

I Only Have Eyes for You: Does Implicit Social Pressure Increase Voter Turnout?

Richard E. Matland
Department of Political Science
Loyola University Chicago
rmatlan@luc.edu

Gregg R. Murray
Department of Political Science
Texas Tech University
g.murray@ttu.edu

March 17, 2015

Forthcoming, *Political Psychology*

Get-out-the-vote mailers using explicit social pressure consistently increase electoral turnout; however, they often generate a negative reaction or backlash. One approach to increase turnout, yet alleviate backlash, may be to use implicit social pressure. An implicit social pressure technique that has shown promise is to display a set of eyes. Researchers contend eyes generate a feeling of being watched, which cues subjects to act in more pro-social ways to demonstrate compliance with social norms. Several studies support this argument, including two voter mobilization studies. The technique has not been widely tested, however, in the political context. In five randomized field experiments, we test the impact on turnout of mobilization mailers using eye displays. We extend previous research by testing for differences in effects between male and female eyes and across political cultures. The effects are substantively and statistically weak at best and inconsistent with previous findings.

Key words: Field Experiment, Voter Mobilization, Eyespots, Implicit Social Pressure, Voter Turnout

Voter mobilization is central to many campaigns' strategic planning, and its significance is likely to increase as the public becomes more polarized and the number of swing voters drops (Abramowitz 2010).¹ While there are a variety of techniques for turning out voters, one used by campaigns from the simplest local operation to the most sophisticated national campaign is get-out-the-vote (GOTV) mailers (Issenberg 2012). Mailers are low technology and require limited human and financial resources, so their use is ubiquitous. There is, however, a significant flaw with this tactic. The research consensus is they are not particularly effective. Green, McGrath, and Aronow (2013) conducted a meta-analysis of 147 distinct evaluations of GOTV mailers. They found that "non-advocacy" messages that primarily appealed to an individual's sense of civic duty concerning voting increased turnout by a small amount (mean = 0.19 percentage points), while "advocacy" messages telling people how to vote had no discernible effect on turnout.

The best documented exception to these limited effects is for mailers using social pressure, which are designed to gain social compliance by "play[ing] upon a basic human drive to win praise and avoid chastisement" (Green and Gerber 2010: 331). Gerber, Green, and Larimer (2008) were the first to test explicit social pressure as a technique to mobilize voters. In their study, they sent subjects a letter presenting the subject's voting history and the voting histories of the subject's neighbors. The letter also indicated an intention to send updated vote histories to everyone in the neighborhood after the election. This resulted in a striking 8.1 percentage-point increase in turnout. Since that first social pressure experiment many social pressure messages have been tested. All with a core message that vote histories are public records, so it is possible to observe people's voting behavior. The effects have rarely been as

¹ We would like to thank the electoral officials who provided us with data and Don Green for his support. This research project was pre-registered with Experiments in Government and Politics at <http://e-gap.org/design-registration>.

dramatic as those recorded in the first experiment, but in their meta-analysis Green and co-authors (2013) identify 29 studies using social pressure GOTV mailers. On average these techniques increase voter turnout by 2.85 percentage points, which is 14 times the average strength of the effect of non-partisan civic duty mailers.

While a social pressure mailer provides an effective method to increase voter turnout, it comes at a significant cost. People often react negatively to being pressured to act in a specific manner (Brehm and Brehm 1981; Wicklund 1974). In earlier work we found significant backlash against using social pressure to get voters to the polls (Matland and Murray 2013; see also Mann 2010). To gauge voters' reactions to social pressure, we treated experimental subjects with one commonly used format, the "self-mailer." This mailer includes the subjects' vote history over the last several elections (indicating the sender is aware if the subject voted in the past and suggesting the sender will follow up to see if the subject votes in the upcoming election). The self-mailer does not, however, include the vote history of one's neighbors or threaten public shaming. In our experiment we asked one set of respondents to evaluate a standard civic duty postcard, while the other group of respondents evaluated the "self-mailer." The results were strong and unambiguous: subjects who evaluated the self-mailer were significantly more likely to say the mailer made them angry and they answered they were significantly less likely to vote for the candidate who employed such tactics.

In short, campaigns face a dilemma. Mailers that fail to include social pressure appear to be largely ineffective, but messages that include significant social pressure lead to the danger of alienating voters and setting off a significant backlash against a candidate. One alternative that has proven effective at raising turnout is to express gratitude and praise voters for voting (Gerber, Green, and Larimer 2010; Panagopoulos 2010, 2011;). Another possible solution is

implicit social pressure, which subtly encourages individuals to engage in socially appropriate behavior without raising their ire by pressuring them directly.

The purpose of this research is to evaluate, using a series of field experiments, an implicit social pressure technique, the display of a set of watchful eyes, that has been shown to increase compliance with social norms. In this endeavor, we start by presenting a review of literature on the effects of eye displays on pro-social behavior. We go on to discuss the two studies done on the effects on voter turnout of mobilization mailers displaying eyes (Panagopoulos 2014a, 2014b). Then we extend this research by formally proposing hypotheses regarding heterogeneous effects related to the sex of the eyes used and the political culture of the message recipients. Next, we report the results of our voter mobilization field experiments, which were designed to retest the use of eye displays on GOTV messages and to test our hypotheses regarding the effects of gendered eyes and political culture. Finally, we discuss the unexpected results of the field experiments and attempt to put them into the broader context of the literature.

Eye Displays as Implicit Social Pressure

The social psychology literature provides evidence that implicit cues can result in individuals acting in a more other regarding manner. It makes sense for individuals to be concerned about being watched as it can affect their reputation, which directly impacts the willingness of others in the community to interact with them. We know individuals are more likely to act in a socially approved manner when they believe they are being watched (Kurzban 2001; Penner, Dovidio, Piliavin, and Schroeder 2005; Rind and Benjamin 1994), even when the perception of being watched is not triggered by an actual person but merely by an image (Haley and Fessler 2005; Oda, Niwa, Honma, and Hiraishi 2011; Powell, Roberts, and Nettle 2012; Shariff and Norenzayan 2007). A meta-analysis of seven published studies of the effect of

implicit eye cues in the dictator game in lab settings (Nettle et al. 2013) finds they reliably lead to an increase in the proportion of people voluntarily giving resources to another.

In the first study designed to test the watchful eyes effect outside of the laboratory, Bateson, Nettle, and Roberts (2006) used a set of photocopied eyes intended to induce greater support for a collective responsibility by manipulating the sense of being watched. In a university department break room there was an “honesty box” where individuals were supposed to pay for the cups of coffee or tea they took. For a period of ten weeks the researchers alternated the picture above the honesty box between a picture of eyes and a picture of flowers. They found the eyes effectively increased pro-social behavior, with contributions averaging 275 percent higher with pictures of eyes compared to the pictures of flowers. Debriefings indicated subjects had not guessed the reason the pictures appeared above the honesty box, nor were they aware of their change in behavior. This strongly suggests the effect occurred at a sub-conscious level and therefore generated no conscious reactance to this pressure.

Additional tests have found this mechanism works effectively in a variety of settings. Powell, Roberts, and Nettle (2012) found that using two eye stickers on charity collection buckets at the end of a supermarket checkout aisle, as compared to a control image of three star stickers, led to a 48 percent increase in charity contributions over an 11-week period. Ernest-Jones, Nettle, and Bateson (2011) found posters including images of eyes were more effective in getting university students to clean up their trash in the cafeteria than posters including images of flowers. This occurred both when the text message on the poster urged students to clean up and when the text message was unrelated to cleaning up.

On the other hand, as the literature on the effect of eyes as an implicit cue has developed, several studies have detected no effects (Cai, Huang, Wu, and Kou 2014; Fehr and Schneider

2010; Jolij and de Haan 2014; Raihani and Bshary 2012). This has led to further efforts to identify the conditions under which eye cues have an impact. Several studies have found context and timing matter. Ekstrom (2011) tested the effect of eyes versus other images on supermarket recycling machines that gave shoppers the opportunity either to receive their recycling deposit back as a cash credit or to donate their deposit to a charity. His effects are distinctly moderated by traffic volume. On busier days when there was substantial traffic, and effectively there were numerous “real eyes” to view a person’s other regarding behavior, the eye images had little to no noticeable effect. On slower days when there were relatively fewer “real eyes” to observe the person’s behavior, the eyes images had substantially larger effects on the behavior of those who were exposed to them. Both the study of eye spot posters in the university cafeteria and the study of eyes on charity baskets at the end of the supermarket checkout aisles also found that effects varied markedly by traffic volume. The eye spots raised contributions on days with limited traffic but not on days with heavy traffic.

Implicit eye cues appear to be ineffective when dealing directly with an individual in a dyadic relationship. Fehr and Schneider (2010) find implicit eye cues have no effect on the actions of an individual in a trust game where reciprocity is specific to the person one is playing with and not a generalized public. They argue the assessment of the individual one is interacting with dominates the decision independent of any implicit cooperation cue.

Further, most eyes studies, whether in the laboratory or the field, test the impact of eyes on a decision that is taken merely seconds after the eyes are introduced. Sparks and Barclay (2013) argue the effect of implicit eye cues is distinctly time bound with a rapid decay rate. Running their own experiment and conducting a meta-analysis of 25 laboratory studies afterwards, they find short initial exposure to an image of eyes just prior to making a decision

leads to the expected pro-social behavior, but extended exposure to an image of eyes prior to decision making leads to the eyes having no effect.

To summarize, these studies find an effect for eyespots but make clear the effect is not universal. Some of the caveats found in this literature are unlikely to be of concern for our GOTV experiments. Fehr and Schneider's finding that eye spots do not have an impact in binary interactions is of little concern as doing ones' civic duty and voting is not a binary interaction. On the other hand, Sparks and Barclay's work showing the effect of eye spots tends to be transitory could be highly relevant. We will have more to say about the existing literature and our GOTV field experiments after we have presented the initial results.

Implicit Social Pressure: Using Eyes To Increase Voter Turnout

The possibility of applying implicit social pressure to voter turnout was first recognized by Costas Panagopoulos (2014a) and tested using a set of GOTV mailers in a mayoral election in Key West, Florida. Panagopoulos tested three separate visual images, which are presented in Figure 1. On one placebo mailer he used a picture of a sandy beach with a palm tree, on a second placebo mailer he used a picture of the U.S. flag, and on the third mailer he used a picture of female eyes. All three mailers included the same generic civic duty message. Panagopoulos found in the context of a very low turnout election that the mailer with the eyes increased turnout by a statistically significant 1.1 percentage points. Turnout was 3.9 percent for the control group, which received no mailer, and 3.7 percent and 4.0 percent in the two placebo groups. Turnout was 5.0 percent for those receiving the eyes mailer.

<FIGURE 1 ABOUT HERE>

Panagopoulos replicated this study (2014b) in Lexington, Kentucky, during the 2011 gubernatorial election, which was a much more visible and competitive election. He used the

same message and placebo mailers. For this test, though, he used a different image: the eyes of a somber looking man, which are also presented in Figure 1. Once again the mailer generated statistically significant results. The control group turned out at 23.2 percent, while the two placebo groups turned out at 22.3 percent and 24.7 percent. The treatment with eyes produced a statistically significant 2.3 percentage point increase in turnout to 25.5 percent.²

This is where the literature on implicit social pressure and voter turnout stands. There are two experimental studies both showing positive effects. As Panagopoulos (2014a: 30) notes, “extension and replication [of his results] are necessary to converge on underlying parameters of interest with greater precision and to examine further the external validity and generalizability of the findings.” More broadly, two experimental tests are insufficient to establish the external validity of a newly proposed theory. McDermott (2002: 335) indicates the external validity of experiments “is established over time, across a series of experiments that demonstrate similar phenomena using different populations, manipulations, and measures. External validity occurs through replication.” Therefore, our first hypothesis is simply a replication hypothesis:

H1: Mailers using eyes (implicit social pressure) will increase turnout in comparison to a control group not receiving a mailer and in comparison to a group receiving a placebo mailer.

Beyond Replication: The Effects of Female Versus Male Eyes.

We also desire to develop this literature further both theoretically and practically by testing for effects using a different set of eye treatments and in contexts that are notably different with respect to political culture and voter turnout. In reviewing the research, we find notable parallels between the research on social pressure mechanisms in the mobilization literature and

² As one reviewer pointed out, this 2.3 *percentage point* increase in turnout represents a 9.9 *percent* increase over the control group. Similarly, while the 1.1 *percentage point* increase in his first turnout experiment is relatively modest, proportionately it represents a dramatic 28.2 *percent* increase over the control group.

the research on the role of pro-social behavior in reputation building in the evolutionary psychology literature (Barclay 2010; Piazza and Bering 2008; Sylwester and Roberts 2013). Generally speaking, the evolutionary psychology literature suggests modern humans possess psychological mechanisms that evolved as a result of the social and environmental conditions and selection pressures experienced by ancestral humans during the Pleistocene Epoch beginning about 2.6 million years ago and ending about 12,000 years ago (Tooby and Cosmides 1992). According to this view, these psychological mechanisms, which evolved to solve social and physical problems faced by ancestral humans, influence but do not determine modern human behavior. This lingering effect is the result of the ponderous speed of evolution in stable environments—small evolutionary changes usually take between 1000 and 10,000 generations or, in human terms, between 20,000 and 200,000 years (Mayr 2001).

Anthropological evidence suggests humans' evolutionary environment consisted of smaller groups of 10 to 30 individuals that existed within larger social groups of 150 to 500 individuals. These groups were loosely held together through male kin-bonding, in which generations of males stayed together to form reliable social relationships while females dispersed from their natal groups to other groups (Foley 1995). A plausible consequence of male kin-bonded ancestral groups is that individuals would have been more concerned with their reputation with their group's males, who formed the core of their group's social network, than their group's females, who lacked the depth of kin connections and permanence in their group. Translating to present day behavior, it is plausible that these vestigial forces could lead to prioritizing male approval in groups. If so, then male eyes would exert more implicit social pressure than female eyes. Such an effect would be consistent with work that finds a distinction in the impact of eyes images on sharing with in-group members, where a strong positive effect is

generated, and sharing with out-group members, where no effect is created (Mifune, Hashimoto, and Yamagishi 2010). A cursory glance at existing data finds initial support for this idea.

Bateson and her co-authors' (2006) results indicate increases were always greater in response to male eyes than female eyes. Furthermore, the absolute increase in turnout in Panagopoulos's experiment using male eyes (2014b) was twice as large as the increase in the experiment using female eyes (2014a), but this difference, while substantively large, is not statistically significant. More importantly, these results are from two different elections, which makes direct comparisons problematic. Panagopoulos's results and the Bateson team's results, as well as the potential evolutionary motivation to more strongly protect one's reputation among male group members, justify a more rigorous test of the effect of male eyes compared to female eyes. Hence our second hypothesis is:

H2: Mailers using male eyes will lead to a greater increase in turnout than mailers using female eyes.

Beyond Replication: Tests Across Political Cultures.

We also wish to test the impact of this form of implicit social pressure across multiple political cultures. Elazar (1972) distinguishes between moralist, individualist, and traditional political cultures. Panagopoulos (2014a, 2014b) found similar effects in his two experimental locations: Lexington, Kentucky, and Key West, Florida. Elazar classifies Kentucky as a traditional political culture as does Miller (1994). Key West is more difficult to classify. There are many ways to describe Key West, but one thing we feel quite confident in is that it is not a "typical American town." Key West's marketing campaign describes it as "close to perfect, far from normal" (FloridaKeysTV 2012). It would be hard to identify a large number of places in the United States that are similar to Key West.

Social psychologists find that as a norm becomes more salient, norm-consistent behavior increases (Cialdini and Goldstein 2004). Applying Elazar's framework we believe voting will be more highly valued in moralistic states, which emphasize the importance of collective action and wide political participation, than in individualistic and traditional states, which emphasize individual action and participation dominated by political elites. One indication this is true is that average voter turnout is considerably higher in moralist than individualist and traditionalist cultures. We have no clear expectation about the relative effect of implicit pressure in traditionalist versus individualist states, but we are interested in testing the effect in all three cultures to determine if there are variations by political culture. Therefore, our third hypothesis is:

H3: Mailers using eyes will be more effective in increasing turnout in localities representative of moralist cultures than in localities representative of traditionalist or individualist cultures.

Experimental Design

Site Selection. To test the impact of implicit social pressure using eyes on voter turnout, we selected five research sites that vary by political culture, election turnout, and intensity of the campaigns. Table 1 presents the locations, election dates, Elazar's classification of political culture, the election involved, and the official turnout. Four of the five elections took place in November 2013. Most of the sites are easily identifiable in terms of political culture. The four Virginia counties were in far Western Virginia, which we define as a traditional political culture.³ Minneapolis is a stronghold of moralist political culture, while Toledo falls solidly in

³ The four counties were Montgomery County (excluding the university town of Blacksburg), Pulaski, Smyth, and Tazewell.

the individualist belt. We define Midland as individualist. Midland is in West Texas, and the farther West one goes in Texas the weaker the traditionalist culture gets and the more dominant the individualist culture becomes. Finally, we debated how to define El Paso. While it is in West Texas, the city has changed so radically in the almost 50 years since Elazar developed his system that it is difficult to be confident in his classification.⁴ El Paso is a large city with a Latino majority that, contrary to much of the rest of the state, votes heavily Democratic. Nevertheless, El Paso retains significant elements of Elazar's individualism with politics seen as a manner for groups and individuals to get ahead. El Paso is clearly more individualist than either moralist or traditional, but it is not unreasonable to argue the political culture is distinct from conventional individualism and is typical of cities along the Texas/Mexico border.

<TABLE 1 ABOUT HERE>

Treatment Description. At each site we used three different mailers, which are presented in Figure 2. Each included the same standard civic duty text, but one displayed female eyes, the second male eyes, and the third a flag. Should we find a strong effect for the eyes mailers, the flag mailer serves as a placebo allowing us to test whether the effect is a function of implicit social pressure or of subjects who are highly responsive to a GOTV mailing using the civic duty text. The eye images were computer generated to be similar yet distinct enough that one set is obviously female while the other set is clearly male.⁵ We used FaceGen Modeller, which is 3D

⁴ We harbor a fair amount of skepticism towards Elazar's typology. On the other hand, we have seen no classification scheme that we find more convincing, and we believe it makes sense to choose sites to maximize variance on the independent variable of political culture. As such, even if Elazar's typology is not perfect, we believe it helps ensure variation in political context.

⁵ A manipulation check for estimated sex of the displayed eyes strongly suggests the experimental subjects perceived the cue correctly. For each set of displayed eyes, only one of 106 survey respondents misclassified the sex of the eyes ($p < 0.001$ one tailed for each picture). The manipulation check was conducted December 1-3, 2014, using Amazon Mechanical Turk (MTurk) workers, who completed a 23-question survey delivered online via Qualtrics. Each MTurk worker was paid \$0.50, and the average time to completion was just under 3.5 minutes. The convenience sample consisted of 106 respondents who self-identified as at least 18 years old and eligible to vote in the United States and who had a computer server connection physically within the United States.

face modelling software, to create the faces from which the eyes were captured.⁶ The intent was to standardize the faces as much as possible while manipulating their levels of femininity and masculinity. We used only the eyes, not a complete face. We were concerned that a full picture of an individual could be misinterpreted as a message from a specific candidate asking the subject to vote for her or him.

<FIGURE 2 ABOUT HERE>

Sample Selection. At each of our five sites election officials provided current lists of registered voters approximately 30 days before the relevant election. We followed the same procedure at each site. First, we eliminated addresses with more than four registered voters on the assumption these were largely apartment buildings where there would be a great deal of turnover and, therefore, a large number of incorrect addresses. Second, we randomly selected one voter in each of the remaining households, which ensures no household is included in more than one condition. Finally, we blocked subjects based on voting propensity at each site, when we had such data (everywhere but Virginia), and randomly assigned subjects to the three treatment groups or the control group, insuring each subsample was representative of the overall population with respect to voting habits. After the election the election officials provided validated voter turnout information, which we use as the dependent variable.

Randomization should provide statistically equivalent experimental groups. We tested for equivalency across the experimental groups at each site. Online Appendix A shows comparisons on a number of variables including turnout rates for earlier elections. Online Appendix A indicates one of the 20 variables tested manifests a statistically significant difference across

⁶ The “Generate” settings were approximately female (or male) for Gender, 35 for Age, “average” for Caricature, “symmetric” for Asymmetry, and “all races” for Race Morphing We followed this process to ensure any differences in effect were the result of differences in the eyes and not differences in other facial features.

samples (sex is unevenly distributed across samples in Virginia).⁷ Of course 1 out of 20 is what we would expect by chance. There are no other statistically significant differences across the samples. Multinomial logistic models regressing group assignment on the independent variables presented later yield tests of the joint significance of the covariates with p-values of 0.19 (Virginia), 0.43 (Minneapolis), 0.71 (El Paso), 0.75 (Midland), and 0.85 (Toledo). These indicate the subsamples are largely equivalent statistically. Nevertheless, after considering the bivariate results we report multivariate results, which should provide more precise estimates of the effects.

Study Execution.

All the mailers were produced by the same printing shop and sent via first-class, pre-sorted mail six days before the election. To monitor whether the mailers reached their destinations on time, we also mailed postcards to 19 confederates in close proximity to the research sites. Fourteen confederates received the mailing the Friday before the election, three received it the next day (i.e., Saturday), and the remaining two never received their GOTV mailing.⁸ In terms of treatment group sizes, we used slightly over 3000 subjects per treatment group in El Paso and 2000 per treatment group at the other four sites. Panagopoulos had 1500 subjects per treatment group in Key West, Florida, and 1000 subjects per treatment group in Lexington, Kentucky.

⁷ Voter registration lists varied in terms of the information they provided. One site (Virginia) had no past turnout data. Registrant sex was not included for two sites (Minneapolis and Toledo). For the chi-squared analyses presented in Online Appendix A for voter turnout and sex, the analysis is for the 2×4 cross tabulation formed by the variable and the four treatment groups. STATA would not calculate chi-squared for a 70×4 crosstab. Therefore, for age the samples were split into four age groups: 18-29, 30-44, 45-59, and 60 or greater and cross tabulated with the four experimental groups.

⁸ The two that were never received were part of the Toledo mailing, although both were for addresses outside of Ohio. With full coverage in four of five sites, we feel comfortable asserting the mailings landed on time. The failure of the two mailers to land as part of the Toledo mailing does suggest, however, that if results appear weaker in Toledo than everywhere else the failure to deliver the message might be an explanation. Fortunately, this does not occur.

Bivariate Results

Table 2 presents turnout and treatment effects for the five sites. Columns 1 and 2 identify the site, the treatment condition, and the sample size. Column 3 presents turnout in the control group, which received no mailing. Column 4 presents turnout for the subjects receiving the placebo mailer with a picture of the American flag. Column 5 indicates turnout rates for those receiving the two eyes mailers. Column 6 shows the average treatment effects for the treatments compared to the control group, while column 7 shows the average treatment effects compared to the placebo group.

<TABLE 2 ABOUT HERE>

The results reported in column 6 provide little support for Hypothesis 1, which asserts that a picture of eyes can generate implicit social pressure sufficient to increase voter turnout. In no case do the female eyes increase turnout relative to the control group at conventional levels of statistical significance. There is a single effect in which male eyes, when compared to the control group, led to a statistically significant bump in turnout of 1.83 percentage points in Midland ($p = 0.05$, one-tailed test). With only one of 10 individual tests meeting the statistical standard, the overall picture is one indicating implicit social pressure using eyes fails to raise turnout when compared to the control group.

Column 7 reports the treatment effects relative to the placebo group. While there is one significant effect (male eyes in Toledo) and one close to significant effect (female eyes in Toledo), these are not driven by a strong increase in turnout among those receiving the mailers with eyes. Instead, they are driven by an unusual 2.4 percentage point drop in turnout among those receiving the placebo mailer relative to the control group. While it is possible there is some contextual condition in the mayoral race that explains the unusual negative effect of the placebo

mailer in Toledo, we believe a more likely explanation is that the drop in turnout is an artefact of sampling variability such that the placebo group is composed of individuals who were slightly less likely to vote in this election than individuals in the control group.

In addition to not supporting Hypothesis 1, the bivariate results also fail to support Hypotheses 2 and 3. Hypothesis 2 predicted the male eyes would have a stronger effect than female eyes. This does not occur. The female eyes have a stronger impact than the male eyes in three of the five locations. More relevantly, in no case does the difference in effects between the male and female eyes approach a statistically distinguishable level. Simply put, there is no difference in the effects. Hypothesis 3 predicted that voters in a moralist political culture would be more sensitive to the implicit social pressure than voters in individualist or traditional political cultures. In fact, the results in column 7 suggest moralistic Minneapolis may be the least sensitive to implicit social pressure, as both male and female eyes failed to raise turnout compared to the placebo. More reasonably, we can say there appears to be no noticeable difference in effects based on the political culture of the sites.

Since there is no meaningful difference between the effects for female and male eyes, we can combine them and retest the eyes mailers to see if the increased statistical power leads to a significant effect. As can be seen in Table 3, when we do this the results remain the same. Compared to the control group, those receiving any eye mailer were no more likely to vote at four of the five sites. The average treatment effect is greater than 1 percentage point at one of the five sites (1.46 percentage points in Midland) and that effect is statistically significant ($p = 0.04$, one-tailed test). When comparing the eye mailers to the placebo mailer, there is a single significant effect (2.63 percentage points in Toledo), but this is driven by the very low turnout for the placebo mailer. When we compare the eyes mailers in Toledo to the control group, which

received no mailer, the Toledo eye mailers averaged a meager 0.21 percentage point increase in turnout.

<TABLE 3 ABOUT HERE>

To summarize the existing work we conducted two fixed effects meta-analyses of the results (Gerber and Green 2012: 361-362). First using our five sites, which include more than 20,000 treated subjects, and then our five sites plus Panagopoulos's two sites. To combine these results we must assume each site represents an equivalent test and the samples are drawn from the same population. This is easy to assume for the meta-analysis of our five sites. Including the two studies by Panagopoulos requires a slight loosening of the standards as to what is "equivalent." We use different pictures and sampling frames (single and multi-voter households versus single voter households only), but the text is the same and the theory behind the stimuli is exactly the same.

The first meta-analysis estimates the effect of our eyes mailers compared to the control group, which received no mailer, as a 0.48 percentage point increase in turnout (s.e. = 0.31; $p = 0.06$, one-tailed test). The estimated effect when compared to the placebo is a 0.60 percentage point increase in turnout (s.e. = 0.51; $p = 0.12$, one-tailed test).⁹ These results indicate the eyes mailer fails to reach standard levels of significance. If we use a generous standard for statistical significance we can suggest the mailer increases turnout by about a half percentage point compared to a control group receiving no mail. In other words sending out 1,000 GOTV mailers with eyes is likely to lead to approximately five more people voting, at a cost of approximately \$50 per vote. At this level of impact, even if the results skirt the point at which they are statistically significant, the substantive impact is so small that it is likely many campaigns would

⁹ As previously noted, the effect for the eyes mailers vis-à-vis the placebo flag group is largely driven by the low turnout among those who received the placebo in Toledo. The treatment effect estimate for the four sites excluding Toledo is 0.24 percentage points (s.e. = 0.55; and $p = .34$, one-tailed test).

choose more cost effective techniques such as canvassing (\$29 per vote) or volunteer phone banking (\$38 per vote) (Green and Gerber 2008). Furthermore, while the eyes mailers do appear to produce a small bump compared to receiving no reminder, our tests generate limited evidence they are any more effective than a standard civic duty mailer as represented by our placebo. In other words, including the eyes on the civic duty message fails to increase turnout significantly. More broadly, at none of our sites did the eyes mailers meet the standard of being significant when comparing the treatment group against both the control group and the placebo mailer.

In short our study shows that implicit social pressure using watchful eyes holds little promise as an effective tool, but what is the message if we include all studies? Doing a meta-analysis of all seven tests including the two Panagopoulos studies increases the estimate of the effects, as expected. The pooled estimate of the mailer impact vis-à-vis the control group is a .72 percentage point increase (s.e. = .26, $p < .01$, one-tailed test). The pooled estimate of the mailer impact vis-à-vis the flag placebo is a .82 percentage point increase (s.e.=.41, $p < .03$, one-tailed test). Compared both to a control group that received no mailer and to a group that received the flag placebo mailer, the results show the eyes mailer has a statistically significant effect on voter turnout. The estimated effect increases, although it remains modest and is still below a 1 percentage point increase in turnout. The revised estimate suggests sending out 1,000 eyes mailers would lead to approximately eight more voters.

When we consider all the results, including the Panagopoulos studies, it seems the question of the effect of the eyes mailer is not fully decided. The best estimate is that the watchful eyes mailer does have a statistically significant effect. The effect is quite modest, however, and we do not believe it meets a standard of being substantively significant even if it does meet the statistical standard. From our viewpoint, adding watchful eyes to a civic duty

GOTV mailer does not increase the effect. The totality of the data is ambiguous enough, however, that the question is still open. Just as we did not see Panagopoulos's two positive effects as the final word on this question, we do not see our five negative results as closing the question either.

Multivariate Results

Online Appendix B reports a series of multivariate probits that estimate the treatment effects at each site. The included controls vary depending upon the information provided by the government agency responsible for distributing election data but primarily include demographic characteristics (sex, age, and years registered), partisan identifiers indicating a person has voted in either the Democratic or Republican party primaries, a dummy variable for whether the individual voter included in the sample came from a household with a single registered voter or multiple registered voters (2, 3, or 4 voters), and validated turnout in previous elections.

Figures 3a-e show estimated turnout with confidence intervals for each of the groups across the five sites. The tables associated with each figure show the average treatment effects controlling for the available variables. Figures 3a-e are based on the probits reported in Online Appendix B. The point estimates move slightly with the addition of controls, but the results still lead us to reject Hypotheses 1, 2, and 3. The exceptions in the multivariate analyses are the same ones noted in the bivariate evaluations. There is a statistically significant positive impact for male eyes in Midland and a significant negative impact for the placebo mailer in Toledo. As with the bivariate results, none of the treatments are significant when comparing against both the control group and the placebo mailer. The overall conclusion of the multivariate analyses is the

same as the bivariate analyses: there is very limited evidence that implicit social pressure using eyes increases turnout beyond the impact of the average civic duty mailer.

<FIGURES 3a-e ABOUT HERE>

Discussion

The results are unambiguous. Tests at five different sites offer very limited evidence of an effect of implicit social pressure using eyes on voter turnout. Only one of ten multivariate tests (male eyes in Midland, TX) was statistically significant. The results of the meta-analysis, which substantially strengthens our ability to find significant effects, suggest there may be a small positive effect of the eyes mailers relative to the control group but when we compare the placebo mailer to the eyes mailers the difference is indistinguishable. Even if we ignore the significance level, the eyes mailer does not consistently outperform the placebo mailer. The placebo mailer had a bigger average effect than the combined average effect of male and female eyes at two sites, while the two eyes mailers had a bigger average effect than the placebo mailer at three sites. We reject H1 concerning the impact of implicit social pressure using eye displays. We reject H2 concerning greater sensitivity to male than female eyes, and we reject H3 concerning greater impact in moralist than individualist and traditional political cultures.

The obvious question is, How does one explain our largely null results in light of Panagopoulos's (2014a, 2014b) findings? Panagopoulos found statistically significant increases in turnout of 1.1 and 2.3 percentage points, effects that were two to four times the size of our effects. Two possible explanations can be rejected as inconsistent with our data. First, political culture cannot explain the differences. Panagopoulos found a significant effect in Kentucky, while we failed to find effects in Virginia, which has roughly the same traditional political culture. Second, heterogeneous treatment effects related to base turnout levels cannot explain the

differences. Panagopoulos found an effect in the Kentucky gubernatorial election where turnout was 29 percent. Our null results with official turnout at 25 percent in Toledo and 33 percent in Minneapolis make it implausible that Panagopoulos's significant results in Kentucky are the result of that turnout rate being uniquely sensitive to implicit social pressure. Further, official turnout at our five sites showed considerable variance: from 15.2 percent in the El Paso mayoral election to above 40 percent in Virginia's gubernatorial election, yet none of the outcomes met the standard of being significant when comparing the treatment group against both the control group and the placebo mailer.

Methodological differences may offer some insight into the contradictory results. We drew our sample from households with no more than four voters, while Panagopoulos limits his sample to households with a single voter. Perhaps living alone or being the only voter in a household makes the eyes message more effective. In a multi-voter household there is already someone who is likely to be aware if one fails to vote, while this is not true in a single-voter household. Therefore, the impact of implicit social pressure from being watched may be greater in single-voter households. This hypothesis is consistent with the studies described in the literature review that found the effects of eye spots in field experiments were distinctly moderated by traffic volume (Ekstrom 2011). On days when there was substantial traffic, and effectively there were many "real eyes" to view a person's other regarding behavior, the eye images had no noticeable effect or a very small effect. On days when there were relatively few "real people" in the supermarket to notice charity contributions or in the cafeteria to notice people cleaning up, the eyes had substantially larger effects on the behavior of those who were exposed to them. A reasonable parallel would be that the eyes work in single-voter households

(which Panagopoulos used) but would be less effective across multi-voter households (which represented most of the households we used).

While being in a single-voter household, as opposed to a multi-voter household, meant an individual was less likely to vote (as seen by the strong negative and statistically significant effect of the single-voter household dummy variable on overall turnout for all five sites in Online Appendix B), this does not speak to whether voters in single-voter households are more sensitive to implicit social pressure. To evaluate this possibility we compared the treatment effects for single- and multi-voter households at the five sites. There are two sites where the eyes mailers have a stronger effect in multi-voter households and three where the effect is stronger in single-voter households. Most of the differences are trivial. We formally tested for the difference in effects of the eyes mailers by including an interaction term for single-voter household and receiving the eyes mailer. We ran these probits both separately (testing for the effect of two interactions, one for female eyes and one for male eyes on single-voter households) and together (combining both sets of eye mailers as one variable) across our five sites. When testing the male and female eyes separately, only one of the ten interactions was significant (in Minneapolis the female eyes mailer led to an increase in turnout among voters in single-voter households but had no effect in multi-voter households). An effect continues to appear for Minneapolis when we combine the treatments, as voters in single-voter households in Minneapolis are more likely to respond to the eyes mailer regardless of sex by voting. Outside of the significant effect in Minneapolis, none of the interaction terms were statistically significant. In short there is some limited support for the idea that eyes are most effective in single-voter households where they

might be substituting for pressure from real people, but this does not appear to be a plausible explanation across all of our sites.¹⁰

Looking at the additional research on eye spots, one factor found to be relevant in previous studies was timing. Timing may be central also in the context of GOTV mailers. As noted in the research review, the impact of eye spots has been found to be affected by exposure time. An impact appears if decisions are made immediately after first seeing the eye images, but not when eye images are visible for an extended period of time prior to a decision. Time may be relevant in our case as there is a much longer lag period between exposure to the image and the act we are seeking to influence than for any of the psychology experiments. In our study most people received their GOTV postcards and looked at them three or four days before the election. The failure to generate an effect may be because of the time gap between receiving the stimuli and the act of voting. Note this may explain the failure of our mailers to have an impact, but it cannot explain the difference between our results and Panagopoulos's as his respondents also waited three to four days after getting their mailers to actually voting.

This leaves two plausible explanations. First, the effect may be dependent upon the eyes that are used. There is nothing in the theory that suggests the type of eyes should matter, but this study and Panagopoulos's studies use different eyes and get different results. Perhaps viewers discerned our eyes were computer generated and therefore they had less of an effect. Note, however, the existing research leans against this argument. Much of the research has used "eye spots," which are simple images that are not eyes but can be interpreted as eyes, rather than

¹⁰ A variant on this suggestion is that perhaps high-propensity voters are more sensitive to the eyes mailers, because they would feel more pressure to comply with social norms regarding voting. We tested for heterogeneous treatment effects based on vote propensity using an interaction effect at the four sites where we had voting history. We identified voters as habitual, regular, occasional, rare, or as registered non-voters using vote histories (see Matland and Murray [2012] for a precise description of how voters are categorized). High-propensity voters were defined as habitual or regular voters (i.e., voters who voted in approximately half or more of all elections). The probit coefficients show very little indication that high-propensity voters were more strongly influenced by the eyes mailers. None of the interactions approached statistical significance (probits available from authors).

pictures of completely formed eyes (Haley and Fessler 2005; Powell, Roberts, and Nettle 2012; Rigdon, Ishii, Watabe, and Kitayama 2009). It may be our eyes were less prominent and therefore had less of an impact. Our eyes were at the top of the message and took up about one-quarter of the mailer. Panagopoulos's eyes were next to his message and took up about half the space on his mailers.

Finally, it may be the differences are simply the result of chance. Differences in outcomes can occur because of sampling variability. Experiments are powerful methodological tools, but multiple tests across different conditions are needed to establish the external validity of any experimental outcome (McDermott 2002). Furthermore, with journals biased towards publishing statistically significant results (Gerber, Green, and Nickerson 2001; Gerber, Malhotra, Dowling and Doherty 2010; Ioannidis 2005), there is perhaps a need to re-emphasize the central role that repeated testing across different sites and manifestations of a concept should play in theory assessment. Experimentalists are quick to warn that individual tests do not definitively prove a theory. Panagopoulos for example explicitly calls for further testing of implicit social pressure in both of his articles. The dangers of overgeneralizing effects are hardly unique to experimentalists, but perhaps experimentalists are especially sensitive because of a general awareness regarding external validity concerns. The concern is legitimate—we must be wary of drawing definitive conclusions about a treatment based on limited testing. As demonstrated by this and other research (e.g., Murray and Matland 2014), multiple tests may cause us to reconsider the effectiveness of a mechanism or force us to develop more nuanced theoretical models that can explain when an effect is to be expected or not.

Conclusion

We believe this work is a positive step forward. We find a consistent set of negligible effects across a number of sites. We are not ready to say implicit social pressure in the form of eyes is ineffective, but we do believe our results make it clear the effect is considerably more circumscribed than one might have assumed initially. This pattern replicates what we see with the use of eye images in economic games and public regarding behavior in field experiments. After the initial field experiment by Bateson et al. (2006) there were a series of studies replicating the success of Bateson et al. (Ekstrom 2011; Ernest-Jones et al. 2011; Powell et al. 2012; Rigdon, et al. 2009). The most recent replications, however, include a string of tests where the eye treatments have failed to generate effects (Cai et al. 2014; Jolij and de Haan 2014; Raihani and Bshary 2012; Sparks and Barclay 2013). These newer studies have identified a series of conditions where the expected effects do not appear. Therefore the literature on using eyes is slowly moving in the direction of considerably greater complexity. As Jolij and de Haan (2014: 9) note after reporting a series of non-effects: *“Although it would be the dream of many idealists to make people nicer or smarter using such an easy manipulation of just showing an image, ‘social priming’ simply does not work that way. We are only just beginning to uncover the exact underpinnings of the effect.”* We see our results as a first step towards that greater complexity in understanding implicit social pressure and voter mobilization. At this point there is a need for more tests with different operationalizations and contexts to see if the initial effects were entirely by chance or were due to a combination of identifiable factors that can explain the present inconsistent findings.

ONLINE APPENDIX A: TESTS FOR GROUP EQUIVALENCY: Means and Chi-Squared Tests Across Groups at Each Research Site

Virginia: Counties of Montgomery (not including Blacksburg), Pulaski, Smyth, and Tazewell: November 5, 2013 Gubernatorial Election

<u>Variable</u>	<u>Control Group</u>	<u>Female Eyes Group</u>	<u>Male Eyes Group</u>	<u>U.S. Flag Group</u>	<u>Chi2</u>	<u>Df</u>	<u>p</u>
Turnout in 2012 General	NA						
Turnout in 2012 Primary	NA						
Turnout in 2010 Primary	NA						
Sex (% Female)	54.20 (.22)	54.45 (1.11)	55.75 (1.11)	51.30 (1.11)	8.72*	3	.033
Age (Ave.)	53.4 (.09)	53.5 (.40)	53.4 (.40)	53.3 (.40)	0.14	9	.99
N	43070	2000	2000	2000			

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

Minneapolis, MN: November 5, 2013 Mayoral Election

<u>Variable</u>	<u>Control Group</u>	<u>Female Eyes Group</u>	<u>Male Eyes Group</u>	<u>U.S. Flag Group</u>	<u>Chi2</u>	<u>Df</u>	<u>p</u>
2012 State General	95.14 (.08)	95.65 (.46)	95.20 (.48)	95.65 (.46)	2.16	3	.54
2010 State General	67.36 (.18)	67.10 (1.05)	67.10 (1.05)	67.30 (1.05)	.12	3	.99
2010 State Primary	26.56 (.17)	25.95 (.98)	26.40 (.99)	27.95 (1.00)	2.37	3	.50
Sex (% Female)	NA						
Age (Ave.)	47.7 (.06)	48.0 (.36)	47.6 (.37)	47.3 (.37)	13.14	9	.16
N	68178	2000	2000	2000			

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

Toledo, Ohio: November 5, 2013 Mayoral Elections

<u>Variable</u>	<u>Control Group</u>	<u>Female Eyes Group</u>	<u>Male Eyes Group</u>	<u>U.S. Flag Group</u>	<u>Chi2</u>	<u>Df</u>	<u>p</u>
Turnout in 2012 General	84.99 (.15)	85.45 (.80)	85.04 (.81)	85.47 (.80)	.61	3	.89
Turnout in 2012 Primary	31.77 (.20)	30.23 (1.08)	32.09 (1.10)	30.92 (1.10)	.47	3	.47
Turnout in 2009 Mayoral	53.70 (.22)	53.78 (1.22)	54.40 (1.23)	53.44 (1.22)	.37	3	.95
Sex (% Female)	NA						
Age (Ave.)	49.5 (.07)	49.7 (.43)	49.1 (.41)	49.1 (.43)	7.04	9	.63
N	62635	2000	2000	2000			

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

Midland, Texas: November 5, 2013 Mayoral Election

<u>Variable</u>	<u>Control Group</u>	<u>Female Eyes Group</u>	<u>Male Eyes Group</u>	<u>U.S. Flag Group</u>	<u>Chi2</u>	<u>Df</u>	<u>p</u>
Turnout in 2012 General	67.69 (.42)	68.15 (1.04)	67.60 (1.05)	67.80 (1.05)	.19	3	.98
Turnout in 2012 Primary	19.22 (.36)	20.65 (.91)	20.00 (.89)	20.95 (.91)	4.99	3	.17
Turnout in 2010 Primary	19.43 (.36)	20.00 (.89)	18.65 (.87)	20.50 (.90)	2.54	3	.47
Sex (% Female)	56.7 (.45)	55.7 (1.11)	55.7 (1.11)	56.6 (1.11)	1.22	3	.75
Age (Ave.)	50.9 (.17)	50.9 (.41)	50.9 (.42)	50.7 (.42)	2.73	9	.97
N	12170	2000	2000	2000			

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

El Paso, Texas (June 15, 2013): Mayoral Runoff

<u>Variable</u>	<u>Control</u> <u>Group</u>	<u>Female Eyes</u> <u>Group</u>	<u>Male Eyes</u> <u>Group</u>	<u>U.S. Flag</u> <u>Group</u>	<u>Chi2</u>	<u>Df</u>	<u>p</u>
Turnout in 2012 General	53.11 (.15)	54.05 (.90)	53.4 (.90)	51.60 (.90)	4.01	3	.26
Turnout in 2012 Primary	19.84 (.12)	20.12 (.72)	19.07 (.71)	18.68 (.70)	3.76	3	.29
Turnout in 2010 Primary	15.84 (.11)	16.39 (.67)	15.90 (.66)	15.15 (.65)	1.80	3	.62
Sex (% Female)	54.16 (.15)	54.41 (.90)	53.56 (.90)	53.98 (.90)	.55	3	.91
Age (Ave.)	49.2 (.06)	49.1 (.33)	49.1 (.33)	49.2 (.32)	2.57	9	.98
N	104089	3062	3062	3062			

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

ONLINE APPENDIX B: MULTIVARIATE PROBITS; Dependent Variable Turnout

Variable Treatment	VIRGINIA	MINNEAPOLIS	TOLEDO	MIDLAND	EL PASO
Placebo	-.001 (.030)	.030 (.033)	-.100*** (.037)	.012 (.040)	.039 (.035)
<i>Female Eyes Treatment</i>	.034 (.030)	.038 (.033)	.001 (.037)	.023 (.039)	-.003 (.035)
Male Eyes Treatment	.012 (.030)	-.006 (.033)	.039 (.037)	.071* (.039)	.019 (.035)
Sex	.119*** (.012)	NA	NA	-.027 (.025)	-.049*** (.011)
Age	.062*** (.002)	.046*** (.002)	.041*** (.002)	.025*** (.004)	.021*** (.002)
Age Squared	-.0005** (.00002)	-.0004*** (.00002)	-.0003*** (.00002)	-.0002*** (.00004)	-.0002*** (.00002)
Years Registered	.019*** (.001)	.004*** (.001)	-.006*** (.001)	.007*** (.001)	.010*** (.001)
Single Voter Household	-.330*** (.012)	-.152*** (.011)	-.094*** (.013)	-.077*** (.026)	-.076*** (.012)
General Elec. 11/2012		.945*** (.034)	.506*** (.024)	1.000*** (.043)	.939*** (.017)
General Elec. 11/2010		.742*** (.013)	.381*** (.016)	.446*** (.031)	.470*** (.014)
Primary Elec. 3/12 or 8/12		.653*** (.019)	.450*** (.016)	.711*** (.034)	.714*** (.018)
Primary Elec. 8/10 or 9/11		.588*** (.014)	.218*** (.033)	.449*** (.034)	.492*** (.016)
Primary Elec. 9/2013			1.253*** (.019)		
Muni. Gen. 11/09 or 11/11		.810*** (.015)	.725*** (.015)		
Special Elec. 5/2011			.150** (.065)		
Registered Republican			.170*** (.016)	.266*** (.034)	.078*** (.024)
Registered Democrat			.117*** (.022)	.313*** (.034)	.081*** (.021)
Constant	-2.24*** (.054)	-3.160*** (.057)	-2.647*** (.0655)	-2.835*** (.114)	-3.010*** (.053)
N	49070	74154	68635	18170	113275
LR chi2 (11)	5252.21	29132.62	38067.65	7160.52	38928.86
Prob < chi2	.000	.000	.000	.000	.000
Pseudo R2	.08	.29	.42	.34	.38

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

Works Cited

- Abramowitz, Alan. I. 2010. *The disappearing center: Engaged citizens, polarization, and American democracy*. New Haven, Conn: Yale University Press.
- Bateson, Melissa, Daniel Nettle, and Gilbert Roberts. 2006. "Cues of being watched enhance cooperation in a real-world setting." *Biology Letters* 2: 412-414.
- Barclay, Pat. 2010. "Harnessing the Power of Reputation: Strengths and Limits for Promoting Cooperative Behaviors." *Evolutionary Psychology* 10(5): 868-883.
- Brehm, Sharon S. and Jack W. Brehm. 1981. *Psychological reactance: A theory of freedom and control*. New York, NY: Academic Press.
- Cai, Wei, Xiangqin Huang, Song Wu, and Yu Kou. 2014. "Dishonest behavior is not affected by an image of watching eyes." *Evolution and Human Behavior*.
<http://dx.doi.org/10.1016/j.evolhumbehav.2014.09.007>
- Elazar, Daniel J. 1972. *American Federalism: A View from the States*. New York, NY: Harper & Row.
- Ekstrom, Mathias. 2011. "Do watching eyes affect charitable giving? Evidence from a field experiment." *Experimental Economics* 15: 530-546.
- Ernest-Jones, Max, Daniel Nettle, and Melissa Bateson. 2011. "Effects of Eye Images on everyday cooperative behavior: a field experiment." *Evolution and Human Behavior* 32:172-178.
- Fehr, Ernst, and Frederic Schneider. 2010. "Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity?" *Proceedings of the Royal Society B: Biological Sciences* 277:1315-1323.
- FloridaKeysTV. 2012. "Key West...Close To Perfect." TVad/YouTube video, 0:30,
<http://www.youtube.com/watch?v=makDyycHvmw&feature=fvhl>.
- Foley, Robert. 1995. "The Adaptive Legacy of Human Evolution: A Search for the Environment of Evolutionary Adaptedness." *Evolutionary Anthropology* 4(6): 194-203.
- Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W.W. Norton & Company.
- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence From a Large-scale Field Experiment." *American Political Science Review* 102: 33-48.

- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. 2010. "An Experiment Testing the Relative Effectiveness of Encouraging Voter Participation by Inducing Feelings of Pride or Shame." *Political Behavior* 32 (3): 409–22.
- Gerber, Alan S., Donald P. Green, and David W. Nickerson. 2001. "Testing for Publication Bias in Political Science." *Political Analysis* 9(4): 385-392.
- Gerber, Alan S., Neil Malhotra, Conor M. Dowling and David Doherty. 2010. "Publication Bias in Two Political Behavior Literatures." *American Politics Research* 38(4): 591-613.
- Green, Donald P., and Alan S. Gerber. 2008. *Get Out the Vote! How to Increase Voter Turnout*, 2nd ed. Washington, DC: Brookings Institution.
- Green, Donald P., and Alan S. Gerber. 2010. "Introduction to social pressure and voting: New experimental evidence." *Political Behavior* 32: 331-336.
- Green, Donald P., Mary C. McGrath, and Peter M. Aronow. 2013. "Field Experiments and the Study of Voter Turnout." *Journal of Elections, Public Opinion and Parties* 23(1): 27-48.
- Haley, Kevin J. and Daniel M.T. Fessler. 2005. "Nobody's watching? Subtle cues affect generosity in an anonymous economic game." *Evolution and Human Behavior* 26(3): 245-256.
- Ioannidis, John P.A. 2005. "Why Most Published Research Findings Are False" *PLOS/Medicine*. Vol. 2(8): e124. doi:10.1371/journal.pmed.0020124.
- Issenberg, Sasha. 2012. *The Victory Lab: The Secret Science of Winning Campaigns*. New York: Crown Books.
- Jolij, Jacob, and Tineke de Haan. 2014. "Being watched doesn't make you nicer: no effect of visible and invisible eye primes on pro-social behavior in a masked priming study" unpublished manuscript.
- Kurzban, Robert. 2001. "The social psychophysics of cooperation: nonverbal communication in a public goods game." *Journal of Nonverbal Behavior* 25(4): 241-259.
- Mann, Christopher (2010) "Is There Backlash to Social Pressure? A Large-scale Field Experiment on Voter Mobilization." *Political Behavior*, 32:387-407.
- Matland, Richard E. and Gregg R. Murray (2012) "An Experimental Test of Mobilization Effects in a Latino Community" *Political Research Quarterly*, 65(1): 192 - 205.
- Matland, Richard E. and Gregg R. Murray. 2013. "An Experimental Test for 'Backlash' Against Social Pressure Techniques Used to Mobilize Voters." *American Politics Research* 41(3): 359-386.
- Mayr, Ernst. 2001. *What Evolution Is*. New York: Basic Books.

- McDermott, Rose. 2002. "Experimental Methodology in Political Science." *Political Analysis* 10(4): 325-342.
- Mifune, Nobuhiro, Hirofumi Hashimoto, and Toshio Yamagishi. 2010. "Altruism Towards In-group Members as a Reputation Mechanism." *Evolution and Human Behavior* 31(2): 109-117.
- Miller, Penny M. 1994. *Kentucky Politics & Government: do we stand united?* Lincoln, Nebraska: University of Nebraska Press.
- Murray, Gregg R., and Richard E. Matland. 2014. "Mobilization Effects Using Mail: Social Pressure, Descriptive Norms, and Timing." *Political Research Quarterly* 67(2): 304-319.
- Nettle, Daniel, Zoe Harper, Adam Kidson, Rosie Stone, Ian S. Penton-Voak, Melissa Bateson. 2013. "The watching eyes effect in the Dictator Game: it's not how much you give, it's being seen to give something." *Evolution and Human Behavior* 34(1): 35-40.
- Oda, Ryo, Yuki Niwa, Atsushi Honma, Kai Hiraishi. 2011. "An eye-like painting enhances the expectation of a good reputation." *Evolution and Human Behavior* 32(3): 166-171.
- Panagopoulos, C. (2010). Affect, social pressure and prosocial motivation: Field experimental evidence of the mobilizing effects of pride, shame and publicizing voting behavior. *Political Behavior*, 32, 369–386.
- Panagopoulos, Costas. 2011 "Thank You for Voting: Gratitude Expression and Voter Mobilization." *Journal of Politics* 73(3): 707-717.
- Panagopoulos, Costas. 2014a. "I've Got My Eyes on You: Implicit Social Pressure Cues and Prosocial Behavior." *Political Psychology* 35(1): 23-33.
- Panagopoulos, Costas. 2014b. "Watchful eyes: implicit observability cues and voting." *Evolution and Human Behavior* 35(4): 279-284.
- Penner, Louis A., John F. Dovidio, Jane A. Piliavin, David A. Schroeder. 2005. "Prosocial Behavior: Multilevel Perspectives," *Annual Review of Psychology* 56: 365-392.
- Piazza, Jared, and Jesse M. Bering. 2008. "Concerns about reputation via gossip promotes generous allocations in an economic game." *Evolution and Human Behavior* 29(3): 172-178.
- Powell, Kate L., Gilbert Roberts, and Daniel Nettle. 2012. "Eye Images Increase Charitable Donations: Evidence from an Opportunistic Field Experiment in a Supermarket." *Ethology* 118(11): 1-6.
- Raihani, Nichola J., and Redouan Bshary. 2012. "A positive effect of flowers rather than eye images in a large-scale, cross-cultural dictator game." *Proceedings of the Royal Society B: Biological Sciences* 279: 3556-3564.

- Rigdon, Mary, Keiko Ishii, Motoki Watabe, Shinobu Kitayama. 2009. "Minimal social cues in the dictator game." *Journal of Economic Psychology* 30(3): 358-367.
- Rind, Bruce, and Daniel Benjamin. 1994. "Effects of public image concerns and self-image on compliance." *Journal of Social Psychology* 34(1): 19-25.
- Shariff, Azim F., and Ara Norenzayan. 2007. "God is Watching You: Priming God Concepts Increases Prosocial Behavior in an Anonymous Economic Game." *Psychological Science* 18(9): 803-809.
- Sparks, Adam, and Pat Barclay. 2013. "Eye Images increase generosity, but not for long: The limited effect of a false cue." *Evolution and Human Behavior* 34(4): 317-322.
- Sylwester, Karolina, and Gilbert Roberts. 2013. "Reputation-based partner choice is an effective alternative to indirect reciprocity in solving social dilemmas." *Evolution and Human Behavior* 34(3): 201-206.
- Tooby, John, and Leda Cosmides. 1992. "Psychological Foundations of Culture." In J. H. Barkow, L. Cosmides and J. Tooby, eds., *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press.
- Wicklund, Robert A. 1974. *Freedom and Reactance*. Potomac, MD: Lawrence Erlbaum.

Table 1. Characteristics of Research Sites

<u>Location</u>	<u>Election Date</u>	<u>Political Culture</u>	<u>Election Type</u>	<u>Official Turnout</u>
Rural Virginia, VA	11/5/13	Traditional	Governor	42.0%
Minneapolis, MN	11/5/13	Moralist	Mayor	33.0%
Toledo, OH	11/5/13	Individualist	Mayor	25.0%
Midland, TX	11/5/13	Individualist	Mayor & Referenda	20.2%
El Paso, TX	6/15/13	Individualist/ Undetermined	Mayoral Runoff	15.2%

Table 2. Bivariate Results for Four Treatments across Five Sites

Condition		Turnout			Average Treatment Effects (ATE)			
(1) Study Site	(2) Treatment Condition	(3) Control Turnout	(4) Placebo Turnout	(5) Treatment Turnout	(6) v. Control		(7) v. placebo	
					ATE	p	ATE	p
Rural Virginia	Control (N=43,070)	44.23 (.24)			-	--	--	--
	Placebo (N=2,000)		44.35 (1.11)		.12 (1.14)	.46		
	Female Eyes (N=2,000)			45.35 (1.11)	1.12 (1.14)	.16	1.00 (1.57)	.26
	Male Eyes (N=2,000)			44.30 (1.11)	.07 (1.14)	.48	-.05 (1.57)	.49
Minneapolis, MN	Control (N=68,178)	41.56 (.18)			--	--	--	--
	Placebo (N=2,000)		42.70 (1.11)		.114 (1.12)	.15		
	Female Eyes (N=2000)			42.65 (1.10)	1.09 (1.12)	.17	-.05 (1.56)	.49
	Male Eyes (N=2000)			41.25 (1.11)	-.31 (1.11)	.39	-1.45 (1.56)	.18
Toledo, Ohio	Control (N=62,635)	37.47 (.19)			--	--	--	--
	Placebo (N=2,000)		35.05 (1.07)		-2.42* (1.08)	.02		
	Female Eyes (N=2000)			37.25 (1.08)	-.22 (1.10)	.42	2.20 (1.52)	.08
	Male Eyes (N=2000)			38.10 (1.09)	.63 (1.10)	.29	3.05* (1.52)	.03
Midland, Texas	Control (N=12,170)	25.77 (.40)			--	--	--	--
	Placebo (N=2,000)		26.60 (.99)		.83 (1.06)	.22		
	Female Eyes (N=2000)			26.85 (.99)	1.08 (1.07)	.16	.25 (1.40)	.43
	Male Eyes (N=2000)			27.60 (1.00)	1.83* (1.08)	.05	1.00 (1.41)	.24
El Paso, Texas	Control (N=104,089)	16.49 (.12)			--	--	--	--
	Placebo (N=3,062)		16.39 (.67)		-.10 (.68)	.44		
	Female Eyes (N=3,062)			16.75 (.67)	.26 (.68)	.35	.36 (.95)	.35
	Male Eyes (N=3,062)			16.69 (.67)	.20 (.68)	.39	.29 (.95)	.38

ATE = turnout rate in treatment group – turnout rate in control group (or placebo group).

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

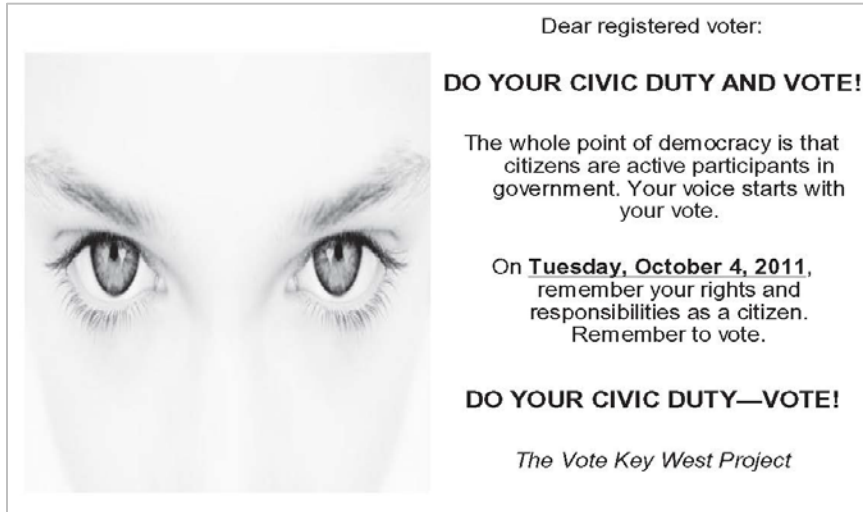
Table 3. Bivariate Results: Testing the Combined Effects of Female and Male Implicit Social Pressure Mailers across Five Sites

Study Site		Turnout	Average Treatment Effects (ATE)			
(1) Study Site	(2) Treatment Condition	(3) Treatment Turnout	(4) v. Control		(5) v. Placebo	
			ATE	p	ATE	p
Rural Virginia	F&M Eyes (N=4,000)	44.83 (.79)	.59 (0.82)	.23	.48 (1.36)	.36
Minnea- polis, MN	F&M Eyes (N=4000)	41.95 (0.78)	.39 (.80)	.31	-.75 (1.35)	.71
Toledo, Ohio	F&M Eyes (N=4000)	37.68 (.77)	.21 (.79)	.40	2.63* (1.31)	.03
Midland, Texas	F&M Eyes (N=4000)	27.23 (.70)	1.46* (.81)	.04	.63 (1.21)	.31
El Paso, Texas	F&M Eyes (N=6,124)	16.72 (.48)	.23 (.49)	.32	.33 (.82)	.35

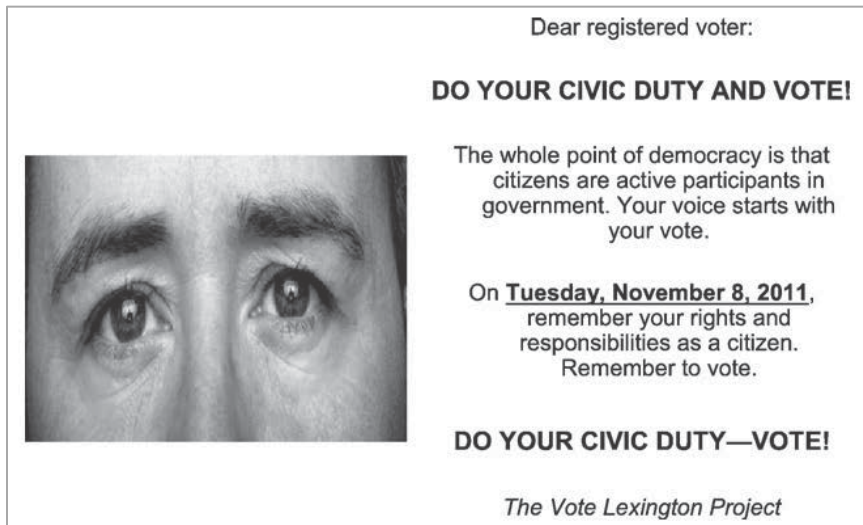
ATE = turnout rate in treatment group – turnout rate in control group (or placebo group).

* p < .05, one-tailed test, ** p < .01, one-tailed test, *** p < .001, one-tailed test

Figure 1. Panagopolous Eyes Treatments/Mailers



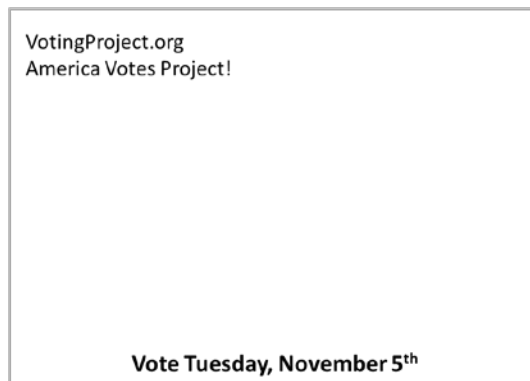
Panagopolous (2014a)



Panagopolous (2014b)

Figure 2: Matland and Murray Eyes Treatments/Mailers

Address Panel



Placebo Treatment

Do your Civic Duty and Vote!



SUPPORT LOCAL DEMOCRACY!

*Dear Registered Voter:
The whole point of democracy is that citizens are active participants in government. Your voice starts with your vote. Remember your rights and responsibilities as a citizen. Remember to vote.*

**On November 5, 2013
Do Your Civic Duty –VOTE!**

Female Eyes Treatment

Do your Civic Duty and Vote!



SUPPORT LOCAL DEMOCRACY!

*Dear Registered Voter:
The whole point of democracy is that citizens are active participants in government. Your voice starts with your vote. Remember your rights and responsibilities as a citizen. Remember to vote.*

**On November 5, 2013
Do Your Civic Duty –VOTE!**

Male Eyes Treatment

Do your Civic Duty and Vote!

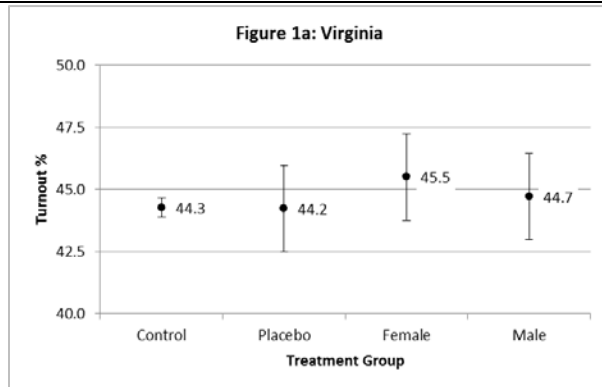


SUPPORT LOCAL DEMOCRACY!

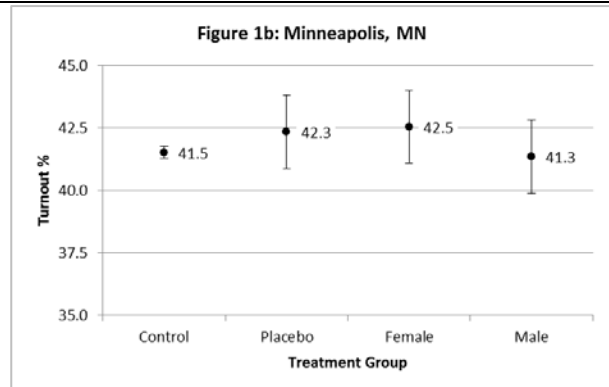
*Dear Registered Voter:
The whole point of democracy is that citizens are active participants in government. Your voice starts with your vote. Remember your rights and responsibilities as a citizen. Remember to vote.*

**On November 5, 2013
Do Your Civic Duty –VOTE!**

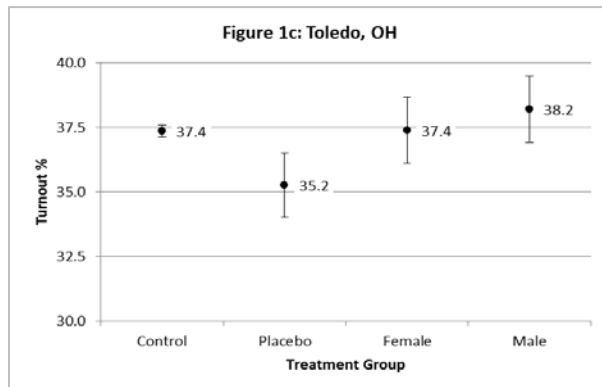
Figure 3. Estimated Turnout with 90% Confidence Intervals from Multivariate Models for Treatment Group Turnout Across Five Research Sites, with Tables of Average Treatment Effects (ATEs)



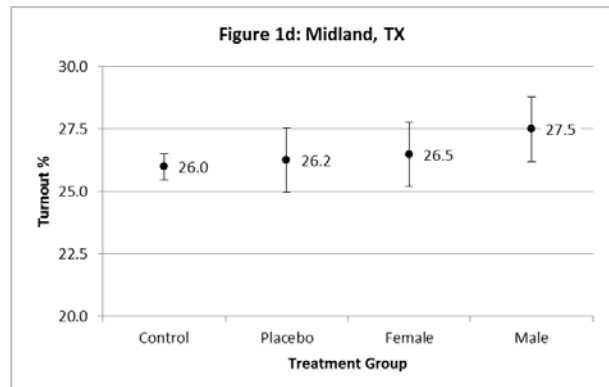
	<u>ATE</u>	<u>Std. Error</u>	<u>P</u>
Placebo (Flag)	-.03	1.08	.49
Female Eyes	1.24	1.08	.13
Male Eyes	.45	1.08	.34



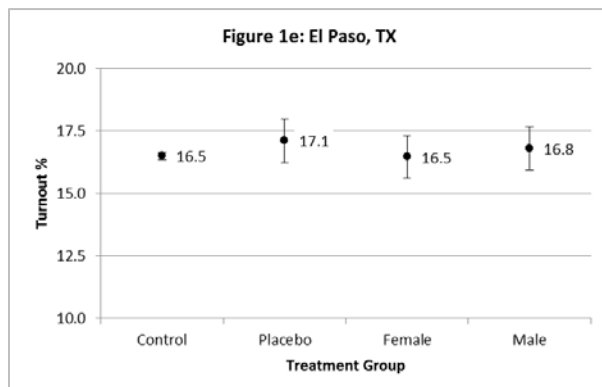
	<u>ATE</u>	<u>Std. Error</u>	<u>P</u>
Placebo (Flag)	.81	.91	.19
Female Eyes	1.03	.90	.13
Male Eyes	-.17	.90	.43



	<u>ATE</u>	<u>Std. Error</u>	<u>P</u>
Placebo (Flag)	-2.11**	.78	.004
Female Eyes	.02	.78	.49
Male Eyes	.85	.79	.15



	<u>ATE</u>	<u>Std. Error</u>	<u>P</u>
Placebo (Flag)	.26	.84	.38
Female Eyes	.49	.83	.28
Male Eyes	1.50*	.85	.04



	<u>ATE</u>	<u>Std. Error</u>	<u>P</u>
Placebo (Flag)	.61	.54	.14
Female Eyes	-.04	.53	.47
Male Eyes	.28	.53	.30

NOTE: P values in tables represent one-tailed tests