The effects of negative political advertisements: A meta-analytic assessment
Lau, Richard R;Sigelman, Lee;Heldman, Caroline;Babbitt, Paul
The American Political Science Review; Dec 1999; 93, 4; ProQuest Political Science
pg. 851

American Political Science Review Vol. 93, No. 4 December 1999

The Effects of Negative Political Advertisements: A Meta-Analytic Assessment

RICHARD R. LAU *Rutgers University*
LEE SIGELMAN *George Washington University*
CAROLINE HELDMAN *Rutgers University*
PAUL BABBITT *Rutgers University*

The conventional wisdom about negative political advertisements holds that no one likes them, but they work, that is, they have the consequences their sponsors intend. Moreover, many analysts have expressed concern over the detrimental effects of such negativism on the American political system. We examine the accuracy of the conventional wisdom and the legitimacy of the fears about the consequences for the political system via meta-analysis, a systematic, quantitative review of the literature. The data do not support either contention. Negative political ads appear to be no more effective than positive ads and do not seem to have especially detrimental effects on the political system. Eleven subsidiary hypotheses about particular circumstances in which significant effects are likely to be found are tested and rejected. Discussion focuses on why negative political advertisements have become so popular in practice when there is so little evidence that they work especially well.

Running for political office in the United States has never been prescribed for the faint of heart. Lord Bryce ([1888] 1995, 879) characterized late-nineteenth-century American campaigns as "thick with charges, defences [sic], recriminations, till the voter knows not what to believe." In the late 1970s and early 1980s, when political action committees targeted a number of congressional incumbents for attack, negative campaigning entered the current high-intensity phase of the "emotional sine wave" it has ridden over the course of American history (Gronbeck 1994, 61). The attack strategy quickly carried over into presidential campaigns, which induced a rise in the "negativism" trend line.[1] A prominent political consultant has summarized the new rules of political engagement as follows:

> 1. Advertise early if you have the money.... 2. Go negative early, often, and right through election day, if necessary. 3. Appeal to the heart and the gut, rather than to the head. 4. Define your opponent to the voters before he or she can define him-/herself or you. 5. If attacked, hit back even harder. 6. It's easier to give voters a negative impression of your opponent than it is to improve their image of you.... The best way to win is by bringing the other guy down, not by bringing yourself up (Kamber 1997, 46–7).

This upsurge in negativism is commonly attributed to television's increasing dominance of modern political campaigns. Television, it is said, has "granted the manufacturers of campaign discourse some Svengalian powers that print and radio lacked" (Jamieson 1992, 9). The mass media were once thought to have minimal political effects, largely reinforcing existing attitudes and commitments (e.g., Berelson, Lazarsfeld, and McPhee 1954; Lazarsfeld, Berelson, and Gaudet 1948), but television apparently has changed all that:

> When skillfully used, television's multiple modes of communication and powerful ability to orient attention can invite strong, unthinking negative responses in low-involvement viewers. And, by overloading our information-processing capacity with rapidly paced information, televised political ads can short circuit the normal defenses that more educated, more highly involved viewers ordinarily marshal against suspect claims (Jamieson 1992, 50).

Recognizing the medium's ability to transmit messages dramatically, rapidly, and widely, candidates for high office have increasingly built their campaign strategies around television in general and television advertising in particular. In the 1996 presidential campaign, for example, the Clinton and Dole forces pumped well over $100 million—more than 60% of their combined budgets—into electronic advertising (Chinoy 1997). But the high cost of television advertising means candidates want to make sure they get their money's worth, and evidently that translates into more and more negativism. According to political scientist Herbert Alexander, commenting on the June 1998 primaries: "The high cost of television means now that you have to go for the jugular," a tendency that leads him to predict "even more negativity than we've experienced now" (quoted in Purdum 1998, 4).

What are they getting for their money? Although some candidates avoid negative appeals altogether and

---

Richard R. Lau is Professor of Political Science, Rutgers University, New Brunswick, NJ 08901-1411. Lee Sigelman is Professor of Political Science, George Washington University, Washington, DC 20052. Caroline Heldman is a Ph.D. candidate, Department of Political Science, Rutgers University, New Brunswick, NJ 08901-1411. Paul Babbitt is a researcher, Center for Public Interest Polling, Eagleton Institute of Politics, Rutgers University, New Brunswick, NJ 08901-1411.

An earlier version of this paper was presented at the annual meeting of the American Political Science Association, Washington, DC, August 28–31, 1997. We thank Karen Hartman for helping locate relevant studies, and Allen Wilhite and the members of the ASPA panel and audience for their helpful comments on the original paper.

[1] The sharpness of this rise has been a matter of some dispute, as can be seen by comparing the trend lines estimated by Jamieson, Waldman, and Sherr (1998), Kaid and Johnston (1991), and West (1993).

others engage only in mild forms of "comparative advertising," many advertisements amount to "30-second snarls" (Will 1994) designed to "criticize, discredit, or belittle their opponents rather than promoting their own ideas and programs" (Ansolabehere et al. 1994, 829). What are the consequences of this barrage of negativism? Public discussion of negative political advertising has been dominated by two responses to this question—one short term and pragmatic, the other long term and despairing.

First, both practitioners of negative advertising and its harshest critics believe that it "works." Citing dramatic instances in which attack ads have been instrumental in turning a campaign around, such as Mitch McConnell's upset of Walter Huddleston in a 1984 Senate race and George Bush's come-from-behind victory over Michael Dukakis in 1988, campaign strategists portray negative advertising as a potent political force. For example, Republican pollster Richard Wirthlin contends that "a negative attack can take a virtual unknown against an apparently strong incumbent and provide a tremendous and strong margin" (quoted in "Negative Spots Likely to Return in Election '88" 1987, 3); and Democratic consultant Philip Friedman agrees: "The big question in most campaigns ... is whose negative campaign is better. If it's negative, it works. If it's positive, save it for your tombstone" (quoted in Henneberger 1994, 45).[2]

Second, negative political advertising is seen as a corrosive influence on participatory democracy, as "the electronic equivalent of the plague" (West 1993, 51). Beset by the unseemly spectacle of candidates doing "whatever it takes" to win (as George Bush vowed in 1992), citizens are repulsed. As one political consultant has observed, "in a campaign of negative ads fighting negative ads, what incentive is there for the viewer to go to the polls? Obviously, both these candidates are turkeys. The potential voter is left with a disgruntled sentiment that it's a shame someone has to win" (quoted in Laczniak and Caywood 1987, 21).

Is negative political advertising really as potent a political force as it has been made out to be, simultaneously shaping election outcomes and causing citizens to "tune out" and "turn off" on the political process? In theory, that is, in both social-psychological (Lau 1985) and rational choice (Riker 1997; Skaperdas and Grofman 1996) theory, negative advertising *should* work, and there clearly is no shortage of dramatic examples and expert testimony that it *does* work. Even so, the most sophisticated research to date on political advertising is only partially consistent with these two ideas. In a series of carefully controlled laboratory experiments, for example, Ansolabehere and Iyengar (1995; see also Ansolabehere et al. 1994) uncovered surprisingly little evidence that campaign ads shape voters' choice between opposing candidates. Yet, these experimental studies, buttressed by analyses of state-by-state voter turnout patterns in the 1992 Senate elections, did bear out the charge that negative advertising sours many citizens on politics, leading to a significant drop in voter turnout.

Why not consider these matters settled, or at least the idea that negative ads undermine public support for and participation in the electoral process? First, both practitioners and critics dwell on instances in which negative advertising is thought to have been decisive, but they tend to ignore counterexamples. The failed presidential campaigns of Bush in 1992 and Dole in 1996, Huffington's losing Senate bid in 1994, and the abortive gubernatorial races of Checchi and Christensen in 1998 serve as reminders that attackers do not always win.[3] Second, no single research study should be treated as definitive. Tellingly, reinterpretations of Ansolabehere and Iyengar's findings have begun to appear (Bartels 1996), as have new studies (analyzed below) that report contrary findings.

Given that political consultants and pundits should not be taken at their word when they testify to the potency of negative advertising, and that any particular research study should not be treated as authoritative, what is the appropriate next step? The usual response would be to propose yet another round of empirical studies, and well-designed new research on negative political advertising certainly would be welcome. We argue, however, that the highest immediate priority is to sift carefully through the available evidence to see what can be learned from data already at hand. In recent years a substantial body of research has accumulated on political advertising in general and negative political advertising in particular. What do these studies, taken as a whole, reveal? Does the preponderance of evidence bear out the claims about negative advertising that have so often been made and seem to be so widely accepted?

This question is not just of passing interest to casual observers of the American political scene. For those who believe that politics matters—that it makes a difference whether a Democrat or a Republican sits in the White House or in the governor's mansion, or which party controls Congress—knowing whether a popular campaign tactic "works" is important information. In the heat of a campaign, long-term qualms about the erosive effect of negative ads on participatory democracy may well give way to the immediate goal of winning an election. Yet, the charge that the increased negativism of campaigns is undermining the American political system cannot be taken lightly. Does anyone doubt that the stability of the American system of government itself might be threatened if turnout continues to fall and if cynicism toward those in office continues to rise? These are vital questions for political scientists to address.

These are also questions that political scientists *should* be able to answer. A criticism often leveled against the social sciences is that research findings do not accumulate. One set of researchers addresses a certain question with a particular method, another group addresses the same question with a different method, a third set addresses a related question with yet another method, and no conclusion is reached. Realistically, however, no one study can ever provide all the answers to any worthwhile scientific question, and no one method is superior in every situation. Science progresses when multiple researchers employ different techniques to explore an issue. Yet, there comes a time when we must take stock of what has been learned. To date, no comprehensive, systematic attempt has been made to review research on negative political advertising, to reconcile or even document contradictory findings, or to determine where the bulk of the evidence lies.[4]

---

[2] Perloff and Kinsey's (1992) survey of political consultants makes it clear that these opinions are well in keeping with the consensus among consultants.

[3] Nor, in principle, *can* attackers always win. When two opponents attack each other, as occurred in the North versus Robb Virginia Senate race in 1994 and the Torricelli versus Zimmer race for the New Jersey senate in 1996, only one can succeed (Lau, Pomper, and Mazeika 1995, 2).

To provide such a review, we have conducted a meta-analysis, an undertaking common in some fields but still rare in political science. Meta-analysis involves the statistical integration of research findings, usually from separate studies conducted independently of one another.[5] In a meta-analysis the findings themselves are treated as primary data, the goal being to establish the consistency and magnitude of the relationships in question. In the following sections, we describe the procedures we used to locate and integrate research findings, present the findings, and consider what these findings suggest about the validity of prevailing understandings of the consequences of negative political advertising.

METHOD

Locating Studies

The first step in a meta-analysis, as in any literature review, is a comprehensive inventory of the body of research. Our starting point was the large number of studies of political advertising accumulated by the two senior authors over the years. To these were added articles we identified by searching pertinent databases and documents, including *ABC Pol Sci*, *Communications Abstracts*, *Current Contents*, *Dissertation Abstracts*, *PsycINFO*, *Psychological Abstracts*, *Social Science Index*, *United States Political Science Documents*, and the programs for meetings of various professional associations. We also combed through the literature cited in each paper to identify additional studies that might contain pertinent findings. Our goal was to access every relevant study. We believe that we have located and analyzed the great majority of them.

Although these search methods are excellent for locating published studies, they are less effective for unearthing unpublished research. It is well known that published studies tend to be biased toward over-reporting of statistically significant results (see, e.g., Begg 1994), so a literature review that underrepresents "fugitive" studies runs the risk of overestimating the true effects of the phenomenon in question. To minimize such bias, we (1) included in the meta-analysis pertinent unpublished papers and (2) contacted the authors of studies we located, described our project, explained the need to consider *all* applicable findings (regardless of the statistical significance of the results), and requested relevant papers. In response, we received a score of new papers, several of which met our criteria for inclusion and were incorporated into the meta-analysis.

Criteria for Inclusion

Every paper, article, chapter, or book in our inventory was initially screened by one of the authors to determine whether it contained findings that met each of five criteria.

1. A focus on *negative* political advertising. We included findings in the meta-analysis only if they pertained specifically to negative political advertisements rather than generically to political advertisements without regard to valence. Recognizing that "negative advertising" is a contested concept, we based this criterion on whether the authors of a given study themselves categorized an ad as negative. Excluded by this criterion was, for example, Brians and Wattenberg's (1996) finding that those who recalled watching more campaign ads in 1992 knew significantly more about the candidates' issue positions, because that finding says nothing about the effects of *negative* ads in particular. Because the findings included in the meta-analysis were based on several different types of negative ads, we undertook a follow-up analysis, described below, to ascertain whether certain types of negative ads, e.g., "comparative" rather than pure attack ads, may be more effective and other types less so.
2. A focus on negative *political* advertising. Dozens of studies have been undertaken of negative *product* advertising (e.g., Muehling, Stoltman, and Grossbart 1990; Putrevu and Lord 1994), but we concentrated exclusively on findings concerning advertising in election campaigns.
3. A focus on negative political *advertising*. We included studies only if they contained an explicit advertising element. This criterion disqualified, for example, findings about the effect of scandals on the outcome of congressional elections (e.g., Dimock and Jacobson 1995; Welch and Hibbing 1997) and of candidate attacks in campaign debates (Roese and Sande 1993).
4. A means of gauging the *effects* of negative political advertising. This criterion eliminated studies that did not contain an outcome measure (e.g., Hale, Fox, and Farmer 1996; Kaid and Johnston 1991; West 1993). It also eliminated analyses lacking an element of *comparison*, because if a study focused exclusively on negative ads, then the possibility

---

[4] Lin (1996) provides a descriptive overview of research on negative political advertising, and Hale (1998) presents a meta-analysis of a small subset of the studies reviewed here.

[5] Standard sources on meta-analysis include Cooper and Hedges 1994; Glass, McGaw, and Smith 1981; and Hunter and Schmidt 1990.

could not be ruled out that the ostensible effects of these ads might hold for positive ads as well. Thus, to be included in the meta-analysis a study had to compare negative advertisements to something else—either a "no advertisements" control group or a "positive advertisements" comparison group. This criterion eliminated some worthwhile studies of the relative effectiveness of different types of negative ads (e.g., Budesheim, Houston, and DePaola 1996; Karrh and Halpern 1997).

5. *Single counting* of a given finding. When we located nonindependent reports of the same finding (e.g., in a convention paper and in a published version of the same paper), we included only the later, and presumably more authoritative, version. By the same token, if one report included a subset of data also incorporated into a more comprehensive data set analyzed and reported upon elsewhere, we excluded the former and focused on the latter. For example, we excluded the experimentally based findings reported by Ansolabehere et al. (1994) because the 1,655 subjects on which these findings were based were a subset of the 2,216 subjects in a more comprehensive presentation that we did include (Ansolabehere and Iyengar 1995). Yet, a data set could provide multiple findings for the meta-analysis if more than one type of outcome measure was involved (e.g., subjects' evaluations of both the sponsor and the target of an ad).

After this screening, one of the senior authors reexamined each excluded study to ensure that pertinent findings had not been inadvertently eliminated. Thus, every item in the original inventory was read at least twice before being dropped from consideration.[6]

Recorded Variables

We recorded a wide array of descriptive information about the researchers and the research design associated with each finding. For present purposes the most critical information was the operational definition of "negative advertising" employed by the researcher and the category into which a given outcome measure fell. Both of these issues relate to an "apples and oranges" criticism often leveled at meta-analysis. Studies made it into our review because the authors stated that their study involved negative advertising. Yet, just as the actual negative (and positive) advertisements used by real candidates vary widely, so do the conceptual and operational definitions of negative advertisements used in these studies. What is the point, one might ask, of trying to summarize a diverse literature by treating findings based on different types of ads as though they were indistinguishable?

This question misses one of the greatest strengths of meta-analysis. Diversity is not a problem in meta-analysis as long as such diversity can be coded and taken into account in the analysis. In defining different types of campaign ads, the crucial factor seems to be how much negative information an ad must contain before it is called "negative." Whereas some researchers treat as negative any ad that mentions the opponent, others distinguish among ads that mention only the sponsor (positive or advocacy ads), ads that focus exclusively or primarily on the opponent (negative or attack ads), and ads that focus on both the sponsor and the opponent (comparative or contrast ads) (Jamieson, Waldman, and Sherr 1998). We coded this information for every finding in our analysis. In our initial presentation of results, we ignore these differences, taking at face value each research team's claim to be studying negative political advertising. Subsequently we take these potentially crucial differences in definition of the independent variable into account to see whether they have any bearing on the results.[7]

We divided outcome measures into 13 inductively based categories: (1) affect for an ad itself; (2) affect for the target of an ad; (3) affect for the sponsor of an ad; (4) differential affect for the target and the sponsor; (5) intention to vote for the target or the sponsor; (6) actual vote for the target or the sponsor; (7) memory for the ad; (8) intention to vote; (9) actual vote turnout; (10) trust in government; (11) political efficacy; (12) knowledge about the candidates running in an election; and (13) "public mood." Of these categories, 2 through 7 pertain to the *intended consequences* of political ads. That is, a successful negative ad should reduce affect for the target (the opponent of the ad's sponsor); not reduce affect for the sponsor; or at least increase affect for the sponsor relative to the target; enhance the likelihood of voting for the sponsor rather than the target; and convey a memorable message.[8] Categories 8 through 13 pertain to *unintended consequences*, that is, potential effects not directly related to the voter's choice between the sponsor and the target; the measures in these categories gauge the systemic effects of negative political advertisements—their influence on voter turnout, trust in government, and the like. We present our results separately for these different categories of dependent variables.

The remaining data required for the meta-analysis (in addition to the effect size for each finding, as explained below) were the number of experimental subjects or survey respondents on which a given finding was based, which was easy to determine in most cases; the reliability of the outcome measure, which we had to estimate in many instances; and the strength of the political advertising "treatment" (i.e., the number and the nature of political ads to which subjects were exposed). Sample size, measurement reliability, and strength of treatment come into play when effect sizes are adjusted for sampling error, measurement error, and variation in treatment effects, as described below and detailed in the Appendix.

---

[6] We have a supplementary table that lists another 199 papers examined for this analysis, along with a brief explanation for their exclusion. This table and a comprehensive bibliography of articles on negative advertising are available for examination at the APSR web site.

[7] We also recorded a wide array of information about each study, including inter alia, the disciplinary affiliations of the researchers; the nature of the subjects (students, general public, etc.) and of the advertisements (video and audio, written text, etc.); the substantive focus of an advertisement (issue or image); whether the content of the negative ads was directly coded or manipulated by the researcher or was inferred from secondary sources; the office being contested; whether actual candidates were used in the advertisement; the gender, race, and incumbency status of each candidate; and so on. In theory, each of these variables could be used as a moderator variable in the meta-analysis, but in practice the number of findings was too small for us to consider more than a few of them; more important, our primary goal was a bottom-line assessment of the effects of negative ads. We consider these additional variables below, after presenting the main results.

[8] For a critique of recall as an indicator of the effectiveness of an ad, however, see Haskins 1964.

Calculating Effect Sizes

The problem of different outcome measures is resolved by calculating "effect sizes" that translate results based on different measures into a common metric. There are two major varieties of effect sizes, *r* measures (based on correlations) and *d* measures (based on mean differences). We chose the latter because most of the findings analyzed here were derived from comparisons of group means.

In the simplest case, the effect size measure is defined as:

$$d = (\bar{X}_E - \bar{X}_C)/S_y.$$

That is, *d*, the measure of effect size, is the difference between the means of the experimental and control groups, divided by the standard deviation. Thus, *d* is expressed in standard deviation units, closely paralleling *z*-scores.[9] Surprisingly often, researchers fail to report the information involved in the formula for *d*, even when the design of a study is as simple as a comparison between an experimental and control group. Generally, however, if even a modicum of more or less exact information (e.g., group means and an *F* statistic) is reported, an algebraic path can be followed to a reasonable approximation of an effect size.[10]

About one-quarter of our data points come from ordinary least squares (OLS) or logistic regression equations, and there is no universally accepted method for handling such data in a meta-analysis (see the Appendix). In general, we proceeded along two lines. The first was to ignore the magnitude of an effect and consider only its statistical significance (Becker 1994). This nonparametric "combined significance" approach is quite conservative, for it tests only the very specific null hypothesis that the effect of interest is not present in *any* of the populations studied. Second, we used the *t* value associated with a regression coefficient as a parametric estimate of the magnitude of the effect (Stanley and Jarrell 1989) and combined it with the effect sizes calculated from other studies in the analysis. This approach is much more informative about the magnitude of an effect and permits more sophisticated hypothesis tests, while considering data from all relevant studies. This point is discussed further in the Appendix.

Adjusting for Errors and Bias

After calculating an effect size for each finding in the meta-analysis, one has to decide how to combine them. Some experts recommend analyzing raw, "unadjusted" effect sizes, while others advocate performing a variety of adjustments prior to analysis. The underlying issue is whether to treat all studies equally. For a large and fairly homogeneous research literature, it could well be appropriate to treat all studies equally, but for a literature as diverse as the one we are considering, this would be problematic. Accordingly, we followed the adjustment guidelines established by Hunter and Schmidt (1990). In what follows we report unadjusted parametric effect sizes, followed by effect sizes adjusted first for sampling error (that is, sample size), then for unreliability of measurement in the dependent variable, and finally for variation in the strength of the independent variable. Our conclusions were not greatly affected by these adjustments, which are described in greater detail in the Appendix.

RESULTS

In all, we located 117 pertinent findings reported in 52 different studies. The Appendix identifies the studies and describes the design of each, the main variables in the design, and the pertinent findings. Table 1 summarizes the statistical analyses of the unadjusted effect sizes, and Table 2 summarizes the results for the adjusted effect sizes. We present these results in three sections, based on the nature of the outcome measure: affect for the negative ads themselves, the intended consequences of negative ads, and the unintended consequences of negative ads.

Affect toward Negative Political Ads

If there is one point on which virtually everyone seems to agree, it is that no one really likes negative political ads.[11] The data analyzed here do not speak directly to this point, for findings concerning evaluations of negative ads were not coded against an absolute "neutral" point. Yet, these data do enable us to assess the closely related claim that negative ads are liked less than positive ads.

---

[9] The formula for *d* becomes less straightforward in research involving factorial designs but is algebraically equivalent. These formulas can be found in the standard sources on meta-analysis cited above.

[10] For example, only significance levels were reported in some studies. In such instances, we adopted the following conventions. If the *p* value was reported as $<.01$, we assumed that $t = 2.75$; if the reported *p* was $<.05$, we assumed that $t = 2.25$; if the effect was "not significant" but sufficient information was provided to determine its direction, we assumed that $t = \pm 1.0$; and if the effect was "not significant" and no indication was given for the direction of the effect, we assumed that $t = 0$. These approximations assumed sample sizes of $N \geq 100$, in which case the *t* distribution closely approximates the normal distribution. For smaller sample sizes we adjusted the presumed *t* value upward.

[11] The idea that voters profit from hard-hitting presentations of the differences between candidates does have its defenders (e.g., Mayer 1996). It is not the sheer existence but the seeming ubiquity of negativism and the excesses to which it is often carried that excite the greatest criticism.

TABLE 1. Summary of Unadjusted Parametric Results

Dependent Variable	Number of Findings	Mean	Median	Range	Standard Error
Affect for Ad Itself[a]	10	−.44	−.34	−3.12 to 1.01	.35
Intended Effects					
Memory for ad[b]	14	.34	.17	−1.15 to 3.86	.32
Affect for target[c]	16	.38*	.34	−.48 to 1.90	.16
Affect for sponsor[d]	25	−.51**	−.35	−2.05 to .75	.13
Differential affect[e]	4	−.87	.10	−4.38 to .72	1.19
Vote intention[f]	13	−.06	.00	−2.40 to 1.77	.29
Actual vote[g]	5	−.24	−.17	−.58 to −.06	.09
Unintended Effects					
Intended turnout[h]	4	−.04	−.05	−.18 to .12	.07
Actual turnout[i]	15	−.06	.02	−1.39 to .49	.11
Other systemic effects[j]	11	−.21	−.08	−1.45 to .14	.13

*$p < .05$, **$p < .01$.

[a] The studies in Table A-1 are numbered consecutively. Data relevant to affect for the ad itself come from studies 14, 15, 24, 35, 37, 40, 42, 44, 48, and 49.

[b] Data on memory for ads come from studies 4, 5, 15, 22, 23, 24, 25, 27, 32, 33, 39, 42, 43, and 44.

[c] Data on affect for the target of negative ads come from studies 4, 6, 13, 14, 20, 21, 24, 27, 29, 32, 35, 36, 40, 42, 43, and 52.

[d] Data on affect for the sponsor of negative ads come from studies 4, 6, 13, 14, 15, 16, 17 (two data points), 20, 22, 24, 27, 29, 30, 32, 34, 35, 36, 40, 41, 42, 43, 44, 49, and 52.

[e] Data on differential affect come from studies 10, 26, 47, and 52.

[f] Data on vote intention come from studies 1 (two data points), 6, 20, 22, 24, 30, 34, 36, 40, 41, 42, and 44.

[g] Data on actual vote come from studies 7, 26 (two data points), 46, and 51.

[h] Data on intended turnout come from studies 1, 22, 31, and 37.

[i] Data on actual turnout come from studies 2, 8 (two data points), 9, 10, 11 (two data points), 12 (two data points), 18, 26, 28, 31, 45, and 50.

[j] Includes two analyses of "public mood" (studies 38 and 45), four of political efficacy (studies 1, 9, 12, and 45), two of trust in government (studies 29 and 45), and three of knowledge about the candidates running in an election (studies 3, 19, and 45).

If affect is markedly lower, on average, for negative than for positive ads, then the average size of the ten pertinent effects should be well below zero as we have coded the data; that is, subtracting affect for positive ads from affect for negative ads should produce a negative effect size estimate. The evidence for the often-asserted abhorrence of negative political ads turns out to be surprisingly weak, however. When the combined significance approach is followed, the null hypothesis that no significant effect exists can be rejected ($z = -9.43$, $p < .001$), but this is due almost entirely to one extremely significant result. When we take a parametric approach (as shown in the top of Table 1), even allowing for this one extreme negative value (−3.12), the mean uncorrected effect size for these ten outcome measures is only −.44. This signifies an average difference of only about four-tenths of a standard deviation between affect for positive and negative political ads. This is in the direction that would be expected based on the frequent expressions of disgust with negative ads. Yet, the variability of the ten effect sizes sounds a caution against interpreting this difference as consistent with expectations. More formally, the standard error is nearly as large as the mean effect size itself, and the 95% confidence interval for the unadjusted mean effect size extends well into

TABLE 2. Summary of Adjusted Parametric Results

			Corrected for Sampling Error		Corrected for Attenuation Due to Measurement Error		Corrected for Sampling Error, Measurement Error, and Variation in Strength of IV	
	N of Studies	Total *N* of Subjects	Effect Size	Standard Error	Effect Size	Standard Error	Effect Size	Standard Error
Affect for Ad Itself	10	1,580	−.52	.44	−.61	.51	−.63	.63
Intended Effects								
Memory for ad	14	7,529	.88	.69	1.11	.79	.55	.49
Candidate-centered intended effects	35	14,458	−.14	.09	−.16	.11	−.15	.12
Unintended Effects								
Combined	22	45,948	.05*	.02	.07**	.05	.04	.05
Just turnout	17	44,644	.03	.07	.04	.08	.02	.09

Note: This analysis aggregates across multiple findings (within category) reported by any study. *$p < .05$, **$p < .01$.

positive territory (-1.23 to $+.35$). Thus, considered as a body of evidence bearing on the issue of affect for negative ads, *the research literature provides no reliable statistical basis for concluding that negative ads are liked less than positive ones.* Adjusted for sampling error, measurement unreliability, and variation in treatment strength (shown in Table 2), the mean effect size increases to $-.63$, but the standard error also rises, to .63. For the adjusted means, the 95% confidence interval runs all the way from -2.06 to $+.80$. For neither the unadjusted nor the adjusted effects, then, can we safely reject the null hypothesis of no difference between negative and positive political ads. It may be that most people do not like campaign ads very much, but whether a political advertisement is positive or negative seems not to be the crucial factor. Or it may be that the widespread hand-wringing about negative ads is largely ritualistic or is focused on a few negative ads that are seen as excessive rather than on negative ads as a genre. If anything, this initial result should be encouraging to those who employ negative ads despite the "fact" that people dislike them. According to the evidence that we have located, this "fact" is considerably overstated.

Intended Consequences: The Effectiveness of Negative Political Ads

If negative political advertising works, it should have positive consequences for its sponsors and negative consequences for its targets. More specifically, negative messages should be more memorable than positive ones; they should cause affect for the opponent to decline; they should, at the very least, not greatly deflate affect for the sponsor; they should have a net positive effect on evaluations of the sponsor relative to those of the opponent; and, most important, they should increase the probability of voting for the sponsor rather than the opponent.

The evidence detailed in the Appendix does not bear out these suppositions. Most of the effect sizes fall very close to the zero point, and about as many are below as above zero. In fact, under the nonparametric combined significance test the null hypothesis that the hypothesized effect is not found in any of the populations studied cannot be rejected ($z = -1.10, p > .20$)—and the bulk of the findings are in the direction *opposite* of what the sponsors of these ads intend. Of the findings for the specific dependent variables in this category in Table 1, the only two with mean positive effect sizes (i.e., the only two that run in the hypothesized direction) are memory for the ad and affect for the target of the ad, and only the latter is statistically significant. That is, it does appear that the sponsoring candidate's opponent is liked less when s/he is attacked by political advertising. Even so, this intended effect is counterbalanced by an even stronger and highly significant decrease in liking of (that is, a "backlash" against) the sponsor, an effect that sponsors of such ads certainly do *not* want to achieve.

We collapsed the multiple indicators of intended effects into two broad categories before adjusting for sampling error and so on: findings about the ad itself (i.e., memorability), which are presumably a secondary consideration for those who run the ads, and findings about the candidates (the rest of the findings in this group).[12] The mean unadjusted effect size is $-.28$, counter to what the sponsor of an ad would want. When the estimates are adjusted for sampling and measurement error and variation in treatment strength (Table 2), the mean candidate-centered effect shrinks somewhat but is still negative, albeit not statistically significant; zero is easily encompassed by the 95% confidence interval, which ranges from $-.39$ to $+.09$. The average corrected effect size for memory of the ads rises as high as 1.11, but the standard error rises commensurately, and this moderately high mean effect is still not different from 0, with the 95% confidence interval ranging from $-.51$ to 1.61. In sum, across these multiple criteria, *there is simply no evidence in the research literature that negative political advertisements are any more effective than positive political ads.*

Unintended Consequences: Do Negative Ads Damage the Political System?

Finally, we turn to the claim that negative political ads sour citizens on politics in general and on campaigns in particular. If this is true, then the effect sizes for the "unintended consequences" outcome measures should be well below zero, as we have coded the data; that is, the unintended consequences of negative ads should be more negative than those of positive ads.

Of the findings included in the meta-analysis, thirty bear on this claim. For the nonparametric combined significance approach, the null hypothesis of no significant effect in any of the populations studied cannot quite be rejected with a one-tailed significance test, which is appropriate here ($z = -1.07, p < .07$). The sample sizes of the thirty data points tend to be very large, however, so it is not surprising that some of them are significant. More telling is the pattern of results: Fifteen are negative, including five significant results, but fourteen are positive, including two significant findings. When these findings are adjusted for sampling error, unreliability of measurement, and treatment strength in Table 2, the mean effect size is slightly greater than 0. The story does not change if we limit the analysis to studies of intended and actual turnout.[13]

---

[12] Most studies in this literature report multiple findings that are relevant to the meta-analysis. When those findings fell into distinct categories, we treated them separately. When multiple findings involving the same set of subjects were in a single category, however (as is the case for the candidate-centered intended consequences), we averaged across the multiple effect sizes that were based on the same set of subjects to arrive at a single effect per data set. Occasionally a study reports multiple relevant findings from distinct sources or data sets (e.g., Ansolabehere and Iyengar 1995), and in such cases we retained the distinct findings. Thus, the 63 individual candidate-centered effects summarized in Table 1 reduce to 35 effect sizes in Table 2.

[13] The best-known findings, those reported by Ansolabehere and Iyengar (1995), are not far removed from a mean effect size of zero. Ansolabehere and Iyengar found that viewing a single negative ad

In sum, *we uncovered little evidence to warrant the fears of those who believe that electoral participation is imperiled by the increasingly widespread use of negative political advertisements.* Participatory democracy may be on the wane in the United States, but the evidence reviewed here suggests that negative political advertising has relatively little to do with it.

Follow-Up Analyses

At this point we need to consider three potential problems that, if present, could undermine the results reported above. First, are these results unduly influenced by a few findings based on atypically large samples? This seems most likely for the "unintended consequences" findings, a few of which involve an extremely large number of cases. To probe this possibility, we first counterfactually assumed that the three studies with extremely large sample sizes (Finkel and Geer 1998: N = 12,252; Geer and Lau 1998: N = 8,069; Kahn and Kenny 1998b: N = 6,110), which together accounted for almost 58% of the total cases in this category, had "only" half their actual total of 26,431 cases. This caused the adjusted means reported in Table 2 to decrease slightly, but they remained positive. Then we reduced the sample sizes even more, by a factor of five, and halved the recorded sample sizes in the other six studies with more than 1,000 cases. This reduced the adjusted means slightly more, but they remained above zero, and our conclusions were left essentially unchanged. Thus, our results seem fairly insensitive to extreme values, for the estimated effect size remained small even when we "fixed" the large-sample problem.

Second, how would our conclusions change if findings that have escaped our notice were taken into account? Earlier we referred to a bias toward statistical significance in published studies. To cope with this problem, Rosenthal (1979) devised a technique for estimating how many extra findings with effect sizes averaging zero would have to be located to reduce an observed significant mean effect size to nonsignificance (see also Hunter and Schmidt 1990; Orwin 1983). Here, we face the opposite problem. Having uncovered mostly nonsignificant mean effects for the "intended consequences" measures in particular, we are concerned that large-effect findings may exist which have not made their way into the published literature. It is conceivable, for example, that some political consultants possess such evidence but have not published it or presented it publicly.

To pursue this possibility, we reversed Rosenthal's logic by calculating how many such findings would have to exist to convince us to alter our conclusions. In light of the observed range of effect sizes for "intended consequences" outcome measures, we would classify an effect size of +1.00 as very strong—stronger, in fact, than any we actually observed. Assuming that the newly available effect sizes averaged +1.00 and were based on the same average number of cases, measurement reliabilities, and combined variances as the findings considered here, at $\alpha = .05$ it would take *25 of them*, not counterbalanced by any contrary new evidence, to increase the mean unadjusted effect size significantly above zero. Given the general bias toward significance in published research, it seems much more reasonable to assume that additional *small*-effect findings exist that so far have escaped our notice than that we have missed so many sizable effects. All in all, it would take a mass of very strong and entirely uncontradicted new evidence to lead us to conclude that negative political ads "work." Because the possibility that such evidence exists seems remote, we place considerable confidence in our conclusion that negative political ads fall significantly short of achieving their intended results.[14]

Third, would stronger effects emerge if we were to disaggregate the data? In aggregating findings across studies, have we ignored crucial differences among studies that, if taken into account, would bring stronger effects to the surface? The best context for addressing this issue is provided by findings on the intended consequences of negative ads, which are extensive enough to allow some subsidiary hypothesis testing. We considered eleven possibilities.

1. Perhaps, to follow up on a point raised earlier, the definition of "negative ad" matters, such that entirely negative ads are no more effective than entirely positive ads, but comparative or contrast ads are most effective. From a rhetorical point of view, a case has been made that comparative or contrast ads are optimal for effective electoral decision making (Jamieson, Waldman, and Sherr 1998). Empirically, however, it makes no difference whether entirely negative or more comparative political ads are contrasted to positive ads, $t(38) = -.73$, ns.

2, 3. Perhaps design quality has increased over time, in which case more recent studies should be more likely to report, correctly, that negative ads are more effective. Or perhaps, just as military leaders often rely on outdated strategies because they worked so well in the last war, negative ads, although once effective, have become so commonplace that they have lost their effectiveness (Lau 1985), in which case there should be a negative correlation between time and effect size. Contrary to both of these ideas, however, the correlation between year and effect size ($r = -.11$) does not differ from zero.

4. Perhaps experiments, because they tend to be artificial, produce unrealistic results, while sur-

---

(compared to viewing a single positive ad) decreased intention to vote in the upcoming election by about 4.5%. With more than 2,200 subjects in the analysis, this effect is statistically significant ($p < .05$) but very "noisy," and it translates into an effect size of only $-.10$, that is, one-tenth of a standard deviation.

[14] With the same set of assumptions, it would take four additional studies of the effects of negative political advertising on turnout with average effect sizes of -1.0 to reverse our conclusion about Ansolabehere and Iyengar's "demobilization" hypothesis, that is, negative advertising, if anything, increases turnout.

vey-based studies produce the expected results. Yet, a t-test contrasting the effect sizes of findings based on experiments versus survey-based findings reveal no significant differences, $t(36) = 1.07$, ns.
5. Perhaps actual ads and/or ads featuring real candidates are more likely to be effective than ads created solely for research purposes. The observed differences are small ($ts < 1$), however, and run in the wrong direction.
6. Perhaps studies that either manipulate or directly code the contents of political ads are better indicators of the true strength of negative ads, compared to studies in which the "negativism" of the ads is inferred from secondary (e.g., newspaper) accounts. Yet, the observed differences are small in magnitude ($t(38) = .57$, ns) and run counter to this hypothesis.
7. Perhaps findings from studies based on student samples are idiosyncratic, while studies that use adult subjects and representative samples produce the expected results. The use of student samples, however, does not significantly affect the effect size, $F(3,36) = .77$, ns.
8. Perhaps televised negative ads work better than printed counterparts. Yet, studies using video ads as stimuli produce no greater effects than studies that present ads through other media, $t(38) = -.80$, ns.
9. Perhaps researchers in certain disciplines are more sensitive to the context of a campaign and thus are more likely to uncover the predicted positive effects. A one-way analysis of variance, however, contrasting studies published in the fields of communications, political science, and psychology uncovers no such differences, $F(2,37) = .05$, ns.
10. Perhaps better-designed studies are more likely to produce the expected results, but there is no significant difference in effects between studies we subjectively categorized as of higher quality and those we considered of lesser quality, $F(2,37) = 1.47$, ns.
11. Perhaps exposing subjects to a larger number of ads produces greater effects. There was no correlation, however, between effect size and the number of ads to which subjects were exposed ($r = .09$, ns).[15]

Having considered and rejected all these possibilities, we conclude that if there is some critical factor that must be present for the expected effects of negative advertising to emerge, we have not been able to identify it.

DISCUSSION

Our synthesis of findings reported in the literature does not bear out the main claims made about the effects of negative advertising. The great majority of the effects reported are of modest magnitude, with effect sizes clustered in a narrow band that extends from slightly above zero to slightly below zero. We observed no significant tendency for negative ads to evoke lower affect than other campaign ads, contrary to the common contention that citizens reserve special disdain for negative ads. As noted earlier, this does not mean that negative political ads are well liked. Indeed, there is abundant evidence that they are not; for example, 75% of those interviewed in a 1994 poll said they were "turned off" by negative ads (Brack 1994). Rather, it simply means that, according to the available evidence, negative political ads are not disliked significantly more than other political ads or, for that matter, than ads in general. In an era when majorities or substantial minorities of adult Americans consider television advertising unhelpful, unbelievable, and misleading, and respond by leaving the room, attending to chores, or channel-surfing during commercial breaks (Mittal 1994), should the unpopularity of negative political ads be considered especially noteworthy?

More important, we did not uncover consistent, let alone strong, evidence that negative ads work to the advantage of their sponsors and/or the disadvantage of their targets. In this respect, it appears that, à la Newton's third law, for every research finding there is an equal and opposite research finding. Only a handful of the positive effect sizes we cataloged are large, and these are counterbalanced or even overbalanced by another handful of effect sizes on the negative side. There simply is no compelling evidence that negative advertising "works."

Of course, the effects of negative campaign ads need not be statistically significant in order to be politically significant or even decisive. Even a tiny advantage to the sponsor can be enough to determine the outcome of a close election, and even an attack that fails to sway voters can cause the target to divert precious resources in order to answer it. The results of our meta-analysis should not, then, be interpreted as saying that negative advertising is invariably a poor tactic. In general, however, negative campaign ads appear to be no more effective than positive campaign ads and even somewhat less so. Thus, while we concede that a well-conceived negative advertising campaign can be a key to electoral success, the same can be said, and with somewhat greater confidence, about a well-conceived positive advertising campaign.

Finally, our meta-analysis also fails to confirm the widely held view that negative advertising should bear a major share of the blame for the widespread political disaffection of recent decades. The effects we observed for the "unintended consequences" measures are too small in magnitude and too mixed in direction to provide empirical support for heated claims that negative ads are undermining public confidence and participation in the electoral process. We should note, however, that all the studies analyzed here focused on the immediate or short-term effects of viewing negative ads rather than on the long-term consequences of being subjected to a continuing barrage of such ads.

---

[15] The statistics reported in this paragraph are for unadjusted effect sizes. None of the conclusions changed when we examined the adjusted effect sizes. (See the Appendix for further explanation of these adjustments.)

A quarter of a century ago, McCombs and Shaw (1972) attributed the prevailing lack of understanding of the effect of political advertising to the dearth of research on the subject. Since then, a great deal of research has been completed, but if the findings reported here are to be believed, widespread misunderstandings remain, at least in the form of overly expansive claims about the effects of negative ads.

Why are claims about the effectiveness of negative advertising so far removed from the findings reported in the literature?[16] Part of the answer is undoubtedly that until the publication of Ansolabehere and Iyengar's *Going Negative*, academic research had hardly dented the consciousness of those who shape public discourse concerning negative advertising. Also, campaigners, consultants, and pundits are not immune to a wide array of perceptual and attributional biases (see, e.g., Fiske and Taylor 1991; Kahneman, Slovic, and Tversky 1982). We cite three such biases.

First, in campaigns in which both sides go on the attack, the well-known tendency toward internal attributions of success and external attributions of failure can lead winners to credit their own "brilliant campaign strategy" and losers to blame their opponents' "vicious attacks." Both claims bolster the impression that negative advertising "works," even though it obviously did not work for the losers.

Second, a different bias that produces the same result is the tendency to overgeneralize from a vivid example that is easily retrieved from memory. For instance, a reasonably well-informed American, if asked to name a presidential campaign in which negative advertising was especially prominent, might well mention the 1988 race, with its images of revolving prison doors and Michael Dukakis riding in a tank. Of course, in both presidential campaigns since then, the main attacker (Bush in 1992, Dole in 1996) lost, but it is the vivid exception—1988, when Bush's attack ads "worked"—that probably would spring to mind, which forges an illusory correlation between going negative and winning.

Third, more broadly, people often misperceive, reinterpret, or ignore information that is inconsistent with their preconceptions, and any or all of these tendencies can lead candidates, consultants, journalists, and political reformers to exaggerate the effectiveness of negative political advertising.

Which, if any, of these accounts is most accurate must remain a matter for speculation at this point. What should not be seen as speculative is the conclusion that prevailing understandings of the effects of negative political ads are in need of fundamental rethinking—a process already heralded by the recent spate of research on negative political advertising. Contrary to ideas that are currently widely accepted, the literature provides no significant support for the suppositions that negative political ads are especially disliked, are especially effective, or substantially undermine public support for and participation in the electoral process.

APPENDIX

The studies reviewed in this article are listed in Table A-1. For each study, we briefly describe the nature of the crucial independent variable; the subjects and basic design; the dependent variable(s); and the results, including our estimate of the raw effect size d.

Most of the findings in our review involve mean differences (usually from experiments), for which the formula for calculating d presented in the text was used. A few come from OLS or logistic regression analyses, however, for which there are no universally accepted means of translating coefficients into effect sizes. The problem is twofold. First, different measurement scales in the independent or dependent variables produce regression coefficients of different magnitudes. Second, even when the focal independent and dependent variables are measured on the same scale, some experts insist that regression coefficients not be compared unless all other variables in the equation are identical. In practice, this requirement is almost never met, in which case all such data must be excluded from a meta-analysis. We rejected this approach out of hand.

A second approach is to ignore the magnitude of an effect and consider only its sign and statistical significance. This nonparametric approach was one of the first meta-analytic techniques to find its way into the statistical literature (Fisher 1932), although it has been greatly refined since then. Becker (1994) describes several "combined significance" tests, of which we employ one of the simplest, converting reported significance levels to their normal (z value) equivalent (Mosteller and Bush 1954). When the null hypothesis is true, the sum of z values is normally distributed; it is divided by its standard deviation, $\sqrt{k}$ (where k is the number of studies); and the ratio is compared to the critical values in the standard normal table.

A third approach is to devise parametric estimates of effect sizes from the regression coefficients and then combine them with the effect sizes computed from the remaining studies. Although a regression coefficient is dimensional (i.e., the measurement scales of the independent and dependent variables affect its magnitude), the t statistic associated with a regression coefficient has no such dimensionality but is a standardized measure of the effect of interest (Stanley and Jarrell 1989; see also Kanetkar et al. 1995; Phillips and Goss 1995; Raju, Pappas, and Williams 1989). Indeed, even when research designs contrast differences in means, it is often the case (when the means are not actually reported) that the effect size is calculated from a t statistic (using the standard formula $d = 2^{*}t/\sqrt{df}$). We used that formula to calculate effect size estimates from findings presented in a regression format.

Although the statistical reasoning seems clear, it may be useful to provide an intuitive justification for this practice. In an experiment, the effects of all possible "third" variables on the dependent variable are assumed to be controlled by random assignment. In the typical research design in which some form of regression is employed, however, subjects have not been randomly assigned to conditions. Instead, the effects of possible third variables are controlled statistically by being included in the regression analysis along with the focal independent variable. As long as a regression model is reasonably well specified, we argue, it provides a situation sufficiently similar to the typical experiment so that the effects of comparable independent variables—whether they are manipulated in an experiment or measured in a survey—can be compared across research designs.

---

[16] Interestingly, although advertising practitioners perceive comparative ads for commercial products as effective and are making increasingly frequent use of them (Rogers and Williams 1989), research indicates that comparative ads are generally no more effective than other ads at generating favorable attitudes toward products and intentions to purchase them (e.g., Barry 1993). The result is a gap between research and practice in product advertising that closely parallels what we have observed for negative political advertising.

Adjusting Estimated Effect Sizes

If the population of interest is not constant across studies, the best estimate of the population effect size, Δ, is not the simple mean effect size across studies (as in Table 1) but a weighted mean that corrects for sampling error, with weights determined by the proportion of all participants who were in a given study.[17] Thus, our initial adjustment to effect sizes was to weight each mean by N_i/N_T, where N_i is the sample size for a study, and N_T is the total number of participants across all studies. If D is the mean unadjusted effect size, and d_i is the unadjusted effect size for each study, then D_1 is the mean adjusted effect size, controlling for the number of participants in each study. D_1 is used on the left-hand side of Table 2.

$$D_1 = \sum (N_i/N_T)^* d_i.$$

The variance of d_1 is a weighted average squared error.

$$\mathrm{Var}(d_1) = \sum (N_i/N_T)^* (d_1 - D_1)^2.$$

This sample variance is a biased estimate of the population variance, however. To correct for this bias the variance of the sampling error is estimated and subtracted from the sample variance. If e is the sampling error, and N_a is the average sample size across studies, then

$$\mathrm{Var}(e) = (4/N_a)^* (1 + D_1^2/8),$$

and the standard error is

$$SD(\delta_1) = \sqrt{\mathrm{Var}(d_1) - \mathrm{Var}(e)}.$$

To correct for measurement unreliability, an adjusted effect size, d_1, is divided by an estimate of the square root of the measurement reliability of the outcome measure, α_i. If we call this doubly adjusted effect size d_2, then

$$d_2/d_1/\sqrt{\alpha_i}.$$

Because in practice α_i is always less than 1, $d_2 > d_1$. Yet, the standard error of d_2 also increases.

$$SD(\delta_2) = SD(\delta_1)/\sqrt{\alpha_i}.$$

D_2, the mean of the d_2 estimates, is used in the middle panel of Table 2.

Although we did not have information on the reliability of every outcome measure in the meta-analysis, the information we did have enhanced our confidence that missing reliabilities could safely be estimated as the average reliability within a category of outcome measures. We adopted the following convention. If reliability was not reported for a particular finding based on a multiitem outcome measure, then we assigned the mean reliability of other findings within the same dependent variable category. For example, for the eight findings for which measurement reliability was reported concerning affect toward the sponsor of the ad, the mean reliability was .89. If another finding concerning affect toward the sponsor was based on a multiitem outcome measure but no reliability coefficient was reported, then we assigned it a reliability of .89. If a finding concerning affect toward the sponsor was based on a single-item outcome measure, however, we set its reliability at .60, two-thirds of the mean reliability for measures in the same category, the premise being that single-item measures are generally less reliable than multiitem measures.

We attempted one additional adjustment for the data, controlling for variation in the strength of the negative advertisement "treatment" to which experimental subjects or survey respondents were exposed. Our estimate of strength of treatment involved two factors, the definition and operationalization of negative advertisements by each research team, and the number of ads to which subjects were exposed. Although we have focused primarily on the differences between "positive" ads (those that focus on the sponsor) and "negative" ads (those that attack the opponent), a third category is sometimes distinguished, that of "contrast" or "comparative" ads (those that explicitly mention information about at least two candidates). Some researchers devise separate experimental conditions for all three types of ads,[18] while other researchers combine contrast ads and negative ads into a single category.

These three types of political advertisements may be qualitatively different (and hypothesis 1 in the text treats them as such), but they certainly are *quantitatively* different in the amount of negative information they convey about an opponent. That is, a positive ad presents no negative information about an opponent, a negative ad is composed largely of negative information about the opponent, and a contrast ad falls roughly halfway between these two extremes. Embracing this typology, we defined a "full-strength" manipulation as one that contrasts negative ads to positive ads, and we treated studies that contrast positive ads to contrast ads (or to some unknown mixture of negative and contrast ads[19]) as if they were half that strength.

We further refined this strength of treatment effect measure by multiplying it by the number of ads to which subjects were exposed—presuming that exposure to two positive or negative ads sponsored by a candidate has twice the effect of exposure to a single ad, and so on.[20] These adjustments are analogous to those that must be made in interpreting unstandardized regression coefficients when predictors have different ranges. The adjustment itself is simple: Each estimated effect size is divided by our estimate of treatment strength before means and standard errors are computed. These final adjusted effect sizes are reported in the last two columns of Table 2.

---

[17] This discussion is based on the presentation by Hunter and Schmidt (1990), who provide numerous examples and describe a wide range of possible adjustments to individual effect sizes.

[18] In such cases, the effect sizes we calculated compare the positive ad condition to the negative ad condition.

[19] A few studies contrast negative ads to a "no advertisement" control condition. These studies were also assumed to involve a "half-strength" treatment.

[20] Very few of the nonexperimental studies present any estimates of the number of positive or negative ads to which a typical survey respondent was exposed. For these studies we somewhat arbitrarily set the level at two. Our conclusions do not change if this level is set at either one or three.

TABLE A-1. Description of Studies Included in the Meta-Analysis

Study	Independent Variable	Subjects and Design	Dependent Variables	Results
1. Ansolabehere and Iyengar 1995	Positive or negative ad for actual candidates inserted into regular commercial break of local news program	Experiment with a convenience sample of 2,216 residents	Intended turnout	Negative ads depressed intended turnout, $d = -.10$.
			Vote intention in primary election	Negative ads decreased intended vote for their sponsor during primary elections, $d = -.14$.
			Vote intention in general election	Negative ads increased vote intention for their sponsor during general elections, $d = .10$.
			Political efficacy	Negative ads decreased efficacy, $d = -.12$.
2. Ansolabehere, Iyengar, Simon, and Valentino 1994	Positive/negative "tone" of 1992 Senate campaigns, coded from newspaper accounts	Aggregate analysis of turnout in 34 Senate elections	Actual turnout	States with more negative Senate election campaigns had lower turnout, $d = -1.39$.
3. Babbitt and Lau 1994	Positive/negative "tone" of 1988 and 1990 Senate campaigns, coded from newspaper accounts	Information about candidates from 1988 and 1990 U.S. Senate elections	Knowledge of major party candidates running in election	Negative issue-based campaigning associated with more issue-based knowledge of candidates, but not more general knowledge about incumbents, average $d = .04$.
4. Basil, Schooler, and Reeves 1991	Positive and negative ads from two senatorial campaigns in another state	Repeated measures design; convenience sample of 24 local community residents who saw two "campaigns" consisting of three positive or three negative ads from each candidate in the Senate election	Affect for sponsor of ad	Candidate was liked better and perceived as stronger when he presented positive ads, $d = -.30$.
			Affect for target of ad	The target was liked better when the opponent used positive rather than negative ads, $d = .46$.
			Memory of ad	Positive ads were recalled better than negative ads, $d = -.30$.
5. Brians and Wattenberg 1996	Exposure to television news, positive ads, and negative ads during the 1992 presidential election	51% of nationally representative survey (ANES data, $N = 1{,}263$) who could recall some political ad	Memory of ad	Negative political ads more likely to be recalled relative to an estimate of their prevalence during campaign, $d = .51$.

TABLE A-1. (Continued)

Study	Independent Variable	Subjects and Design	Dependent Variables	Results
6. Bullock 1994	Exposure to ads for two hypothetical state senate candidates, varying by the type of attack ad (image or issue) and the ambiguity of the ad	Experiment with 451 randomly selected prospective jurors awaiting assignment	Affect for target	Targeted candidates were rated less favorably after exposure to negative ads compared to positive ads, $d = 1.40$.
			Affect for sponsor	Candidates were rated more positively when sponsoring positive ads than attack ads, $d = -1.52$.
			Vote intention	Negative ads caused the likelihood of voting for the targeted candidate to drop significantly, $d = 1.09$.
7. Capella and Taylor 1992	Which candidate initiated negative ads in 25 Senate campaigns with "substantial amounts of negative advertising"	Vote totals in 25 Senate elections from 1986 through 1990	Authors' judgment of whether negative ad campaign "worked" or "failed" (i.e., how final results differed from projected results before ad campaign began)	Negative ad campaign decreased proportion of vote obtained by initiator of negative ads in 18 of 25 elections, $d = -.58$.
8. Finkel and Geer 1998	Proportion of negative ads used by two major-party presidential candidates, 1960–92	Aggregate turnout levels in nine presidential elections	Actual turnout	Turnout decreased with higher proportion of negative ads, $d = -.51$ (although due entirely to one outlier).
		Reported turnout by 12,252 ANES respondents, 1960–92	Actual turnout	Reported turnout was higher for survey respondents in election years with higher proportions of negative ads, $d = .01$.
9. Freedman and Goldstein 1999	Very sophisticated estimate of number of negative ads seen by survey respondents	Second wave ($N = 290$) of representative panel study of 1997 Virginia gubernatorial campaign	Reported turnout	Viewing more negative ads associated with higher turnout, $d = .25$.
			Political efficacy	Exposure to negative ads slightly lowers political efficacy, $d = -.05$.
10. Garramone, Atkin, Pinkleton, and Cole 1990	Exposure to various combinations and numbers of positive and negative biographical profiles and political commercials for two fictional U.S. Senate candidates	Experiment with 372 students assigned to one of six conditions: control, double-positive, single-positive, negative-positive, single-negative, or double-negative	Differential affect	Exposure to negative ads caused greater image discrimination (the difference between candidate image evaluations) than exposure to positive ads, $d = .38$.
			Voter turnout	Negative ads did not significantly affect voter turnout, $d = -.18$.

TABLE A-1. (Continued)

Study	Independent Variable	Subjects and Design	Dependent Variables	Results
11. Geer and Lau 1998	State-based estimates of proportion of negative ads used by two major-party presidential candidates, 1960–96	State-level turnout in ten presidential elections, 1960–96	State-level turnout	The higher the proportion of negative advertising used by both candidates in the state, the higher the turnout, *d* = .48.
		Reported turnout by 15,632 ANES survey respondents, 1960–96	Reported turnout	The higher the proportion of negative advertising used by both candidates in the state, the higher the turnout, *d* = .08.
12. Goldstein 1997	Number of negative ads shown in 75 largest media markets during 1996 presidential campaign	Aggregate analysis of 1,588 counties, followed by individual-level analysis of 879 ANES respondents living in 75 largest media markets	County-level turnout	The more negative ads run in county, the lower the turnout, *d* = −.27.
			Reported turnout	More negative ads to which a respondent exposed, the greater probability of voting, *d* = .09.
			Political efficacy	More negative ads to which a respondent exposed, the lower political efficacy, *d* = −.05.
13. Haddock and Zanna 1997	Impressions of actual candidates before and after controversial attack ads aired during 1993 Canadian national election	"Natural" experiment with 110 college students	Affect for sponsor of attack ads	Affect toward sponsor of ads decreased after airing of ads, *d* = −.32.
			Affect for target of attack ads	Affect toward target of ads increased after airing of ads, *d* = −.35.
14. Hill 1989	Positive or negative ad from Bush or Dukakis campaign	Experiment with 120 college students	Affect for sponsor of attack ads	Sponsor of ad was liked less if ad was negative rather than positive, *d* = −.65.
			Affect for target of attack ads	Target of ad was liked more if ad was negative rather than positive, *d* = −.13.
			Affect for the ad itself	Negative ads were liked more than positive ads, *d* = .73.
15. Hitchon and Chang 1995	Exposure to positive, neutral, and negative ads from female and male gubernatorial candidates	Experiment using a 3 (positive, negative, neutral) × 2 (female, male) within-subject factorial design with 75 undergraduate subjects	Affect for ads	Negative ads received more negative evaluations than positive ads, *d* = −.50.
			Affect for sponsor of attack ads	More negative affect for candidates who attacked their opponents, *d* = −.81.
			Memory for ads	Positive ads produced highest candidate recall, while negative ads produced lowest candidate recall, *d* = −.58.

TABLE A-1. (Continued)

Study	Independent Variable	Subjects and Design	Dependent Variables	Results
16. Hitchon, Chang, and Harris 1997	Exposure to positive, neutral, and negative ads in a gubernatorial race	Experiment using a 3 (positive, neutral, negative) × 2 (male, female) within-subject factorial design with 72 undergraduate subjects	Affect for sponsor of ads	Negative ads produced less favorable responses than positive or neutral ads, $d = -.80$.
17. Kahn and Geer 1994	Actual positive or negative ads from out-of-state gubernatorial candidates inserted in regular ad breaks during a TV sitcom	Experiment with 209 college students; subjects saw one or two positive or negative ads	Affect for sponsor of ads	Sponsor of ad was liked less after one negative compared to one positive ad, $d = -.28$. Sponsor of ads was liked less after two negative compared to two positive ads, $d = -.74$.
18. Kahn and Kenney 1999	Coding of sample of campaign ads from 1988–92 U.S. Senate elections	ANES Senate election study, $N = 2{,}256$	Reported turnout	Relatively greater use of negative ads by both candidates associated with higher turnout, $d = .02$.
19. Kahn and Kenney 1998b	Coding of sample of campaign ads from 1988–92 U.S. Senate elections	ANES Senate election study, $N = 6{,}110$	Knowledge of Senate candidates	Respondents had more awareness of major party candidates when relatively more negative ads used, $d = .07$.
20. Kaid 1997	Exposure to actual ads from 1996 Clinton or Dole campaigns	Experiment with 116 undergraduates as subjects	Vote intention	Subjects were much more likely to say they intended to vote for a candidate after viewing one of his negative ads compared to one positive ad, $d = 1.77$.
			Affect for target of ads	Target of ads liked less after negative ad compared to positive ad, $d = .68$.
			Affect for sponsor of ads	Sponsor of negative ad liked slightly more than sponsor of positive ad, $d = .28$.
21. Kaid and Boydston 1987	One of five actual ads used by congressional candidate from another district	Convenience sample of 428 residents rated candidate before and after seeing one of his ads	Affect for target of ads	Affect for target of ads dropped after viewing negative ad, $d = .36$.