the ads at home. A typical viewer does not have the benefit of any formal training as to how he or she should subjectively perceive the purpose of a political television ad. In short, Professor Gibson

provides neither social science authority nor any logical reason to believe that the lack of formal training in any way hampered the coders from making what Professor Gibson acknowledges was a subjective assessment.

B. Representativeness of the Coders.

Immediately after complaining that the coders did not receive adequate training, Professor Gibson takes an inherently inconsistent position and argues that the undergraduate student coders were not representative of "average viewers". However, the use of undergraduate subjects in studies measuring subjective perceptions of external stimuli is well-established and accepted social science procedure. Professor Gibson makes no attempt to demonstrate that the identity of the coders led to any systematic bias that might undermine the general conclusions drawn from their coding. Absent such a demonstration, there is no reason to doubt the validity of using this common social science practice.

C. Method of Selecting the Coders.

I do not believe that the composition of the student coders introduced any systematic bias in to the *Buying Time* studies. The data contained in the *Buying Time* databases were coded primarily by teams of undergraduate students. In 1998, all data was coded by undergraduate honors students from the Arizona State University. In 2000, data was primarily coded by a team of six undergraduates students enrolled in my upper-level Interest Groups course at the University of Wisconsin.[18]

[18] A number of ads captured relatively early in the 2000 election cycle were coded by my graduate student research assistants, as explained in Appendix E to my

The undergraduate students were *not* informed that the data they were coding would be used by the Brennan Center, or anyone else, to study the effects of campaign finance legislation. And, while some of the coders were students of mine, I do not believe that I ever expressed to them any policy preference as to the desirability or undesirability of campaign finance legislation, either in the classroom or during the coding process. To the best of my knowledge, the coders were unaware of the policy implications of their answers to Questions 6 and 11. If anything, they thought the project was about the tone of political advertising—one of my primary scholarly interests.

D. Inter-Coder Reliability.

To explore the accuracy of the undergraduate students' coding, I undertook a test of inter-coder reliability on the 2000 data.[19] I randomly selected 150 unique advertisements, which were then each coded by five separate undergraduate coders. In general, I found inter-coder reliability to be extremely high.

To further measure the reliability of the coders' assessment of Question 11—asking them to classify the purpose of the ad as either providing information/urging action on a bill or issue or generating support or opposition for a candidate—I recently conducted another test of inter-coder reliability using all ads sponsored by interest groups as the population. From these over 350 advertisements, I randomly selected 50 and asked 10 undergraduate students to

---

expert report. However, all ads airing with 60 days of the election the primary focus of BCRA—were coded by undergraduate coders.

[19] The procedures and results of this test are set forth in detail in Appendix I to my expert report of September 23, 2002.

code them on three attributes. These were 1) whether the ads "generate support or opposition for a particular candidate" or "provide information or urge action"; 2) their tone (attack, contrast, or promote); and 3) their focus (a candidate's personal attributes or policy). Each student was given written instructions and a spreadsheet in which to enter his or her answers. Although none of the coders had experience coding political advertisements, each had previously done coding of local news broadcasts. Each student took approximately one hour to code the 50 advertisements. Four of the ads were subsequently dropped from the analysis because their codes were missing in the original dataset. In other words, they were creatives that were provided to us by CMAG in summer of 2001.

I analyzed only interest group advertisements in this reliability exercise, as opposed to interest group and candidate-sponsored ads, because there is much more variation in interest group ads in the answer to Question 11, concerning the purpose of the advertisement (to generate candidate support or opposition or to provide information or urge action). In the sample of 46 cases, sixty-four percent of the time the original code was "generate candidate support," twenty-six percent of the time it was "provide information or urge action," and two percent of the time it was "unsure/unclear."

We found a high degree of agreement between the original codes and those given by the ten coders. Considering the 50 ads as a whole, 75 percent of the time when the original code was "generate support or opposition," the ten coders agreed. In twenty-three percent of the cases, they believed the ad was designed to "provide information," and in two percent of the cases they were

"unsure." Similarly, 75 percent of the time when the original code was "provide information" the coders agreed. In 18 percent of the cases, they believed the ad's purpose was to "generate support or opposition," and in 7 percent of the cases they were "unsure." Given that the coders agreed with the original code in 75 percent of the cases, *regardless of what that original code was*, there is no hint of systematic bias in the original coding.

## V. THE DATABASES

Professor Gibson does not deny that the databases created from the 1998 and 2000 CMAG data are the largest, most comprehensive and most systematic collections of political television advertising ever assembled. Nor does he identify any more comprehensive or reliable collection of data on which the *Buying Time* authors should have relied.

Professor Gibson does, however, voice a number of criticisms regarding the *Buying Time* databases in general. First, he claims that the databases are "riddled with internal inconsistencies and errors". (Gibson p. 23.) Second, he complains that he was unable to replicate the conclusions contained in the two *Buying Time* studies. (Gibson pp. 23-24, 47-48.) He also asserts that the fact that the *Buying Time* studies were not formally peer-reviewed casts doubt on their integrity. (Gibson pp. 4, 45.) None of Professor Gibson's criticisms undermine the reliability or validity of either the CMAG databases or the *Buying Time* studies.

A. Errors in the CMAG Database.

Each of the two CMAG databases contains information on over 970,000 separate airings of political advertisements, each coded for over 40 variables and a slew of contextual variables. As a result, each database contains over 40 million data points.

Some errors are inevitable in any database of this size. Some data will be missing, some variables will be miscoded, and other entries will be duplicated. However, contrary to Professor Gibson's insinuation, the presence of such errors does not necessarily mean that the database is invalid or unreliable.

Throughout most of his report, Professor Gibson fails to make the rudimentary distinction between the two types of error that can affect social science research: random error and non-random error (systematic bias). It is universally recognized that random error does not undermine the validity of a data set because random error, by definition, occurs in all directions. In the aggregate and over a large set of data, random errors are expected to cancel each other out. Although random errors may make the coding of a particular data point inaccurate, their aggregate effect over the whole data set is not expected to undermine conclusions.[20]

---

[20] For example, it invariably occurs that, during the data entry process, data processors sometimes incorrectly enter information from the coding sheet into the computer database. Although this would result in an individual data point being inaccurate, assuming that such misentries occurred randomly, they would not affect the aggregate results of the study.

Systematic bias, on the other hand, is error that always errs in one direction. The presence of systematic bias poses a serious threat to the validity of a study.

Although Professor Gibson's report is replete with accusations of error, he typically does not specify whether these alleged errors were random or systematic. For example, Professor Gibson asserts that the presence of incomplete or illegible visuals may have introduced error into the report. (Gibson, p. 7-8.) However, he never explains how such error would systematically bias the study's results. Similarly, Professor Gibson refers (without much further specification or explanation) to the alleged presence of "missing data" (Gibson p. 46), but never suggests that such missing data would introduce systematic bias in to the data set. And, he asserts that the merging of the coding data with the CMAG frequency data may have been the source of error, but never suggests how such error might affect the conclusions drawn from the data. (Gibson p. 24 n.27.) In short, it appears that the great majority of the errors that Professor Gibson alleges are, at most, the result of random "noise" which would not have systematically biased the study's results or undermined its validity.

Moreover, Professor Gibson makes no attempt to show that his subjective assertions about the alleged prevalence of error in the database are a valid basis for questioning the validity of the principal conclusions drawn from the database by the authors of *Buying Time 1998* or *2000*. For example, Professor Gibson does not dispute that the *Buckley* magic words are rarely used in political

advertising, or that group-sponsored ads that mentioned candidates tended to be concentrated before an election.

B. Peer Review.

Professor Gibson also claims that the fact that *Buying Time* "is not the product of any peer-review process . . . . seriously limits the confidence one can place" in the studies. (Gibson p. 4.) Although Professor Gibson is correct that the *Buying Time* studies themselves were not formally peer-reviewed, I have used CMAG databases in the past as the basis for a number of other scholarly articles, several of which have been published in the top-rank of peer-reviewed political science journals. I list some of these articles at page 9, above.[21] During the peer review process for these articles, none of the academic reviewers shared Professor Gibson's concerns about the validity or reliability of the CMAG databases.

Moreover, much of the CMAG data relied upon by *Buying Time 2000* was released to the public in real time during the 2000 election. Indeed, during both the 2000 and 2002 election cycles, the University of Wisconsin Advertising Project, which I head, has been the major source of data for journalists covering political advertising. Much of the data we report can cast the election strategies of particular candidates, parties or interest groups in an unfavorable light; there were many occasions on which it might have been in the political interest of those aggrieved candidates, parties or groups or to question

[21] Incidentally, this list includes articles published in both the *Journal of Politics* and the *American Journal of Political Science*, which Professor Gibson himself describes as among "the three most prestigious journals in our discipline." (Gibson p. 2)

the reliability or validity of our data. But at no time have we been challenged on the accuracy of the factual data we have reported on the content and targeting of political advertising.

The Brennan Center has already released the 1998 CMAG database to the public. We plan to release the full 2000 database, with complete documentation, in November of 2002.[22] This will not only allow any scholar to replicate any part of the analysis contained in *Buying Time*, it will allow them to carry out their own analyses of the data—including to completely redo the study, if they wish. I know of no other comparable social science data set that has released so much, to so many, so soon.

[22] Contractual obligations to CMAG have prevented the release of the 2000 micro-data and storyboards until two years after the 2000 election.

Kenneth M. Goldstein

10/14/02

Date