Field Experiments and the Study of Voter Turnout

DONALD P. GREEN*, MARY C. MCGRATH** & PETER M. ARONOW**
*Columbia University, USA; **Yale University, USA

ABSTRACT Although field experiments have long been used to study voter turnout, only recently has this research method generated widespread scholarly interest. This article reviews the substantive contributions of the field experimental literature on voter turnout. This literature may be divided into two strands, one that focuses on the question of which campaign tactics do or do not increase turnout and another that uses voter mobilization campaigns to test social psychological theories. Both strands have generated stubborn facts with which theories of cognition, persuasion and motivation must contend.

For more than a century, voter turnout has attracted scholarly attention from every corner of political science. Some are drawn to the normative question of whether democratic institutions are legitimate when large segments of the electorate fail to vote or are prevented from doing so (Lijphart, 1997; Lipset, 1981; Piven & Cloward, 1988; Teixeira, 1992). Others are troubled by the distributive consequences of unequal turnout rates among socioeconomic or ethnic groups (Burnham, 1982; Hicks & Swank, 1992; Hill & Leighley, 1992; Mebane, 1994; Rosenstone & Hansen, 1993; Verba et al., 1978). Still others regard unusually high rates of turnout with concern, either because high rates suggest coercion (Shi, 1999; Zaslavsky & Brym, 1978) or fraud (Jockers et al., 2010; Myagkov et al., 2007), or because surges in turnout mean greater influence for new voters with weaker commitments to democratic values (Bennett & Resnick, 1990). The positive literature on voter turnout is no less diverse. Many scholars have sought to describe and explain over-time, cross-national or within-country variations in turnout rates (Alford & Lee, 1968; Blais & Dobrzynska, 1998; Putnam, 2000), linking turnout patterns to registration laws (Brians & Grofman, 2001; Jackson et al., 1998; Rosenstone & Wolfinger, 1978), representative institutions (Bowler et al., 2001; Endersby & Kriekhaus, 2008; Jackman, 1987; Powell, 1986), party competition (Key, 1949),
socialization and culture (Almond & Verba, 1963; Litt, 1963), patronage (Blaydes, 2010; Heckelman, 1995), and campaign tactics (Blydenburgh, 1971; Cain & McCue, 1985; Caldeira et al., 1985; Caldeira et al., 1990; Goldstein & Freedman, 2002; Goldstein & Ridout, 2002; Huckfeldt & Sprague, 1992; Kramer, 1970; Rosenstone & Hansen, 1993). Finally, a vast number of scholars have approached voter turnout from a social-psychological vantage point, explaining individual variation in voter turnout by reference to social attributes such as affluence (Rosenstone & Hansen, 1993; Verba & Nie, 1972), education (Leighley & Nagler, 1992; Nie et al., 1996; Rosenstone & Hansen, 1993; Wolfinger & Rosenstone, 1980), age (Goerres, 2007; Strate et al., 1989), or ethnicity (Leighley & Nagler, 1992; Logan et al., 2012; Timpone, 1998), or to psychological attributes such as a sense of civic duty (Campbell et al., 1954; Riker & Ordeshook, 1968), partisan attachment (Campbell et al., 1960), interest in politics (Brady et al., 1995; Palfrey & Poole, 1987; see also Denny & Doyle, 2008), or feelings of internal and external efficacy (Abramson & Aldrich, 1982; Acevedo & Krueger, 2004; Rosenstone & Hansen, 1993).

A generation ago, prior to the publication of influential monographs such as Mobilization, Participation and Democracy in America (Rosenstone & Hansen, 1993), Voice and Equality (Verba et al., 1995), To Vote or Not to Vote? (Blais, 2000), or Voter Turnout and the Dynamics of Electoral Competition in Established Democracies Since 1945 (Franklin, 2004), the extensive literature seemed to leave few openings for new contributions. After these books appeared, one might have guessed that there would be very little left to add.

Nevertheless, the literature on voter turnout has enjoyed something of a renaissance in recent years, propelled in part by field experimentation, a research method that traces its origins in the early works of Gosnell (1927) and Eldersveld (1956). An experiment is a study in which the units of observation are assigned at random with known probability to treatment and control conditions. Field experiments are randomized studies that take place in real-world settings. In the context of voter turnout research, field experiments tend to share one or more distinguishing characteristics: studies take place unobtrusively in the context of actual campaigns; the experimental interventions are campaign tactics (e.g., door to door canvassing) that campaigns deploy; the participants are actual voters; and outcomes are measured using administrative records. For example, in his 1953 and 1954 studies of voter mobilization in Ann Arbor, Michigan, Eldersveld randomly called, mailed or canvassed registered voters prior to an election, and gauged each individual’s turnout by examining public records compiled by local officials.

From the 1950s to 1980s, a small number of studies used experiments to study the effectiveness of campaign tactics (Adams & Smith, 1980; Gross et al., 1974; Miller et al., 1981), but the method was seldom used or discussed, and no field experiments on voter turnout – or any other topic – appeared in a political science journal during the 1990s. Interest in field experimentation was revived with Gerber and Green’s (2000a) study of the effectiveness of canvassing, phone calls and direct mail. Just eight years later, the second edition of Get Out The Vote: How to Increase Voter Turnout (Green & Gerber, 2008) summarized the results of dozens of experiments.
assessing the mobilizing effects of campaign tactics and messages. Dozens more have since appeared in journals or conference proceedings. The recent book *Mobilizing Inclusion: Redefining Citizenship through Get-Out-the-Vote Campaigns* (García Bedolla & Michelson, 2012) alone reports more than 100 experiments. Unlike the sporadic research that preceded it, the recent wave of experimental studies represents a broad and methodologically sophisticated research enterprise that extends beyond the United States and increasingly draws inspiration from psychological theories of persuasion and cognition. The aim of this review article is to call attention to the empirical and theoretical contributions of this burgeoning experimental literature. We begin by briefly describing how field experiments overcome important methodological weaknesses of survey research, which for decades provided the basis for causal inferences about voting. Next, we provide an overview of the experimental literature on voter mobilization tactics, summarizing the findings and their implications for theoretical propositions concerning social norms, habit formation and interpersonal influence. Finally, we discuss the ways that experiments have begun to shed light on the effects of factors other than campaign tactics, such as education, representation and institutional design.

**Field Experimentation as a Departure from Survey-Based Approaches**

Since the 1940s, the research literature on voter turnout has been dominated by analyses of cross-sectional surveys. Scholars have drawn upon a variety of academic surveys (Butler & Stokes, 1969; Campbell et al., 1960; Verba et al., 1995) and government-run surveys (Leighley & Nagler, 1992; Wolfinger & Rosenstone, 1980), but the basic features of this style of investigation tend to be similar. Voter turnout is usually measured by asking respondents whether they voted, although occasionally surveys have “validated” these reports by consulting official records of who voted. Respondents’ exposure to the “treatment” – campaign activities or electoral laws – is measured in one of two ways. The first method employs questions that ask respondents to recall their encounters with mobilization campaigns or the mobilizing efforts of friends and family. For example, Rosenstone and Hansen (1993) use the following survey item from the American National Election Studies to measure exposure to mobilization efforts:

> As you know, the political parties try to talk to as many people as they can to get them to vote for their candidates. Did anyone from one of the political parties call you up or come around and talk to you about the campaign?

The second method categorized the institutional rules that apply in the respondent’s jurisdiction. For example, several studies dating back to Kelley et al. (1967) have classified US states according to the stringency of their registration requirements, and more recent studies have measured the cost of voting by reference to the availability of mail-in balloting or early voting (Berinsky 2005; Berinsky et al., 2001; Knack, 1995; Mitchell & Wlezien, 1995; Nagler, 1991). When applied to country-
level data, this approach compares voter turnout rates among countries with varying electoral rules or systems of representation (Franklin, 1999, 2004; Jackman, 1987; Jackman & Miller, 1995; Powell, 1986).

An important limitation of this line of research is that the treatments — the campaign appeals, electoral rules and systems of representation — are not randomly assigned. On the contrary, the treatments that survey respondents receive may be systematically related to their unmeasured attributes. For example, it may be that mobilization efforts increase turnout, or it may be that campaigns direct their mobilization tactics at likely voters. In the case of electoral laws, it may be that easier registration requirements promote higher turnout, or it may be that easier registration requirements coincide with other factors that are associated with high turnout. In order to demonstrate that turnout is affected by campaign activities or electoral laws, researchers typically perform a multivariate analysis in which turnout is predicted by one or more of these treatments as well as controls for background characteristics, such as partisanship or past turnout. Whether covariate adjustment is sufficient to generate unbiased estimates of the intervention’s effect is a matter of conjecture, as it depends on the relationship between the treatment and unmeasured causes of turnout. Even if statistical controls were adequate to eliminate bias, uncertainty about whether a statistical analysis is unbiased is sufficient to undercut the credibility of the findings (Gerber et al., 2004).

Randomized experiments are designed to address the inference problems caused by self-selection, strategic targeting or unobserved confounders. Field experiments in particular are designed to assess the effects of a randomized intervention that is deployed in the context of an actual election. The use of a randomly allocated treatment has two important advantages. First, random assignment is a procedure that ensures that the potential outcomes of those in the treatment group are, in expectation, the same as the potential outcomes in the control group.2 Comparability of treatment and control groups is no longer a matter of speculation. Second, when researchers (or collaborating campaigns) administer the treatment, the nature and timing of the intervention is usually known with a high degree of precision; by contrast, survey questions that ask subjects to recall whether, when and how they were contacted by campaigns are susceptible to misreporting (Vavreck, 2007). Field experiments aid generalizability by assessing the impact of actual campaign treatments on actual voters in the midst of an election. Field experiments on voter turnout also tend to have a further advantage: they almost invariably measure turnout using administrative records rather than self-report, which means that outcomes for treatment and control groups are measured in a symmetrical fashion. When voting is measured using self-report, there is the risk that respondents who falsely report that they voted will also falsely recall being contacted by a campaign. Finally, most field experiments on voter turnout are unobtrusive: subjects are unaware that they are participating in an experiment, and outcomes are measured without their involvement. The combination of strong internal and external validity makes field experiments on voter turnout an unusually fruitful line of research, shedding light on the causes of political participation and social psychological theories more generally.
What Tactics Mobilize Voters?

The 1998 New Haven study (Gerber & Green, 2000a), which tested the effects of canvassing, commercial phone calls and direct mail, indicated that canvassing produced substantial increases in turnout, mail had weak but detectable positive effects, and phone calls from commercial phone banks had no positive effect. In the wake of that study, dozens of experiments have evaluated these tactics and considered others, such as volunteer phone banks, pre-recorded calls, e-mail and text messages. Given the sheer volume of studies, we summarize the results using meta-analysis rather than describing each study individually.

Meta-analysis is a statistical method that pools the experiments in a given domain, weighting each study by the precision with which it estimates the treatment effect. This fixed effects weighting scheme means that smaller experiments tend to be accorded less weight than larger studies. The voter turnout literature is in many ways well-suited to meta-analysis insofar as the outcomes are measured along the same metric (percentage point increases in turnout), and the treatments within a given domain are relatively similar. The main challenge is the “file-drawer problem” whereby experiments with null findings go unreported. We have made a concerted effort to compile a comprehensive set of studies, including unpublished studies. Building on the meta-analyses conducted by Green and Gerber (2008), we searched for subsequent publications and conference papers, and we contacted researchers working in this area in order to obtain information about studies that were neither published nor presented. (As expected, unpublished studies tend to have smaller effect sizes than published studies.)

We begin our discussion with direct mail because it presents the fewest technical complications. In all, 147 distinct experimental treatments assessed the effects of direct mail, sometimes sending several mailings to a given recipient. Using linear regression, we estimated the percentage-point increase in turnout per mailing. We begin by restricting our attention to what might be termed “conventional” mailings, those that neither exert social pressure (by promising to monitor whether a person actually votes) nor express gratitude for past voting or political involvement. In this restricted set of 110 studies, we find a weighted average treatment effect of 0.162, with a 95% confidence interval of (0.078, 0.247). If we further divide these 110 studies according to whether the mailing conveyed a nonpartisan encouragement to vote or rather advocated voting for a given issue or candidate, we find the two types of mailings have significantly different average effects. Nonadvocacy mailings (a diverse collection of 79 treatments that include, for example, forceful assertions of civic norms) on average increase turnout by 0.194, with a 95% confidence interval of (0.106, 0.282), whereas the 31 advocacy mailings have negligible effects on turnout: an estimated effect of 0.213 with a 95% interval of (−0.515, 0.088).

This pattern of results has an interesting implication: telling people how they should vote may affect vote preference but has little effect on turnout. Encouragements to vote, on the other hand, do boost turnout somewhat.
The effects of unconventional mailings are best appreciated when contrasted with the relatively weak effects of conventional mailings. As discussed below, “social pressure” mailings that scold voters for failing to vote, confront recipients with official records that document whether they voted, and promise to monitor whether they vote in an upcoming election produce much larger treatment effects. The 29 social pressure treatments generated a weighted average treatment effect of 2.850 percentage points, 95% CI = (2.686, 3.014). Weaker but still substantial effects were evident from the eight treatments that thanked voters for past participation or urged them to become part of an honor roll of people who vote: 1.332 percentage points, 95% CI = (0.916, 1.747). Evidently, whether mail works to increase voter turnout depends on the content of the mailing. Conventional mailings tend to have relatively weak effects; mailings that tap into social norms can have dramatic effects.

As we turn our attention to canvassing and phone calls, we confront a complication. Canvassers and callers sometimes fail to reach their intended targets. Experiments are unable to assess the effects of a treatment on the type of person that Angrist et al. (1996) dub “Never-takers,” those who would not be contacted if called or visited. Instead, experiments that fail to treat some of the members of the assigned treatment group are limited to estimating the average treatment effect among “Compliers,” those who would be contacted if visited or called. The average treatment effect for this subgroup is estimated by calculating the difference in turnout rates between those assigned to treatment and control and dividing this difference by the fraction of the assigned treatment group that actually receives the treatment (see Gerber & Green, 2012: chapter 5). For each canvassing and calling study in our meta-analysis, we use this formula to estimate the Complier average causal effect (CACE).

Analyzing the effects of 71 canvassing treatments reveals a weighted average CACE of 2.536, with a 95% confidence interval of (1.817, 3.255). In some ways this estimate understates the influence of canvassing because many of the studies target low propensity voters. For example, the Michelson (2006) study found that canvassing increased turnout among young Latino voters from 7.0% to 9.4%. Given a contact rate of 50%, the estimated CACE is 4.8 percentage points. This statistically significant increase would be even more impressive in percentage terms given the low base rate of voting.

Studies that assess the effects of GOTV phone calls fall into three categories. The first gauges the effect of assignment to pre-recorded phone calls. Excluding two experiments that tested the effectiveness of a pre-recorded social pressure message, we find a weighted average effect of 0.156 percentage points among the remaining eleven treatments, which tested the effects of messages recorded by politicians, celebrities or local clergy. The confidence interval, which ranges from -0.118 to 0.430 percentage points, indicates that the treatment effects of this tactic are very close to zero.5

Live calls from commercial phone banks have been the subject of extensive investigation, 25 treatments in studies comprising hundreds of thousands of subjects. Contrary to the findings of the initial Gerber and Green (2000a) study, commercial phone banks on average do increase turnout, but the effects are small. The average CACE is
0.980 percentage points with a 95% interval ranging from 0.504 to 1.456 percentage points. Live calls from volunteer phone banks seem to be more effective. Excluding experiments that conducted follow-up calls with respondents who had previously indicated an intention to vote (see below), volunteer phone banks generate an average CACE of 1.936 percentage points, with a 95% CI = (1.298, 2.575), over 37 experimental treatments.

The basic pattern of results seems to suggest that impersonal tactics, such as conventional direct mail or robotic phone calls, tend to produce weak effects, whereas canvassing produces much stronger effects (on those who open their doors when canvassers show up). In a similar vein, tactics that involve live interactions with human beings seem to vary in effectiveness depending on whether GOTV messages are delivered in a routinized way by a commercial phone bank or in a more authentic manner by a volunteer phone bank. This pattern seems to suggest that mass e-mails encouraging registration or voter turnout should have weak effects. That pattern is indeed borne out by several large-scale randomized studies. Nickerson (2007a) reports the results of several experiments involving tens of thousands of college students. Nickerson’s results indicate that nonpartisan e-mails do little to boost turnout rates; in fact, his meta-analysis suggests a weakly negative effect. This negative finding (which applies also to voter registration; see Bennion & Nickerson, 2011) is consistent with Stollwerk’s (2006) large-scale experiment showing that e-mails from the Democratic National Committee to New York City Democrats had a weak and insignificantly negative effect on turnout in a mayoral election. The conclusion that e-mail is ineffective, however, requires two caveats. Davenport (2012) studied the effect of personalized e-mails from senders who were personally acquainted with the people they targeted for mobilization. Although the point estimates are subject to a fair amount of sampling variability, they do suggest strong CACEs among those who open the e-mails. In three large-scale experiments, Malhotra et al. (2012) show that while mass e-mails from a nonpartisan group have no apparent effect, e-mails boost turnout by roughly one-half percentage point when the sender is the local registrar of voters.

Given these relatively weak results, it is surprising to learn that text messages sent to cell phones raise turnout. Dale and Strauss (2009) report the results of a large experiment in which text messages were sent to voters who had previously agreed to receive this type of reminder. The CACE of the text message was 4.1 percentage points, which rivals the effect of a volunteer phone call. Dale and Strauss attribute the strong effect of the text to the fact that this mode of communication provides a “noticeable reminder” to vote, but it is puzzling that this reminder worked when, for example, live reminders from commercial phone banks tend to produce weak effects. One possibility is that reminders of any sort work well given that recipients have previously opted in to receive them. But Malhotra et al. (2011), studying text messages from a local registrar of voters to recipients who did not explicitly opt in, found statistically significant increases in turnout in two low-salience elections. Although turnout increased by less than a percentage point, the base rate of turnout in the control group was very low; Malhotra et al. argue that in percentage
terms, their effects are as large as those reported by Dale and Strauss (2009). The effectiveness of text messaging represents an intriguing anomaly.

Synergy, Diminishing Returns, or Neither?

A few experiments, such as the New Haven study, used a factorial design to assess the effects of treatments separately and in combination with one another. The design allows one to assess two competing hypotheses about how treatment effects change when treatments are administered in conjunction with one another. The diminishing returns hypothesis is that voters who have been previously exposed to one GOTV treatment are less moved by subsequent treatments than they would have been in the absence of prior exposure. This hypothesis is sometimes advanced to explain the limited effectiveness of GOTV treatments in hotly contested elections, in which voters are otherwise exposed to a high volume of campaign communication. The synergy hypothesis, on the other hand, contends that previous GOTV exposure increases the effectiveness of subsequent GOTV communications, perhaps because it piques voters’ interest. Note that the synergy hypothesis is not simply that more communications produce larger effects; rather, the claim is that the marginal effect of a given form of communication is larger when it is preceded by prior communication.

Relatively few studies have employed designs that test the effectiveness of different combinations of GOTV tactics, and only a handful are large enough to be able to detect interactions with reasonable statistical power. That said, the pattern of results offers very little support for the synergy hypothesis. In the New Haven study, the interactions among canvassing, mail and phone are negative but insignificant. Ramirez (2005), studying nonpartisan messages directed at Latino voters in several states, finds weak effects of pre-recorded phone calls and direct mail, regardless of whether they are used in conjunction with one another or with live phone calls. Cardy (2005), studying an advocacy campaign in a gubernatorial election, found some evidence of synergy, but the interaction fell well short of statistical significance. A series of experiments directed at minority voters in California in 2006 and 2008 turned up no evidence of interactions between mailings and phone calls (Michelson et al., 2007). A large partisan study conducted in 2005 also found weak effects of direct mail, and no indication that mailings or phone calls enhanced the effects of canvassing. Outside the United States, the best example of a factorial design is Fieldhouse et al. (2011), which tested the effects of mail and phone and found a positive but statistically insignificant interaction. On balance, the evidence provides no consistent support for either diminishing effects or synergy; interactions between treatments tend to be weak.

Although the literature on multiple GOTV treatments on the whole casts doubt on the synergy hypothesis, a series of experiments reported by Michelson et al. (2009) suggests that early phone calls may enhance the effectiveness of follow-up phone calls. In four experiments, volunteer callers asked respondents to an initial round of calls whether they intended to vote in the upcoming election. In one experiment, everyone contacted was randomly assigned to be called a second time; in the other
experiments, only those who expressed an intention to vote were randomly assigned to be called again. Those who said yes to this initial call were randomly assigned to receive a follow-up call immediately prior to Election Day. In all four experiments, the second round call proved to be highly effective among those who had previously expressed an intention to vote. (No effect was found among those who did not initially express an intention to vote.) The estimated CACE in their experiments is more than twice as large as the typical CACE of a single phone call.6 On the other hand, two subsequent experiments with commercial phone banks in 2008 and 2010 found that follow-up calls produced no special effects on turnout, suggesting that the effect observed in previous studies may have something to do with the rapport achieved by volunteer callers.

Message Effects

To this point, we have focused our attention on the manner in which campaigns contact voters. We now take up the question of how voters respond to different GOTV appeals. Among scholars who study voter turnout, one long-standing question is the extent to which voters are motivated by the prospect of participating in a close election, either because their perceived chances of casting a pivotal vote increase or because closeness signals that something important is at stake (Cox & Munger, 1989). Another long-standing hypothesis is that voters participate out of an internalized sense of civic duty (Riker & Ordeshook, 1968), in which case messages that prime this norm may encourage voter turnout. Yet another hypothesis is that voters participate out of a sense of group solidarity, whether to their neighborhood or ethnic group (Uhlaner, 1989). Several experiments have compared the mobilizing strength of messages falling into these broad categories. Gerber and Green (2000a) found that the effects of direct mail and canvassing did not vary to any significant degree depending on whether messages emphasized the closeness of the election, the civic obligation to vote, or the importance of voting in order to make elected officials pay attention to one’s neighborhood. Michelson’s (2003) canvassing experiment indicated that Latinos were mobilized to approximately the same extent regardless of whether the canvassers’ appeal focused on civic duty or ethnic solidarity. Dale and Strauss’s (2009) text messaging experiment found effects of similar magnitude regardless of whether the message emphasized the closeness of the election or the importance of doing one’s civic duty. Enos and Fowler (2012) found that turnout was no higher in the wake of a closeness message than a reminder about an upcoming election in the context of a town that was holding a rematch of a tied election.

A number of other experiments find negligible differences in message effectiveness. For example, Dale and Strauss (2009) found no evidence that text messages are more effective when they include information about a hotline where voters could find out where and when to vote. Arceneaux and Nickerson (2010) found that criticizing the negative features of a candidate or ballot measure had roughly the same effects on turnout as extolling the positive features of the opposing candidate or opposite side of the ballot measure, a finding echoed by Barton et al. (2011).
Message effects become more substantial when turnout appeals focus on social norms. Building on early studies that decried non-voters as “slackers” (Gosnell, 1927) or reminded voters that their participation in an upcoming election was a matter of public record (Gross et al., 1974), Gerber et al. (2008) conducted a large-scale direct mail experiment that randomly varied the forcefulness with which the norm of voting was asserted. The control condition received no mail; the “Civic Duty” treatment asserted the norm of voting; the “Hawthorne” treatment not only asserted the norm of upholding one’s civic duty but informed recipients that their turnout in the upcoming election was being studied by researchers; the “Self” treatment asserted the norm of voting, presented voters with official records indicating whether they and their housemates turned out in the past two elections; and the “Neighbors” treatment presented not only the household’s past turnout but also that of several neighboring households. This study showed that turnout increased significantly with each increment in social pressure. This study has been replicated in a variety of different electoral contexts (Abrajano & Panagopoulos, 2011; Mann, 2010; Sinclair et al., 2012), and similarly strong effects have been found when voters are presented with their vote histories as part of a canvassing drive (Davenport, 2010) or told that their names will be printed in newspapers along with other voters or non-voters (Panagopoulos, 2010). Appeals that emphasize social norms sometimes fail to produce large effects, perhaps because the assertion of the norm of civic duty is fairly weak (Green et al., 2010; Matland & Murray, 2011) or because the message focuses on descriptive norms (the rate at which others vote) rather than on prescriptive norms (the obligation to vote) (Enos, 2011; Nickerson & White, 2010). When social norms are asserted forcefully, the effects tend to be quite large, and even pre-recorded phone calls conveying social pressure messages significantly increase turnout (Gerber et al., 2010).

The application of social psychological ideas to voter turnout extends well beyond the enforcement of social norms. In a series of large-scale field experiments, Panagopoulos (2011a) has demonstrated that messages that thank the voter for past involvement or interest raise turnout significantly (see also Mann, 2012: experiment 3). Although, as noted above, these effects are smaller than the effects of social pressure mailings, gratitude mailings have the advantage of generating votes without provoking anger among recipients. Another interesting application concerns the effects of asking people to predict whether they will vote and to visualize the steps they will take in order to follow through on their intention to vote. Calling experiments tend to cast doubt on the self-prophesy hypothesis, which states that people become more likely to vote if they report that they intend to vote (Cho, 2008; Nickerson & Rogers, 2010; Smith et al., 2003). On the other hand, experiments suggest that encouraging respondents to describe when, where, and how they plan to vote increases turnout (Goldstein et al., 2008; Nickerson & Rogers, 2010), although the estimated effects are not large, and some studies fail to find an effect (Cho, 2008).

In sum, messaging experiments to date have suggested that the arguments used to encourage voting (civic duty, the closeness of the race, group solidarity) tend to have relatively minor effects, whereas the norms used to frame those arguments (e.g.,...
social pressure, gratitude) have much stronger effects. Somewhere in between are messages that encourage voters to reflect on the details of how they will transform their intention to vote into an actual vote.

**Spatial and Temporal Spillovers**

The discovery of effective methods for increasing turnout opens the door to other research opportunities. One is the investigation of whether and how treatment effects are transmitted from one person to another. Scholars have long suspected that mobilizing one person might indirectly mobilize others as well. The immediate target of mobilization might become more interested in the election and talk to housemates or friends about it. Or the target of mobilization might lower the cost of voting for others, by sharing a ride to the polls or conveying information about where to vote.

Two experimental designs have been used to study interpersonal influence. One is the placebo-controlled design pioneered by Nickerson (2008). In this design, canvassers target households containing two registered voters. The voter who comes to the door is randomly treated with either a GOTV message or a placebo message about recycling. Since the treatment of the person who answers the door is random, the second-hand treatment of that person’s housemate is also random. Interpersonal influence is gauged by comparing the average treatment effect among those who answer the door to the average treatment effect among their housemates. Using this design, Nickerson (2008) finds that 60% of the immediate effect of canvassing is transmitted to the housemate. An alternative design is a multi-level experiment in which one randomly varies the density of treatments administered to geographic areas or within households. Using a social pressure mailing, Sinclair et al. (2012), for example, randomly vary the proportion of residents within nine-digit zip codes that receives treatment and randomly vary which member of a multi-voter household receives treatment. They find some evidence of within-household spillovers but no evidence of spillovers across households in the same zip code.

Another line of investigation tracks whether mobilization effects persist in subsequent elections. Gerber, Green and Shachar (2003) report that the effects of the New Haven study in 1998 persisted to the mayoral election the following year. Davenport et al. (2010) track six social pressure studies over time to see if the substantial immediate effects persist over time. Although effects appear to decay over time, the three large experiments show significant treatment effects in subsequent elections, and the remaining three studies show more equivocal but positive effects. García Bedolla and Michelson (2012) find evidence of persistence when tracking more than a dozen experiments involving minority voters in California.

One interpretation of this pattern of persistent effects is that the act of voting is habit forming. Random inducements to vote cause people to become accustomed to voting and perhaps to think of themselves as voters (see Bryan et al., 2011 for suggestive evidence on the role of self-conceptions). Consistent with this interpretation are the findings that aggregate shocks to voter turnout tend to persist over time (Atkinson & Fowler, 2011; Green & Shachar, 2000), that the introduction of low salience elections
erodes turnout in high salience elections (Franklin & Hobolt, 2011), and that young people whose birthdays make them eligible to vote in a presidential election are significantly more likely to vote in subsequent elections than those whose birthdays narrowly miss the cutoff date (Meredith, 2009). Despite ample evidence suggesting the persistence of treatment effects and the role of habit-formation, scholars have yet to sort out more fine-grained questions. Which kinds of elections are most likely to generate habits? What types of voters are most likely to form voting (or non-voting) habits?

**Experimental Frontiers**

Field experimentation has reshaped the study of voter turnout. As the experimental literature has matured, the evaluation of isolated voter mobilization tactics has gradually coalesced into theoretically-guided research programs. Thanks in part to the vast scale with which turnout experiments are often conducted, behavioral propositions about persuasion, interpersonal influence, norms, and habit formation are now studied with unprecedented methodological rigor. Experiments are said to produce stubborn facts with which any serviceable theory must contend. The facts generated by the field experimental literature on voter turnout rank among the most stubborn in all of social science.

What are the next steps in the development of this literature? One is the systematic study of treatment effects in different political contexts. Variations in electoral salience seem to affect voters’ responsiveness to mobilization appeals (Arceneaux & Nickerson, 2009), and researchers are starting to track voters over time to study their responsiveness to a series of mobilization efforts (Fieldhouse et al., 2011; Malhotra et al., 2012). Another important development is the push to test voter mobilization tactics outside the American context. Recent years have seen a steady increase in field experimentation outside the United States, with randomized voter mobilization studies taking place in Benin (Wantchekon, 2009), Brazil (de Figueiredo et al., 2011), Canada (Loewen & Rubenson, 2010, 2011), China (Guan & Green, 2006), England (Cutts et al., 2009; John & Brannan, 2008), France (Liegey et al., 2010), Georgia (Driscoll & Hidalgo, 2012), Ghana (Ichino & Schündeln, 2012), Japan (Gerber & Yamada, 2009), Mexico (Chong et al., 2011), Nigeria (Collier & Vicente, 2011). More such studies are needed within each country in order to provide a reliable assessment of how each electorate responds to an array of campaign stimuli. Although studies like Cutts et al. (2010) are designed to test whether results from American experiments apply to the United Kingdom, the literature has yet to produce a study that fields the same treatments in different countries.

How institutions shape voter turnout is another topic for further exploration. It is often assumed that experiments have nothing to contribute here because random manipulation of institutions is often difficult, if not impossible. Experimental research opportunities do exist, however. One approach is to bring institutions into existence psychologically (see, for example, León, 2011, on fines for non-voting). Quite often, some segment of the electorate is unaware of electoral rules; sometimes the same may be said of public officials (Hess et al., 2011). An experimental intervention that makes people aware of an institution may be used to assess the behavioral consequences of
introducing a new institution (see, for example, León, 2011). Another approach is to use experiments to study a hypothesis about how an institution affects turnout. For example, Amy (2002) has argued that proportional representation increases turnout by giving voters a broader menu of options; another argument is that votes for small parties are less likely to be wasted under proportional representation, and so supporters of small parties have more reason to turn out. These behavioral hypotheses may be tested in field settings. The prediction is that voters in non-PR systems would be more likely to vote if alerted about an election with a broad menu of candidate options or viable minor party candidates. A third approach is to take notice of naturally-occurring random assignments. Some local constituencies in India are randomly reserved for women candidates, allowing researchers to test whether turnout among women is affected by the viability of women candidates (Bhavnani, 2009); in some cases, lotteries determine which families are able to send their children to high-achieving schools, allowing researchers to test whether randomly assigned school quality affects parents’ likelihood of voting in local elections (Hastings et al., 2007); and senators in some American state legislatures are subject to randomly varying term lengths (Titiunik, 2011), allowing one to test whether turnout is affected by the addition of an extra race on the ballot.

Another topic on the frontier of experimental research is the causal influence of individual voter attributes. The non-experimental literature on voter turnout is replete with claims about the causal effects of attributes such as age, race or gender. There are good reasons to be skeptical of these causal inferences, but what are the experimental alternatives? It is hard to imagine how one would go about altering physical attributes, but researchers can take advantage of “Mendelian” random assignment that causes siblings to take on different attributes (Davey Smith & Ebrahim, 2003, 2004). In some cases, a randomized intervention may alter certain aspects of the socialization experiences that are thought to contribute to age, race or gender effects. For example, the Moving to Opportunity experiments altered the neighborhood environments of American children who would have otherwise grown up in public housing (Goering & Feins, 2003; Katz et al., 2001). Other attributes, such as education or affluence, are also susceptible to experimental manipulation. For example, a growing literature attempts to gauge the causal effect of schooling on voter turnout by tracing the downstream consequences of random interventions that increase educational attainment (Sondheimer & Green, 2010) or family income (De La O, 2012; Doherty et al., 2006). Even when the search for experimental opportunities comes up empty, the exercise of looking for persuasive identification strategies remains worthwhile, even if it does nothing more than remind us of how challenging it is to draw secure causal inferences in this domain.

Notes

1. Data used for the meta-analyses reported in this article are available at http://isps.research.yale.edu/research.
2. The term “potential outcomes” refers to the outcomes that a subject would express depending on whether he or she receives the treatment. Untreated subjects express their untreated potential
outcome, while treated subjects express their treated potential outcome. When treatments are administered at random, treatment is independent of subjects’ potential outcomes.

3. In order to minimize variation in the nature of each treatment, we exclude from our meta-analysis compound treatments, such as a mixture of direct mail and phone calls, and discuss them below with regard to hypotheses about synergy and diminishing returns.

4. We count experimental treatments at varying levels of aggregation, including averaging across multiple experiments in multiple geographic regions. In so doing, we underestimate the number of total treatments deployed.

5. Because studies of recorded calls have very high contact rates (which include messages left on answering machines), we consider the effect of assignment. To obtain a rough estimate of the CACE, divide the effect of assignment by 0.8. The two pre-recorded call experiments that confronted recipients with their vote histories, on the other hand, had lower contacts because they did not leave messages. These studies found stronger effects among Compliers: a weighted average effect of 1.720 percentage points, with a 95% CI = (0.306, 3.135).

6. These experiments raise the question of whether the effectiveness of campaign contacts grows stronger as Election Day approaches. Relatively little research has addressed this issue directly. (Experiments tend to assess the effects of calls made within a few days of the election.) Panagopoulos (2011b) tested the effects of nonpartisan calls from commercial phone banks conducted four weeks, two weeks and three days before a municipal election and found them all to have roughly the same weak CACE.

References

References marked with an * are included in the meta-analysis, but are not otherwise referenced in the text.


Arceneaux, Kevin, Kousser, Thad & Mullin, Megan (2009) Get out the vote-by-mail? A randomized field experiment testing the effect of mobilization in traditional and vote-by-mail precincts. Paper presented at the Annual Meeting of the American Political Science Association, Toronto, Canada.


León, Gianmarco (2011) Turnout, political preferences and information: experimental evidence from Peru, Unpublished manuscript (University of California, Berkeley).


Mann, Christopher B. (2009) The impact of cost and the role of message in voter mobilization: voter registration, vote by mail, and election day voting field experiments, PhD dissertation, Yale University, Department of Political Science.


Panagopoulos, Costas (Forthcoming) Extrinsic rewards, intrinsic motivation and voting. *Journal of Politics*.


