A meta-analysis was performed of studies of mediated health campaigns in the United States in order to examine the effects of the campaigns on behavior change. Mediated health campaigns have small measurable effects in the short-term. Campaign effect sizes varied by the type of behavior: $r = .15$ for seat belt use, $r = .13$ for oral health, $r = .09$ for alcohol use reduction, $r = .05$ for heart disease prevention, $r = .05$ for smoking, $r = .04$ for mammography and cervical cancer screening, and $r = .04$ for sexual behaviors. Campaigns with an enforcement component were more effective than those without. To predict campaign effect sizes for topics other than those listed above, researchers can take into account whether the behavior in a cessation campaign was addictive, and whether the campaign promoted the commencement of a new behavior, versus cessation of an old behavior, or prevention of a new undesirable behavior. Given the small campaign effect sizes, campaign planners should set modest
goals for future campaigns. The results can also be useful to evaluators as a benchmark for campaign effects and to help estimate necessary sample size.

Introduction

Researchers have long debated the efficacy of mediated health campaigns (Douglas, Westley, & Chaffee, 1970; Hornik, 1988; Hyman & Shetsley, 1947; Mendelsohn, 1973; Wallack, 1981). Only recently have they begun to use meta-analysis to determine how effectively media campaigns bring about health behavior change (Freimuth & Taylor, 1994; Redman, Spencer, & Sanson-Fisher, 1990; Snyder & Hamilton, 2002). On average, campaigns have a small and quantifiable effect on behavior change (Snyder & Hamilton, 2002).

By computing the average campaign effect size under different circumstances, meta-analyses of campaign effectiveness can provide valuable information for funders, campaign planners, researchers, and evaluators. Knowing the average effect size allows campaign planners to set their goals more realistically. Researchers and evaluators can use information about average effect sizes to calculate power, plan for an appropriate sized sample for evaluation purposes, and act as a benchmark when evaluating particular campaigns. Finally, funders can compare the effectiveness of media campaigns with other types of interventions, such as school-based programs or clinical health education.

The purpose of the present article is to estimate the average effects of mediated health campaigns for different health topics and types of behaviors. The focus was on behavioral outcomes because behavior change is often the “bottom line” goal of a campaign, even though is it more difficult to achieve than awareness of a problem, knowledge of a solution, or motivation for change (McGuire, 1981; Prochaska & DiClemente, 1983; Rogers, 1983). Drawing on the diffusion of innovations (Rogers, 1983) and other approaches, we classified behaviors by the adoption goals of commencement, prevention, or cessation, whether the behaviors were addictive or not, and the rate of behavioral compliance before the campaign.

Campaign Topic

Intervention work in health is often organized around medical problems, such that researchers often stay in one health domain (e.g., AIDS) for many years and perhaps their whole career. Within different health subfields, separate journals often exist, different theories may predominate, and interventions may be conducted differently. To the extent that there are differences in campaign effects by topic, researchers in particular subfields would be interested in learning about the average effect size for their own area.

RQ1: What is the average effect size for different campaign topics?

Type of Behavior

Meta-analyses have shown that the effect of clinic-based interventions varies by the type of behavior targeted by the intervention (Mullen, Mains, & Velez, 1992). For example, the average correlation between exposure to a clinic-based intervention and behavior change for cardiac patients was $r = .09$ for diet and exercise behaviors, $r = .03$ for smoking, $r = -.04$ for drug adherence, and $r = .25$ for blood pressure (converted from the g-statistics in Mullen et. al, 1992, Table 7). Unfortunately, it is not feasible to meta-analyze the effects of mediated campaigns on most specific behaviors because there have been too few studies conducted on those specific behaviors.

To deal with this limitation, researchers can classify behaviors based on their characteristics and test the impact of the theoretically-derived constructs on campaign
effect size. The classification approach is practical for topics for which there are not enough campaigns of a particular type to provide reliable estimates. Classifying behaviors according to theoretical constructs may also enable researchers to predict the impact of a new behavior based on the theory. The next three sections present a different classification scheme based on theoretical constructs.

**Adoption Type**

The first classification differentiates behaviors based on the goal of the campaign—whether the campaign goal is to increase the target behavior or not (Mullen, Simons-Morton, Ramirez, Frankowski, Green, & Mains, 1997). However, a meta-analysis found that clinic-based education had the same effect on a wide range of behaviors. There was no appreciable difference in effect sizes between interventions that tried to increase behavior ($r = .29$, seat belt use, exercise, breast self-examination) and interventions that tried to reduce or substitute behavior ($r = .29$ smoking and alcohol reduction; $r = .24$ weight reduction and nutrition; Mullen et al., 1997). Part of the problem may be that it is unclear how to categorize campaigns that aim to prevent a behavior without necessarily promoting a new behavior (e.g., illicit drug use prevention).

A more promising distinction is to differentiate between

(a) the commencement of a new behavior,
(b) the prevention an undesirable new behavior, and
(c) the cessation/reduction of an old behavior.

We call this three-category scheme type of adoption, following Rogers (1983), who uses the term adoption of innovations as a generic term that encompasses all types of behavior changes. The type of adoption includes the direction of the effect (increasing or decreasing) and whether the behavior is desirable or undesirable to the campaign planners. Unlike the Mullen et al. (1997) approach, our proposed scheme treats the substitution of a new behavior for an old behavior as a commencement behavior rather than grouping it with reduction of an old behavior. Once enough substitution campaigns have been conducted, they could be analyzed as a separate category.

More than one adoption type may occur for a given health topic. A nutrition campaign may emphasize prevention (e.g., not giving sugar cereals to children), commencement (e.g., increasing fiber), or cessation (e.g., reducing fat consumption). Smoking campaigns may emphasize prevention (particularly for youth) or cessation.

Using the three categories of adoption, we anticipate that it is easier to convince people to commence a new behavior than to extinguish an old behavior. Prevention behaviors may fall in between, since the undesirable behavior is not yet habitual and might therefore be more susceptible to influence.

H1: Campaigns promoting the commencement of a new behavior will have a greater average effect size than campaigns promoting prevention of an undesirable behavior, which, in turn, will have a greater average effect size than campaigns promoting the cessation of an existing undesirable behavior.

**Addictiveness of Behavior**

Another way to classify behaviors is whether the target behavior is addictive or not. To extinguish non-addictive behavior, campaigns seek to alter an individual’s intentions and
attitudes towards the behavior. We expect that the physiology of addiction would increase a person’s resistance to campaign messages. For example, people who want to quit smoking must cope with the physical discomforts of withdrawal. We predict that addictive behaviors would be more difficult to change in a campaign, given the physiological barriers to changing an addictive behavior.

H2: Campaigns that promote the cessation of addictive behaviors will have a smaller effect size than those promoting the cessation of non-addictive behaviors.

Baseline Behavior Rate

Diffusion theory (Rogers, 1983) states that the percentage of the population that has adopted a new behavior follows a predictable pattern over time. Across disciplines, researchers have found that the adoption process is most often described by curves (or functions) that resemble the normal curve, relating the percentage of new adopters over time. There are very few new adopters at the beginning and end of the diffusion time frame, and many more in the middle (Mahajan & Peterson, 1985; Rogers, 1983). If very few people have changed their behavior, then getting more people to innovate can be a difficult process. Similarly it can be difficult to get the last 10 percent of the population to change their behavior. Rogers calls them “laggards,” and they are not a promising target group for campaign designers (1983). The exact shape of the diffusion curve—where the peak occurs, whether it is symmetrical or not, how steep the increase and decrease—varies from innovation to innovation. Often researchers will examine the cumulative percentage of adopters over time, which follows an S-curve (Rogers, 1983).

We can apply the diffusion curve to make predictions about the impact of initial level of behavior in a population prior to a campaign. Typically, the percentage of the population already compliant with a campaign is measured in a baseline wave of data collection that occurs before the campaign begins. Since adoption rates are low at the beginning of the diffusion period, we can predict that campaigns with a low baseline rate of behavior should be more difficult and have a lower campaign effect size. Since adoption rates are high in the middle of the diffusion period, we can predict that campaigns with levels of baseline behavior levels at about 50 percent of the population should have greater success rates, as evidenced by a greater campaign effect size.

At the end of the diffusion period, when baseline behavior rates are high (approaching 100%), then the shape of the diffusion curve suggests that campaigns will have a more difficult time changing behavior, and so should have a lower effect size. Thus, very low and very high baseline behavior rates should have a lower effect size than baseline behavior rates that are moderate. The curve depicting the relationship between baseline behavior rate and campaign effect size should consist of an upward trend (monotonically increasing function), followed by a downward trend (monotonically decreasing function).

In addition to deriving the shape of the baseline behavior rate and campaign effect size from diffusion theory, it is valuable to consider the underlying processes that may explain the relationship between these two variables. The upward trend, which we call the “bandwagon effect,” represents the positive impact of a baseline behavior rate on campaign effect size. The more people who engage in the new behavior, the more correct the behavior appears to be. In addition, greater baseline rates of behavior may result in increased interpersonal communication about the campaign topic, which should result in higher rates of behavior change. Finally, the more the population is already doing the correct behavior, the greater availability of positive models for the correct behavior, which may increase campaign success rates.
The downward trend, which we call the “resistance effect,” is the negative impact of baseline behavior rate on campaign effect size. Some researchers have found decreasing rates of diffusion over time (Coleman, Katz, & Mendel, 1966; Hamblin, 1973), such that the more people who have adopted an innovation, the lower the effect of the campaign. As more and more people adopt the target behavior, the remaining population becomes more resistant. Resistance may be due to receivers’ involvement with the campaign topic, their accumulated information on the topic, or their reactance against pressure exerted by the campaign (Hamilton & Stewart, 1993). If people respond with reactance to increased pressure to change, then the more people who engage in the behavior, the greater the reactance.

A negative relationship between baseline behavior rates and campaign effects may also be due to decreasing percentages of people who are ready to change. When modest numbers of people have already adopted the behavior, it may be possible to find a larger percentage of people ready to change their behavior, perhaps because they are already further along in the stage of change (Prochaska & DiClemente, 1983), or more susceptible to persuasion (Hovland & Janis, 1953) because of exposure to prior campaigns, interpersonal influences, and life circumstances. For those campaigns that begin with a higher level of baseline behavior rates, people with a propensity to change may have already changed, so a campaign would have a more difficult time persuading the remainder of the population, who are more resistant.

Together, the bandwagon effect and resistance effect help explain the expected relationship between baseline behavior rate and campaign effect size. As baseline rates increase from zero to moderate levels, the bandwagon effect will predominate because resistance is low in the population. As more people adopt the behavior, the percentage of people who are resistant to behavior change increases. As baseline rates increase from moderate levels to high levels, the resistance effect will dominate the bandwagon effect.

A mathematical model can be constructed for the relationship between baseline behavior rate and campaign effect size. A linear increasing trend would be evidence of a bandwagon effect with no resistance. A linear decreasing trend would be evidence of a resistance effect with no bandwagon effect. A mixed model of increasing and decreasing trend (Mahajan & Peterson, 1985) would be evidence of a bandwagon effect predominating at lower levels of baseline behavior, and a resistance effect predominating at higher levels of baseline behavior. The mixed-model curve would identify the point of maximal impact of baseline behavior rate on campaign effect size. The same inflection point represents the point at which resistance begins to have a stronger effect than the social influences of the bandwagon effect. We can test which function best describes the relationship between baseline behavior rate and campaign effect size: the increasing line of the bandwagon effect, the decreasing line of a resistance effect, or a mixed-model inverted-U function.

RQ2: Which function best describes the impact of baseline behavior rates on campaign effect size?

Methods

Selection Criteria for Inclusion of Studies

Studies were included in the meta-analysis if they met our criteria regarding publication, media, variables, design, and measurement. First, all campaign evaluations must have been published in English in refereed journals or in edited scholarly or professional books. Second, the studies must have reported on health campaigns using at least one
form of community-wide mass media. We excluded campaigns that only use interpersonal channels, campaigns that were waged within schools or workplaces, and experiments that exposed people to media in small group settings. Third, the campaigns needed to have taken place in the United States, to control for cultural and media system differences. Fourth, the studies must have included fully specified measures of campaign effects on at least one type of behavior advocated by the campaign. In a few cases physiological measures were included in the meta-analysis because prior research showed that there were no differences in effect size between behavioral measures and physiological measures (Snyder & Hamilton, 1999).

The unit of analysis was the campaign. We located appropriate evaluation information on 48 campaigns, with a total N of respondents across all campaigns of 168,362.

Search Procedure

Four databases—Psychlit, Soclit, Medline, and Eric—were searched in summer, 1998. The keywords “health,” “campaign,” “communication,” were supplemented with “education,” “media,” “mass media,” “television,” “posters,” “billboards,” “newspapers,” “radio,” “intervention,” “anti-drug,” “smoking,” “risk,” “cancer,” “AIDS,” “seat belts,” and “community.” In addition, we used books about campaigns and literature reviews for leads on additional campaigns.

We examined several thousand abstracts and over 300 publications. The majority of the rejected publications were reviews, did not test media effects, or only used media in schools or in the workplace (k = 131). Another set of studies did not report behavioral effects of the campaign, and focused instead on other outcomes (k = 38). When articles reported on the design of a campaign and did not contain evaluation results, we conducted a search for a published evaluation; no evaluation data had been published for 11 of those campaigns. Despite our attempts to contact the authors for additional information, ten campaigns were rejected because the statistics they reported were incomplete.

Measures

Effect Size

The main dependent variable was the effect of exposure to the media campaign on behavior change. When evaluations reported multiple behavioral measures, effect size was averaged across the measures. When multiple journal articles about the same campaign reported on the same behavior, we chose the effect estimate that

1. was the most fully specified statistically;
2. represented the entire target population, rather than a subset; and
3. coincided with the end of media activities (we will deal with the issue of long-term effects in another publication).

To compute the effect size for each campaign, we converted the published statistic into a correlation using standard formulas (Rosenthal, 1994). Software that calculated the conversions included DSTAT 1.11 (Johnson, 1995), VGBARE (Hunter, 1993), and MetaCor (Hamilton, 1991). When studies did not report a statistic but did provide pretest and posttest percentages (k = 24), we used the following formula to compute d:

\[
d = \frac{(i_2 - i_1) - (c_2 - c_1)}{\sqrt{\frac{(i_1 + c_1)}{2} \cdot \left(1 - \frac{(i_1 + c_1)}{2}\right)}}
\]

\[d = \frac{(i_2 - i_1) - (c_2 - c_1)}{\sqrt{\frac{(i_1 + c_1)}{2} \cdot \left(1 - \frac{(i_1 + c_1)}{2}\right)}}
\]
where $i$ is the intervention (or campaign) community, $c$ is the control community, and the subscript specifies the pretest (1) or posttest (2). The formula equates the effect size $d$ to the change over time in the intervention community, minus the change over time in the control community, divided by the standard deviation of the average pretest score (Hunter, personal communication). The statistic $d$ was then converted to $r$.

**Campaign Topic**
For each campaign, we coded the topic of the campaign. The broad categories of topics included smoking, drinking, seat belt use, cardiovascular (diet and exercise), mammography, dental care, and sexual behavior campaigns.

**Addiction**
Smoking turned out to be the only addictive behavior found in the sample of campaigns.

**Adoption Type**
We coded whether the object of the campaign was for people to commence a new behavior (seat belt use, exercise, mammography, dental care, condom use, health status screenings, hypertension control, supportive interpersonal behaviors, fruit and vegetable consumption, and crime prevention behaviors), cease a current behavior (smoking, drinking, infants sleeping with milk bottle, sex with risky partner), or prevent a future behavior (smoking).

**Baseline Behavior Rate**
Baseline behavior rate is the pretest percentage of the desired behavior in the intervention community/communities. Only 28 of the 48 campaigns reported a pretest measure in the intervention site. For campaigns that did not have a baseline measure in the intervention community but did measure comparison community percentage of behavior at post-test, that figure was used, bringing the total number of campaigns with a baseline behavior rate measure to 36.

**Analysis**
The analysis employed the Hunter and Schmidt (1990) approach to meta-analysis, which is slightly more conservative than other meta-analytic techniques (Johnson, Mullen, & Salas, 1995). Campaign effect size estimates were not corrected for attenuation because publications did not provide reliability scores for their measures. In nearly all cases, the behavior measure was a single item. In contrast to other meta-analyses in the field of communication, the present sample showed extreme variation in sample size, ranging from $N = 121$ to $N = 40,493$. We compared the effects with and without the largest studies, to determine whether including the largest studies distorted our results. There was little relationship between sample size and campaign effect size: $r(48) = -.07$. In subsequent analyses, study effects were weighted based on the sample size.

We did not correct for the intra-cluster correlation within studies that sampled individuals within clusters, unless the data were reported that way. Since only a few studies (e.g., Grube, 1997; Hannan, Murray, Jacobs, & McGovern, 1994) reported the ICC, it was not feasible to correct the studies using the ICC (Simpson, Klar, & Donner, 1995). Failing to account for ICC should not affect effect size (Zucker, 1990), the main concern of the present study. A meta-analysis of school-based smoking interventions found correcting for the ICC was a non-significant adjustment (Rooney & Murray, 1996).
The average media campaign effect given as a correlation ($r$) was converted into the average percentage campaign behavior change by first converting $r$ to $d$ with DSTAT (Johnson, 1995). Note that $2 r = d$ when is less than .24 (Hunter & Schmidt, 1990). The $d$-statistic is the average change in behavior divided by the standard deviation ($SD$) of the average behavior rate, and the $SD$ is the average behavior rate * (100 – the average behavior rate).

**Results**

The average media campaign effect on behavior was $\bar{r} = .09$, with a 95% confidence interval of .07 to .10. The total number of participants ($Tn$) was 168,362; the average $n$ per study was 3508, and the number of studies ($k$) was 48. Table 1 presents the list of studies and their attributes. Among those studies measuring change population percentage engaged in the behavior, the average change in percentage was 8% ($SD = .19$, range 5 to 92, $N = 156,654$, $k = 40$). The average rate of desired behavior per campaign (combining pretest and posttest, campaign and control communities as available) was $M = .67$ ($SD = .19$, range 8 to 92, $N = 156,779$, $k = 40$). The test of homogeneity indicated that the set of campaigns was heterogeneous ($SD_r = .06$), with sampling error explaining 7% of the variance across campaigns ($\chi^2(47) = 658.03$, $p < .001$). That is, 93% of the variance across campaigns could be attributable to moderator variables. Next, we examined the extent to which this variance was due to the four moderator variables related to the hypotheses and research questions: campaign topic, adoption type, addiction, and baseline behavior rate.

**Campaign Topic and Addiction**

There was a wide variety of health campaign topics, and topic had a very large impact on effect size ($\eta = .78$; see Table 2). The greatest campaign effect size occurred with seat belt campaigns ($\bar{r} = .15$). Four of the eight campaign topics were homogeneous: sampling error explained 100% of the variance within the drinking, oral health, mammography, and sexual campaigns. The other three campaign topics (seat belts, heart campaigns, smoking) were relatively homogeneous, with $SD_r$ ranging from .03 to .05. The effectiveness of campaigns on different health topics appears in Figure 1.

**Adoption Type**

Half of the 48 campaigns promoted the adoption of a new behavior—seat belts, mammogram or PAP screenings, condom use, hypertension medicine, dental checkups, supportive behaviors in friendships, fruit and vegetable consumption, and crime prevention behaviors. The range of cessation behaviors (40% of the campaigns) was much more limited (smoking, alcohol, unprotected sex, and baby bottle tooth decay) and dominated by anti-smoking campaigns. Only five campaigns (10%) were aimed at prevention of bad health behavior without also promoting the adoption of a new behavior, and their objective was youth smoking prevention.

Within the entire sample of campaigns, adoption type had a substantial impact on the size of the campaign effect: $\eta = .53$. Commencement campaigns were more successful than prevention or cessation campaigns. Commencement campaigns had a relatively larger average effect size ($\bar{r} = .12$, $Tn = 78,351$, $k = 24$), and modest amount of variance across studies ($SD_r = .06$), with sampling error accounting for only 7% of the variance. Prevention campaigns had a small average effect size ($\bar{r} = .06$, $Tn = 33,316$, $k = 5$) with
<table>
<thead>
<tr>
<th>Campaign</th>
<th>Cite</th>
<th>Behavior</th>
<th>$\bar{r}$</th>
<th>Sample Size</th>
<th>Adopt Cease Prev.</th>
<th>Diffuse Curve Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Su Salud</td>
<td>McAlister et al., 1992; Ramirez &amp; McAlister, 1988</td>
<td>Smoking cessation</td>
<td>.20</td>
<td>175</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>AIDS Community Demonstration Project</td>
<td>Fishbein, Guenther-Grey, Johnson et al., 1996</td>
<td>Condom use, vaginal sex, bleach use</td>
<td>.03</td>
<td>6,184</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>AIDS Prevention for Pediatric Life</td>
<td>Santelli, Celantro, Rozsenich et al., 1995</td>
<td>Condom use</td>
<td>.05</td>
<td>1,509</td>
<td>A</td>
<td>30</td>
</tr>
<tr>
<td>America Responds to AIDS</td>
<td>Snyder, 1991</td>
<td>Risky sex</td>
<td>.01</td>
<td>163</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>CA Tobacco Education Media Camp.</td>
<td>Popham, Potter, Hetrick, Muthen, Duerr, &amp; Johnson, 1994</td>
<td>Smoking</td>
<td>.03</td>
<td>10,339</td>
<td>P</td>
<td>13</td>
</tr>
<tr>
<td>Cancer Control in a TX Barrio</td>
<td>McAlister et al., 1995</td>
<td>Mammography screening, pap smear</td>
<td>.05</td>
<td>309</td>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Campaign</th>
<th>Citea</th>
<th>Behaviorb</th>
<th>ř</th>
<th>Sample Size</th>
<th>Cease Prev. c</th>
<th>Diffuse Curve Leveld</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Trials Project: Underage Access Component</td>
<td>Grube, 1997</td>
<td>Alcohol sales to minors</td>
<td>.17</td>
<td>949</td>
<td>C</td>
<td>50</td>
</tr>
<tr>
<td>Decreasing Binge Drinking at College</td>
<td>Haines and Spear, 1996</td>
<td>Binge drinking</td>
<td>.07</td>
<td>4,258</td>
<td>C</td>
<td>43</td>
</tr>
<tr>
<td>Drinking During Pregnancy</td>
<td>Kaskutas and Graves, 1994</td>
<td>Limiting drinking</td>
<td>.11</td>
<td>2,746</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Farm Cancer Control Project</td>
<td>Gardiner, Mullan, Rosenman, Zhu, &amp; Swanson, 1995</td>
<td>Mammography screening</td>
<td>.01</td>
<td>1,545</td>
<td>A</td>
<td>49</td>
</tr>
<tr>
<td>Five A-Day for Better Health, CA</td>
<td>Foerster, Kizer, DiSogra, Bal, Krieg, &amp; Bunch, 1995</td>
<td>Fruit and vegetable consumption</td>
<td>.01</td>
<td>2,002</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Forsyth County Cervical Cancer Prevention</td>
<td>Dignan, Michielutte, Wells, &amp; Bahnson, 1994</td>
<td>Pap smear</td>
<td>.04</td>
<td>1,830</td>
<td>A</td>
<td>45</td>
</tr>
<tr>
<td>Freedom from Smoking in St. Louis</td>
<td>Wheeler, 1988</td>
<td>Quit smoking</td>
<td>.18</td>
<td>429</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Friends Can Be Good Medicine</td>
<td>Hersey, Klibanoff, Lam, &amp; Taylor, 1984</td>
<td>Supportive behavior</td>
<td>.09</td>
<td>340</td>
<td>A</td>
<td>36</td>
</tr>
<tr>
<td>Grand Junction Bike Helmet</td>
<td>Rouzier &amp; Alto, 1995</td>
<td>Wear bike helmet</td>
<td>.41</td>
<td>121</td>
<td>A</td>
<td>6</td>
</tr>
<tr>
<td>Study/Program</td>
<td>Reference</td>
<td>Measure(s)</td>
<td>Effect Size</td>
<td>df</td>
<td>p Value</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------</td>
<td>-------------</td>
<td>----</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Heart to Heart</td>
<td>Goodman, Wheeler, &amp; Lee, 1995</td>
<td>Smoking, inactivity, cholesterol (P), blood pressure (P), weight (P)</td>
<td>0.01</td>
<td>2700</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Know When to Say No</td>
<td>Werch &amp; Kersten, 1989; Werch, Kersten, &amp; Young, 1992</td>
<td>Drinking</td>
<td>0.12</td>
<td>314</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>KY Rural High Blood Pressure Control</td>
<td>Kotchen, McKean, Jackson-Thayer, Moore, Straus, &amp; Kotchen, 1986</td>
<td>Hypertension (P)</td>
<td>0.10</td>
<td>1,044</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Media-Based Mammography in San Diego</td>
<td>Mayer, Krossman, Miller, Crooks, Slymen, &amp; Lee, 1993</td>
<td>Mammography screening</td>
<td>0.05</td>
<td>506</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Minority Smoking Cessation in Chicago</td>
<td>Jason, Tate, Goodman, Buckenberger, &amp; Gruder, 1988</td>
<td>Smoking cessation</td>
<td>0.16</td>
<td>137</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>MMHP, MN Adult Smoking Prevention</td>
<td>Jacobs, Leupker, Mittlemark et al., 1986; Lando, Pechacek, Pierie et al., 1995; Leupker, Murray, Jacobs et al., 1994</td>
<td>Smoking, physical activity</td>
<td>0.05</td>
<td>7,400</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>MMHP, MN Youth Smoking Prevention</td>
<td>Perry, Klepp, &amp; Sillers, 1989</td>
<td>Smokers</td>
<td>0.09</td>
<td>4,090</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>MN Periodontal Awareness TV Campaign</td>
<td>Bakdash, McMillan, &amp; Lange, 1984</td>
<td>Dental visits</td>
<td>0.13</td>
<td>2,000</td>
<td>A</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Campaign</th>
<th>Citea</th>
<th>Behaviorb</th>
<th>( \bar{r} )</th>
<th>Sample Size</th>
<th>Adopt Cease Prev.( c )</th>
<th>Diffuse Curve Leveld</th>
</tr>
</thead>
<tbody>
<tr>
<td>MN/ WI Adolescent Tobacco Use</td>
<td>Murray, Perry, Griffin, et al., 1992; Murray, Pirie, Leupker, &amp; Pallonen, 1989; Murray, Prokhorov, &amp; Harty, 1994</td>
<td>Smoking</td>
<td>.07</td>
<td>15,396</td>
<td>P</td>
<td>12</td>
</tr>
<tr>
<td>Mpowerment Project</td>
<td>Kegeles, Hays, &amp; Coates, 1996</td>
<td>Unprotected anal sex</td>
<td>.12</td>
<td>188</td>
<td>A</td>
<td>41</td>
</tr>
<tr>
<td>Parents Magazine Intervention</td>
<td>Kishchuck, Laurendeau, Desjardin, &amp; Perreault, 1995</td>
<td>Positive, negative interactions with kids</td>
<td>.02</td>
<td>307</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Preventing Baby Bottle Tooth Decay</td>
<td>Bruerd, Kinney, &amp; Bothwell, 1989</td>
<td>Tooth decay (P)</td>
<td>.14</td>
<td>1,465</td>
<td>C</td>
<td>43</td>
</tr>
<tr>
<td>Programma Latino para Dejar de Fumar</td>
<td>Marín, Pérez-Stable, Sabogal, &amp; Otero-Sabogal, 1990; Marín, Perez-Stable, Marín, &amp; Hauck, 1994</td>
<td>Smoking</td>
<td>.06</td>
<td>5,701</td>
<td>C</td>
<td>22</td>
</tr>
<tr>
<td>Rural CVD Program in WV</td>
<td>Farquhar, Behnke, Detels, &amp; Albright, 1997</td>
<td>Wellness score (P)</td>
<td>.09</td>
<td>425</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Seat Belt Contest</td>
<td>Foss, 1989</td>
<td>seat belt use</td>
<td>.09</td>
<td>6,072</td>
<td>A</td>
<td>26</td>
</tr>
<tr>
<td>Seat Belt Use</td>
<td>Robertson, Kelley, O’Neill, Wixom, Eisswirth, &amp; Haddon, 1974</td>
<td>seat belt use</td>
<td>.01</td>
<td>2,720</td>
<td>A</td>
<td>15</td>
</tr>
<tr>
<td>Seat Belt Use in Elmira, NY</td>
<td>Williams, Lund, Preussner, &amp; Blomberg, 1987</td>
<td>seat belt use</td>
<td>.24</td>
<td>3,358</td>
<td>A</td>
<td>49</td>
</tr>
<tr>
<td>Study Title</td>
<td>Authors</td>
<td>Outcome</td>
<td>Data</td>
<td>Setting</td>
<td>Year</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td>------</td>
<td>---------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Seat Belt Use in Modesto, CA</td>
<td>Lund, Stustser, &amp; Fleming, 1989</td>
<td>seat belt use</td>
<td>.22</td>
<td>A</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Seat Belts in VA</td>
<td>Roberts and Geller, 1994</td>
<td>seat belt use</td>
<td>.16</td>
<td>A</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>Smoking Prevention in CA, Prop 99</td>
<td>Jenkins, McPhee, Le, Pham, Ha, &amp; Stewart, 1997</td>
<td>Quit smoking</td>
<td>.04</td>
<td>C</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>Smoking Prevention in School</td>
<td>Flynn, Worden, Secker-Walker, Badger, Geller, &amp; Costanza, 1992; Flynn, Worden, Secker-Walker, Pirie, Badger, Carpenter, &amp; Geller, 1984</td>
<td>Smoking</td>
<td>.10</td>
<td>P</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Smoking: Community Control Center, L.A.</td>
<td>Danaher, Berkanovic, &amp; Gerber, 1984</td>
<td>Quit smoking attempts</td>
<td>.15</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking: VA Hospital Clinic</td>
<td>Mogielnicki, Neslin, Dulac, Balstra, Gillie, &amp; Corson, 1986</td>
<td>Smoking abstinence</td>
<td>.19</td>
<td>C</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Stanford 3 Community Study</td>
<td>Farquhar, Wood, Breitrose et al., 1977; Fortmann, Williams, Hulley, Haskell, &amp; Farquhar, 1981; Meyer, Nash, McAlister, Maccoby, &amp; Farquhar, 1980</td>
<td>Diet, weight, cholesterol, fat (P)</td>
<td>.06</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford 5 Community Study</td>
<td>Fortmann, Winkleby, Flora, Haskell, &amp; Taylor, 1990; Schooler, Chaffee, Flora, &amp; Roser, 1998</td>
<td>Smoking, exercise (P) and weight, cholesterol, blood pressure</td>
<td>.01</td>
<td>C</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Campaign</th>
<th>Cite</th>
<th>Behavior</th>
<th>( r )</th>
<th>Sample Size</th>
<th>Adopt Cease Prev.</th>
<th>Diffuse Curve Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop Smoking Clinic, NY Su Vida Su Salud</td>
<td>Dubren, 1977; Suarez, Nickols, &amp; Brady, 1993</td>
<td>Quit smoking, Screenings: pap, mammograms, and breast</td>
<td>.11</td>
<td>293</td>
<td>C</td>
<td>38</td>
</tr>
<tr>
<td>Take a Bite out of Crime</td>
<td>O’Keefe, 1985</td>
<td>Crime prevention</td>
<td>.10</td>
<td>1,049</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Time to Quit in Buffalo</td>
<td>Cummings, Sciandra, &amp; Markello, 1987</td>
<td>Smoking cessation</td>
<td>.16</td>
<td>321</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>VT Drink Calculator</td>
<td>Worden, Flynn, Merrill, Waller, &amp; Haugh, 1989</td>
<td>Blood alcohol (P)</td>
<td>.08</td>
<td>487</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>Community Education</td>
<td>Wing and Epstein, 1982</td>
<td>Weight loss (P)</td>
<td>.08</td>
<td>189</td>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Weight-a-Thon</td>
<td>Bauman, Brown, Bryan, Fisher, Padgett, &amp; Sweeney, 1988; Bauman, LaPrelle, Brown, Koch, &amp; Padgett, 1991; Bauman, Padgett, &amp; Koch, 1991; LaPrelle, Bauman, &amp; Koch, 1992</td>
<td>Smoking</td>
<td>.03</td>
<td>951</td>
<td>P</td>
<td>32</td>
</tr>
</tbody>
</table>

*Multiple cites were often used. See the complete list in the Appendix.

*Physiological measures are marked by (P).

*Adoption categories: Adopt a new behavior = A, Prevent a new behavior = P, Cessation = C.

*Larger number (max to 50) is closer to middle of diffusion curve at baseline. Studies without a baseline lack a measure of diffusion.
<table>
<thead>
<tr>
<th>Campaign topic</th>
<th>Average effect size ($\bar{r}$)</th>
<th>$SD_\rho$</th>
<th>Number of campaigns (k)</th>
<th>Pooled N</th>
<th>Average % change</th>
<th>$SD$</th>
<th>Number of campaigns (k)</th>
<th>Pooled N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seat belt</td>
<td>.15</td>
<td>.05</td>
<td>5</td>
<td>54,614</td>
<td>.15</td>
<td>.05</td>
<td>5</td>
<td>54,614</td>
</tr>
<tr>
<td>Oral health</td>
<td>.13</td>
<td>.00</td>
<td>2</td>
<td>3,465</td>
<td>.14</td>
<td>.00</td>
<td>1</td>
<td>1,465</td>
</tr>
<tr>
<td>Drinking</td>
<td>.09</td>
<td>.00</td>
<td>4</td>
<td>7,805</td>
<td>.07</td>
<td>.01</td>
<td>2</td>
<td>4,745</td>
</tr>
<tr>
<td>Smoking</td>
<td>.05</td>
<td>.03</td>
<td>17</td>
<td>79,629</td>
<td>.03</td>
<td>.02</td>
<td>13</td>
<td>75,786</td>
</tr>
<tr>
<td>Heart</td>
<td>.05</td>
<td>.03</td>
<td>4</td>
<td>5,282</td>
<td>.04</td>
<td>.04</td>
<td>3</td>
<td>4,857</td>
</tr>
<tr>
<td>Mammography</td>
<td>.04</td>
<td>.00</td>
<td>5</td>
<td>4,566</td>
<td>.03</td>
<td>.02</td>
<td>5</td>
<td>4,566</td>
</tr>
<tr>
<td>Sexual behaviors</td>
<td>.04</td>
<td>.00</td>
<td>4</td>
<td>8,044</td>
<td>.06</td>
<td>.02</td>
<td>2</td>
<td>1,697</td>
</tr>
<tr>
<td>Other</td>
<td>.08</td>
<td>.07</td>
<td>7</td>
<td>4,957</td>
<td>.20</td>
<td>.01</td>
<td>2</td>
<td>1,070</td>
</tr>
</tbody>
</table>

Note. $F(7, 168139) = 36,603; p < .000, \eta = .78$ for average effect size.
sampling error accounting for 26% of the variance across studies ($SD_p = 0.02$). Cessation campaigns had a smaller average effect size ($\bar{r} = 0.05, Tn = 56,695, k = 19$) with sampling error accounting for 17% of the variance across studies ($SD_p = 0.02)(SD_p = 0.04$). The average change in percentage of the population performing the behavior was 12% for commencement campaigns ($M$ desirable behavior = 51%), 5% for cessation campaigns ($M$ desirable behavior = 65%), and 4% for prevention campaigns ($M$ desirable behavior = 90%). Thus, consistent with H3, commencement campaigns had a greater effect than prevention campaigns, but about the same effect as prevention campaigns.

Because Snyder and Hamilton (2002) found that campaigns using messages about the enforcement of a regulation such as seat belt usage were much more successful than other types of campaigns, we reran the analysis for adoption type omitting the four enforcement campaigns (three commencement, one cessation). Excluding the enforcement campaigns, commencement campaigns were no different from the other types of campaigns, with average effect sizes of $\bar{r} = 0.05$ for both commencement and adoption campaigns ($n = 32,529, k = 21; n = 55,746, k = 18$); prevention campaigns remaining unchanged at .06 ($\eta = 0.07$). The mean change in population percentage performing the behavior among enforcement campaigns was 17% ($SD = 0.03, n = 46,771, k = 4$), 5% ($SD = 0.04, n = 20,033, k = 13$) for nonenforcement adoption campaigns, 3% ($SD = 0.03, n = 33,316, k = 5$) for prevention campaigns, and 3% ($SD = 0.03, n = 48,680, k = 11$) for nonenforcement cessation campaigns.

**Addictive Behaviors**

H1 predicted that addictive behaviors would have a smaller effect size than non-addictive behaviors. Unfortunately, smoking was the only addictive behavior represented in the campaigns, making the test of the hypothesis weaker than if there had been a variety of addictive behaviors. Nonetheless, a relatively large number of the campaigns dealt with smoking (37%). A comparison of the effect size for smoking campaigns ($\bar{r} = 0.05,SD_p = 0.03, Tn = 82,329, k = 18$) with the effect size for nonsmoking campaigns ($\bar{r} = 0.12, SD_p = 0.03SD_p = 0.06, Tn = 86,033, k = 30$) was compatible with the hypothesis that addictive behaviors have lower campaign effect size ($r = -.59, Tn = 169,362$). None of the addictive campaigns had an enforcement component. When non-smoking
enforcement campaigns were removed from the analysis, smoking campaigns still had a slightly lower campaign effect size than non-smoking campaigns ($r = -.22, Tn = 121,591$).

Within adoption categories, smoking cessation campaigns, which had an average effect size of $r = .04$ ($SD_{\rho} = .04, Tn = 49,013, k = 13$), had a lower average effect size than other types of cessation campaigns ($r = .10, SD_{\rho} = .02, Tn = 10,382, k = 7, F(1,56,693) = 19167, p = .000, r = -.50$). Note that smoking prevention campaigns ($r = .06$, discussed above) which dealt with the behavior before it was addictive, were more effective than smoking cessation campaigns ($r = .24$). Thus, cessation campaigns dealing with an addictive behavior were less successful than campaigns dealing with nonaddictive behaviors.

**Baseline Behavior Rate**

The second research question concerned which function best described the impact of baseline behavior rate on campaign effect size. We analyzed the data set with and without the enforcement campaigns, because enforcement was strongly related to campaign effect size.

For the entire data-set, baseline behavior rates ranged from 6% to 100% of the population ($M = .57, SD = .25, n = 153,894; k = 36$). We began by testing the linear models—the bandwagon and resistance effects. The linear equation was:

$$\text{Campaign effect size} = b_1 \cdot \text{bb rate} + c;$$

where \( \text{bb rate} \) was the baseline behavior rate, \( b_1 \) was the coefficients, and \( c \) was a constant. The estimated values for the linear model were \( b_1 = -.0005, c = .092, R^2 = .00 \). The standardized effect size was trivial ($\beta_1 = -.02, Tn = 153,893, k = 36$).

Next, we tested the quadratic equation:

$$\text{Campaign effect size} = b_1 \cdot \text{bb rate} + b_2 \cdot \text{bb rate}^2 + c$$

We found that $b_1 = .004, b_2 = -.00003, c = .014, R^2 = .12$, indicating an adequate fit. The shape of the curve followed the predicted inverted-U (Figure 2a).

However, when we removed the enforcement campaigns, the quadratic curve results changed dramatically. The linear model still showed no linear relationship between baseline behavior rate and campaign effect size ($b_1 = -.0002, c = .055, R^2 = .00, \beta_1 = -.014, Tn = 107,122, k = 32$). The quadratic coefficients were $b_1 = -.0026, b_2 = -.00002, c = .10, R^2 = .14$. The shape of the curve changed from an inverted-U to a U-shaped curve, directly contrary to the prediction (Figure 2b). Thus, the results were inconclusive, and hinge on whether enforcement campaigns should be analyzed together with non-enforcement campaigns.

**Discussion**

The results of the meta-analysis should be of value to campaign planners and funders as they establish goals for behavior change in future media campaigns. The results make a clear that campaign’s goals of changing 20% of a population’s behavior would be a very big challenge, and probably result in failure. By establishing more realistic goals, planners can better set funders’ and staff members’ expectations about what can and cannot be accomplished by a single media campaign.

In addition, our figures can become a benchmark against which new campaigns can be measured. Planners are frequently innovating with new channels, message strategies,
FIGURE 2 Campaign effect size by baseline rates of behavior a. All campaigns, \( k = 36 \)
b. Non-enforcement campaigns, \( k = 32 \).
and timing. Now they will be able to judge whether the innovations were helpful or not. Of course, it would be better to design a field experiment to directly test novel campaign designs, but that is not possible for many campaigns.

We examined the impact of health communication campaigns on behavior change across studies and found that campaigns on average have small but tangible effects. The effects ranged from \( \hat{r} = .07 \) to \( \hat{r} = .10 \), and in percentage terms, campaigns changed the behavior of about 8% of the population. It is crucial to remember that small percentage changes may affect very large numbers of people in a community, state, or national campaign. An 8% change in a city of 100,000 targeted adults would yield 8000 more people engaging in the desired health behavior. The modest changes caused by media campaigns could have an important impact on public health.

The campaign effects were heterogeneous, so the overall average tells only a limited part of the story. It is essential to understand the key characteristics of campaigns that moderate effect size.

Among the campaign characteristics tested in the present research, the one that explained the most variance in campaign effect size was the topic of the campaign. Seat belt, oral health, and drinking campaigns tended to be slightly more successful than campaigns on other topics. Campaign planners and researchers can use the estimated campaign effect sizes when they are dealing with the topics covered by the present meta-analysis.

For new campaign topics, it is important to know whether or not there will be enforcement messages, since campaigns with enforcement messages have greater success rates (Snyder & Hamilton, 2002). Although it appeared at first that it was easier to promote the adoption of a new behavior than to prevent a new undesirable one from starting or extinguish an old one, the effect of enforcement messages was largely responsible for that relationship. The mean change in percent of population performing the target behavior was 17% for enforcement campaigns, 5% for nonenforcement adoption campaigns, 3% for prevention campaigns, and 3% for nonenforcement cessation campaigns, among the studies that measured the percentage change.

Cessation of an addictive behavior was particularly difficult to attain in interventions. Smoking was the only addictive behavior in the present meta-analysis, and the effect size for anti-smoking campaigns, especially smoking cessation, was lower than the campaign effect size for non-addictive behaviors.

Interestingly, the results presented here for mediated anti-smoking campaigns were similar to other types of anti-smoking efforts. Our average smoking cessation media campaign effect size was \( \hat{r} = .04 \), representing a 2% gain in non-smokers due to the campaigns. A meta-analysis of phone counseling for smoking cessation found a 6% average difference between the percent of people who had stopped smoking short-term in a phone counseling group versus a control group (Lichtenstein, Glasgow, Lando, Ossip-Klien, & Boles, 1996). A comparison of youth smoking, alcohol, and drug abuse prevention and adult smoking and alcohol cessation campaign effects is in Figure 3. There were only two alcohol media campaigns aimed at adults and another two for children. Aside from the Mullen et al. (1997) meta-analysis, which mixed clinic-based smoking and alcohol cessation interventions, the results are remarkably similar across intervention techniques, with the effect sizes averaging \( \hat{r} = .03 \) to .08 for youth campaigns and \( \hat{r} = .03 \) to .04 for adult smoking campaigns. The effect of mediated adult alcohol campaigns was larger than adult smoking campaigns, although not as large as the Mullen et al. (1997) results for clinic-based interventions.

In general, the effectiveness of media campaigns was substantially less than interventions in clinical settings as reported by Mullen et al. (1997), who estimated that the
effectiveness of clinic-based interventions averaged about \( r = 0.27 \). However, an earlier meta-analysis by Mullen et al. (1992) found clinic-based education effect sizes among cardiac patients were comparable to or lower than our media campaign effect for all behavioral measures. Thus, it should not be assumed that diagnosed patients respond better to interpersonal communication for all topics. It would be interesting to explore differences in clinic-based and mediated interventions in a single study.

Note that media campaigns may be more cost-effective than clinical interventions when the goal is to reach large numbers of people, since clinic-based education is usually far more expensive per person reached than through a media campaign. If the goal is to reach a relatively small number of people, clinic-based approaches may be more cost effective. Unfortunately, neither the meta-analytic studies of clinic-based interventions nor the campaigns in our meta-analysis calculated cost-effectiveness.

The test of whether the likelihood of successful adoption depended in part on initial rates of correct behavior was inconclusive. An inverted-U, consistent with diffusion theory’s mixed-influence model (Mahajan & Peterson, 1985), was found for the entire data set. However, the enforcement campaigns were critical to determining whether the curve was inverted or not. Removing them from the analysis resulted in a U-, or cup-shaped curve. The test needs to be repeated with more studies to reach a clearer understanding of the role of baseline behavior rates.

Using our average campaign effect sizes, researchers and evaluators can estimate the sample size necessary to find media campaign effects when they exist. This is important, since many evaluations in the past seem to have used samples too small to detect the typically small effect sizes found in media campaign evaluations. For example, at 80% power, a campaign evaluation needs 2471 participants to detect an effect size of \( r = 0.05 \) at \( \alpha = 0.05 \), 1-tailed (Borenstein & Cohen, 1988).

The results were limited by the availability of published evaluations. Although comparisons of published and unpublished studies in other domains have not found evidence of a publication bias that would cause an overestimation of effects, there has not been a test of publication bias in the campaign literature. It is important for more journals to publish evaluations of campaigns because it allows peer review and wider dis-
semination of research findings than occurs with evaluation reports. Evaluation data enable scholars, as well as funders, policy-makers, and taxpayers to understand what does and does not work (Guttentag, 1977). In addition, the findings represent short-term effects, and need to be extended to long-term effects of campaigns.

Summary

Mediated health campaigns have small measurable effects in the short term. Campaign planners should set modest goals for future campaigns. Campaign evaluators should be sure to use a large enough sample size to detect small effects. As they try to predict the success rates of mediated health campaigns, researchers and campaign planners should be mindful of the average effect size for that campaign topic, and whether there is an enforcement component to the campaign. When the topic is not one that was included in the present meta-analysis, it could also help to consider

1. the type of behavior change being promoted—commencement of a new behavior, cessation of an old one, or prevention of a new undesirable behavior, and
2. whether or not the behavior in a cessation campaign is addictive.

References

*Included in the meta-analysis.


