

Charitable Giving in Nonprofit Service Associations: Identities, Incentives, and Gender Differences

Heng Qu* and Richard Steinberg**

*Qu (corresponding author): Assistant Professor at Texas A&M University Bush School of Government and Public Service. 4220 TAMU College Station, TX 77843-4220, USA. Email: hqu@uemail.iu.edu.

**Steinberg: Professor of Economics, Philanthropic Studies, and Public Affairs at Indiana University-Purdue University Indianapolis. Address: Cavanaugh Hall 509E, 425 University Blvd., Indianapolis, IN 46202. Email: rsteinbe@iupui.edu.

Authors' Note

We thank the Science of Philanthropy Initiative and its funder, the John Templeton Foundation, for funding and helpful suggestions. We have also received helpful remarks from Sara Konrath, Mark O. Wilhelm, Daniel Jones, Patrick Rooney, Barbara J. Duffy, Ruth K. Hansen, Kevin Snapp, Xiaonan Kou, and Kathryn S. Steinberg. We also thank all the referees for their suggestions that have made this a better paper.

Keywords

nonprofit service associations, charitable giving, social identity, incentives, gender differences

This is the author's manuscript of the article published in final edited form as:

Qu, H., & Steinberg, R. (2017). Charitable Giving in Nonprofit Service Associations: Identities, Incentives, and Gender Differences. *Nonprofit and Voluntary Sector Quarterly*, 46(5), 984–1005. <https://doi.org/10.1177/0899764017703709>

Charitable Giving in Nonprofit Service Associations: Identities, Incentives, and Gender Differences

Abstract

Nonprofit service associations, like the Lions Clubs, Rotaries, and Kiwanis, provide collective goods. Membership in a service association involves two essential elements: members' shared interest in the club's charitable service; and private benefits stemming from social interactions with other members, such as networking, fellowship, and fun. We report results from a laboratory experiment designed to test the effect of membership and priming on charitable giving. The two experimental conditions activate chains of associative memory linked to the service or socializing aspects of membership. We find that female subjects give significantly smaller donations after receiving the socializing stimulus. Male subjects are less sensitive to our experimental conditions, giving slightly more in the socializing condition, but the differences are not statistically significant. We discuss three mechanisms that may explain our results: social identity theory, reputation and image motivations, and quality inference.

Keywords: nonprofit service associations, charitable giving, social identity, incentives, gender differences, lab experiments, priming

Introduction

Nonprofit service associations (or service clubs), such as Lions Clubs, Rotary, and Kiwanis, provide a space where members get together to satisfy their self-interest and social needs, as well as to solve their community concerns (Charles, 1993). Giving circles, where groups of individuals donate their time and money to a pooled fund and collectively allocate their funds to other charities and community projects, are similar (Eikenberry, 2006). This paper is concerned with the effectiveness of service clubs and other similar types of organizations in fostering members' generosity and support for collective goods. A common belief is that the higher the level of member commitment, the more successfully a service club can pursue its mission. We argue that this analysis is too simple, that the way the club builds commitment matters. Our pilot study found that one approach to building commitment actually reduces donations. Hence, this study distinguishes two broad categories of benefits to members. First, "service benefits" – members benefit from their association with an organization that helps nonmembers and from their personal role in helping the club to do so. Members may want that association to enhance their charitable reputation or meet psychological needs, but the intended effect is the provision of a charitable service to nonmembers. Second, "socializing benefits" – members benefit from their interactions with other members, such as gaining networking opportunities, making friends, or having fun.

The dual goals of providing service and socializing benefits may have conflicting effects on members' charitable behaviors. On the one hand, service clubs assemble like-minded individuals to support the provision of collective goods. On the other, the social activities of the service club may de-emphasize the charitable mission, resulting in reduced donations.

The most obvious way to study the effects of service club membership is to compare the charitable donations of members and nonmembers. However, this approach suffers from self-selection bias – those who choose to join are different in unobserved ways from those who do not join. Are differences in generosity due to membership *per se* or the unmeasured factors that make someone become a member? A more sophisticated approach would compare donations before and after joining, but this suffers from the same problem. Is increased generosity due to joining, or did some unobserved variable lead to both joining and giving? In contrast, experiments use random assignment of subjects to treatments, so that differences in unobserved variables average out and estimated effects are plausibly causal.

Hence, we conduct a laboratory experiment designed to distinguish between two different ways of building member commitment: focusing members' attention on the service benefits stemming from helping others, or the socializing benefits stemming from interacting with other members. Our design distinguishes the control group (nonmembers) from the treatment group (members) under two conditions (service and socializing). We find that emphasizing the socializing benefits reduces donations, whereas emphasizing service benefits has a small, statistically insignificant, positive effect when compared with the control group. Although the difference in donations across conditions was statistically significant, this difference is due to the behavior of female subjects. Male subjects have small, statistically insignificant treatment responses. The gender difference in treatment effects was unexpected, and we explore possible explanations in our discussion.

In the next section, we discuss theory and present our research hypotheses. Then, we describe our experiment in detail and report the results. A discussion concludes the paper.

Theory

In this section, we first review the problem of free riding when contributing to collective goods and discuss the ways in which service clubs can reduce this problem. Then we discuss psychological priming, the priming conditions used in our experiment, and the ways these stimuli might affect behavior.

Collective Goods. Collective goods are nonrival in consumption – that is, it costs nothing to let an additional person enjoy them once they are provided (Samuelson, 1954). Like many nonprofits, service clubs provide collective goods and such provision is enshrined in their mission statements. For example, Lions Club International lists its mission as “To empower volunteers to serve their communities, meet humanitarian needs, encourage peace and promote international understanding through Lions clubs.” (Lions Clubs International, 2015). Any additional member or nonmember can benefit from the fact that Lions provide these services to others, and the addition of more consumers of this fact does not diminish the consumption of incumbents. The service club in our experiment provides access to higher education, a service collectively enjoyed by those who value others’ access to education.

A well-known problem with voluntary support for collective goods is free riding, enjoying service provision without personally contributing. This leads to under-provision of collective goods relative to the social optimum. Andreoni’s (1988, 1989) impure altruism model provides a framework that predicts the observed distribution of charitable giving and free riding. In his model, potential donors care about both the level of collective good provision (labeled “altruism”) and their personal contribution to the collective good (“warm glow”). Variation in the relative importance of these two motivations explains why some people free ride, others do not, and still others take an easy ride (less extreme free riding).

Service clubs assemble like-minded individuals to support the provision of collective goods. This reduces free riding through self-selection, repeated interactions with other members, and community enforcement (Kandori, 1992). In addition, service clubs can employ mechanisms known or thought to reduce free riding, like devoting a portion of mandatory dues to funding collective goods or bundling warm glow with entertainment in special-event fundraisers.

Priming. Priming occurs when exposure to a stimulus influences the response to another stimulus (Samson and Miles, 2016). Molden (2014) notes the lack of a precise universally accepted definition of social priming. Most users of the term look at how priming activates various social representations (of, for example, traits, stereotypes, or goals) and consider how that activation affects subsequent behavior. The operational definition of priming has expanded to include (p. 4) “how cues that activate ... the recall of specific social contexts and events alter current preferences and choices.” Morewedge and Kahneman (2010, p. 435) summarize evidence that priming works through stimulation of “associative memory,” the “network of long-term memory for semantic information, emotions and goals that is governed by the spread of activation, as determined by the strengths of interconnecting weights.” Further,

[T]he pattern of automatic activation in memory tends to produce a comprehensive and internally consistent interpretation of the present situations, which is causally embedded in the context of the recent past, and incorporates appropriate emotions and preparedness for likely future events and for future actions (p. 436).

In our experiment, we prime subjects using a task designed to activate chains of association linked to the socializing or service benefits of membership in a service club.

Effects of priming the socializing benefits of membership. Three theories predict that subjects receiving the socializing condition will decrease, and one that they will increase

donations. First, reference to the social activities of the club may activate impressions that the club is not serious in providing collective goods or services, or members' donations may not be used to accomplish charitable purposes. This would presumably reduce their donations.

Second, the condition alters the reputational and or image benefits of donating. One of the main reasons why people give is to maintain their self-image and reputation as good people. Extrinsic rewards, such as money, prestige, or front row seats, for prosocial acts create doubt regarding one's true motivation, crowding out the image and reputational motivations for giving. Social psychologists Lepper et al. (1973) call this the overjustification effect and economists Bénabou and Tirole (2006) model the problem mathematically. Several experiments confirm the empirical relevance of this phenomenon. For example, Ariely et al. (2008) told subjects in a laboratory experiment that donations would be made to a charity in proportion to the distance they pedaled on a stationary bicycle. Their experimental conditions varied whether or not subjects were paid for their time on the bicycle and whether other subjects observed them cycling. Those receiving monetary payments covered less ground when observed than did those in the two unpaid conditions. However, subjects that were not observed and received monetary payments for effort covered more ground than did those in the two unpaid conditions. Field experiments by Karlen and McConnell (2014) found that these effects are due to reputational crowding out, rather than an altruistic desire to serve as a role model and thereby increase the gifts of others. In our experiment, the socializing condition could affect the reputational benefits of donating. Subjects may view the socializing condition, when the socializing benefits of club membership are emphasized, as an extrinsic reward for members that reduces the signaling value of giving. If so, donations would fall.

The closely related theory of motivational crowding out also applies. Deci (1971) labeled the feeling of satisfaction that one gets from acting as “intrinsic,” and provided evidence that intrinsic motivation is crowded out by extrinsic rewards. Applied here, socializing benefits (extrinsic) may crowd out the warm glow from giving, resulting in reduced donations.

Third, social identity theory postulates that individuals contain multiple internal identities that lead to distinct behaviors when activated (Tajfel and Turner, 1979). Akerlof and Kranton (2000) maintain that each internal identity comes with a set of socially prescribed norms of behavior. For example, cultural and social norms, although contested, suggest that Asians should be good at math and women not so good. Both identities are present in Asian women, and either identity could be activated by contextual factors. Shih, Pittinsky, and Ambady (1999) confirmed this story in an experiment, and found that Asian-American women score higher on a math test when primed to identify themselves as Asian than when primed to identify as women. In the present context, subjects may enter the experiment with a “partier” identity that has norms of making friends and having fun rather than making charitable donations. The socializing condition could activate this partier identity.

The fourth theory is that the socializing condition increases generalized commitment to the club. Olson (1965) argues that clubs gain the commitment of members by providing selective material and solidary incentives. Knoke (1988) finds that an index of social incentives has a significant positive effect on an index of internal participation in professional associations. This commitment would plausibly carry over into support for the club’s charitable mission, resulting in increased donations. In light of these competing sign predictions, we offer the following null hypothesis:

H₀¹: Average donations of nonmembers will be the same as average donations of members in the socializing condition.

Effects of priming the service benefits of membership. Two theories suggest the service condition will increase donations. First, according to social identity theory, this condition may activate a “giver” identity, with a prescribed level of giving. Second, as with the socializing condition, this condition may increase generalized member commitment to the club, resulting in increased support. Knoke (1988) also found that an index that includes the goal of attempting to change nonmembers’ lives also explained internal participation. Putting it all together, we have the following additional null hypotheses:

H₀²: Average donations of nonmembers will be the same as average donations of members in the service condition.

H₀³: Average donations of members in the service condition will be the same as average donations of members in the socializing condition.

The Experiment

Sequencing and Summary. First, subjects chose an organization they would like to be associated with from a list of two options. Second, subjects participated in a practice round, using a neutrally framed decision making task. Third, subjects in the treatment groups were welcomed to membership in the fictitious Education Service Club (ESC) and completed one of the two priming tasks. Subjects in the control group skipped this step. Fourth, we read a solicitation message to all subjects and asked them to allocate their tokens. Fifth, subjects designated a real-world charity to receive their donations and filled out questionnaires. We compensated subjects privately before they left. Table 1 summarizes our procedures.

[Table 1]

Design and Assignment. We used a between-subject design, consisting of one control (no club membership) and treatments with club membership under two conditions (service vs. socializing). 113 students at a Midwestern university participated in our experiment, receiving an average compensation of \$7.60.¹ Upon arrival, subjects were asked to choose an organization they would “like to get involved with” from a list of two organizations, with a vague, one-sentence mission statement for each. One of the organizations was always the fictitious Education Service Club (ESC) (or Education Service Association (ESA) in the control group). The other was a fictitious organization, either the Society of Accounting Standards (SAS) or the Society of Warehousing Standards (SWS), randomly applied to sessions, which we intentionally described in a way that led very few subjects to make this choice. Those subjects picking ESC (ESA) continued with the experiment. The 12 subjects who chose SAS/SWS were taken to another room and completed the questionnaires, but did not participate in the later stages of the experiment. We did this because the treatment specifically referred to the ESC (ESA), with some subjects assigned to membership and others to (nonmember) donor status. Because subjects stated that they would like to get involved with this organization, they are less likely to feel forced into membership or donor status by the random assignment procedure. An additional 8 subjects were dropped for failure to complete the priming task as directed.² Our final sample contained 93 subjects, with 30 assigned to the control group, 32 to treatment under the service condition, and 31 to treatment under the socializing condition.

We assigned subjects to 18 single-treatment sessions based on their availability, and each participated in the experiment only once.³ Subjects did not know in advance which treatment they would receive, so there is no self-selection into treatment. Moreover, each treatment had multiple sessions, on different days and times, so that selection based on unobservable attributes

correlated with time should not be a problem. Subjects were instructed to finish their tasks independently without interacting with others during the experiment – they sat behind black paper shields without seeing others’ activities and were not allowed to talk with others.

Membership Manipulation. Following the practice round, subjects in the treatment groups were told: “You have chosen the Education Service Club, and are now being granted the membership in the Education Service Club.” The experimenter then read a new member welcome letter orally while subjects read the printed version. Those in the control group heard nothing about membership. They were told instead: “You have chosen the Education Service Association and here is more information about it.”

Priming Task. The welcome letter was also used for priming. The letter also included basic information about the ESC’s mission, service projects and social activities, and membership obligations. Immediately after reading the welcome letter, subjects in the service condition were asked to highlight in the letter “at least 3 phrases that are directly related to the ESC’s *charitable mission*,” and subjects in the socializing condition were asked to highlight “at least 3 phrases that are directly related to the ESC’s *social activities*.”

Four subjects completed the priming task that was not assigned to them (highlighting only the mission-related phrases) instead of the task that was assigned (highlighting the social-related activities). Our baseline results reassign these subjects to the condition they experienced (service). We check whether results are sensitive to this reassignment against the alternative of dropping the subjects (resulting in a sample size of 89) and report the differences in the results section.

Decision Task. All subjects were endowed with 600 tokens and received a short and neutral fundraising message from the ESC (or ESA in the control group). For the control group,

the fundraising message, which was only four sentences long, introduced the charity's mission and its scholarship fund, and then requested donations by asking "please make a donation to the ESA Scholarship Fund." The message also stated "100% of your donation will benefit students because an anonymous donor is covering all of the administration and fundraising costs associated with this campaign" in an effort to remove heterogeneity resulting from subjects making diverse assumptions about these matters. The message for the treatment groups was identical, except that the organization was called ESC and the charity's mission statement was omitted (as it was presented already in the new member welcome letter). The message was read aloud to subjects by the experimenters.

Next, we asked subjects to complete the allocation task, that is, to divide their endowment between a personal account and the scholarship fund. They were told: "At the end of each session, tokens are converted to real money at the rate of 100 tokens to \$1. Your token donations will become real donations to a scholarship fund you will choose from a list of real-world charities."⁴ The money that a subject kept for herself was added to her \$5 show-up fee to determine her monetary compensation. We presented subjects with multiple examples and the opportunity to ask questions in an effort to minimize misunderstandings and miscalculations. Then subjects were told to "place the tokens you wish to keep for yourself in the envelope labeled 'Personal Account,' and place the tokens you wish to donate in the envelope labeled 'Charity.'" To preserve anonymity, subjects sealed the envelopes and handed them to an assistant who relayed them to a second assistant in another room for processing.

Questionnaires. Subjects completed questionnaires post-treatment. The first was a psychometric scale designed to check whether those in the treatment group identified as ESC members, adopted from Ellemers, Spears, and Doosje (1997).⁵ The second contained 14

questions used to construct the Empathic Concern (EC) and Perspective Taking (PT) scales (Davis, 1983). The EC and PT scales respectively allow us to control for individual differences in the level of dispositional empathy with unfortunate others and tendency to spontaneously adopt others' psychological point of view.

Then, subjects completed the Unlikely Virtues (UV) scale (Patrick et al., 2002), which allowed us to control for social-desirability bias. Although anonymity was guaranteed and explained to subjects, this scale serves as an additional check for subjects who do not understand or trust the anonymity procedure. The UV scale is the sum of "true" responses to 14 true-false items that describe socially desirable behaviors that are rarely true, such as "My opinions are always completely reasonable," or "I have never felt that I was better than someone else."

Finally, a socio-demographic questionnaire asks about subjects' basic socio-demographic characteristics, including gender, age, ethnicity and race, educational level, academic major, marital and parental status, and frequency of religious activities. In addition, subjects were asked if they had made donations to any charities before, and if they have/had membership in any public-service clubs. Table 2 provides summary statistics for those variables employed here.

[Table 2]

Results

We analyze our results descriptively, through parametric and nonparametric tests, as well as three kinds of regressions. The pattern is the same throughout. Donations are lowest in the socializing condition and highest in the service condition. The difference between giving in the control and service condition was small and statistically insignificant. These results are consistent with our hypotheses, but the gender difference was unexpected: all the large and/or

statistically significant treatment effects were for females, not males. Table 3 reports unconditional treatment effects, and table 4 reports regression results.

[Table 3]

Panel A of table 3 best illustrates this story. Average token donations were 353.33 out of 600 in the control group, 365.63 in the service condition, and 275.81 in the socializing condition. The same is true of median donations (300, 400, and 200 respectively) and the percentage of subjects who made the maximum donation of 600 tokens (27%, 38%, and 19% respectively). Panel C of table 3 reports nonparametric Wilcoxon–Mann–Whitney tests to determine whether the distributions of donations in each pair of treatments are consistent with sampling from a single population.⁶ We reject the equal distributions null hypothesis at the 10% level. However, the differences between service and control and between socializing and control are not statistically significant. Results from independent sample t-tests, not displayed, are similar.

For robustness, we compare these results with alternative samples. Recall that four subjects experienced the service condition although they were told to highlight phrases related to socializing, and our baseline sample categorizes these subjects as in the service group. If we instead leave them in the socializing group, the difference between donations under the two conditions becomes statistically insignificant, although donations remain higher under the service condition (313.64 vs 297.14). The service treatment has a lower mean (313.64) than the control group (353.33), but this difference is not statistically significant. If, instead, we simply drop these subjects from our sample, the service and socializing treatment means are 351.79 and 275.81 (the difference is not significant, $p = 0.18$). The control mean is not significantly different from the service mean ($p = 0.98$), although it is nearly significantly higher than the socializing mean ($p = 0.14$). Nonparametric tests (the Mann-Whitney rank-sum test and Kruskal-Wallis

equality-of-populations rank test) provide similar patterns. We take it as a good sign that the effects become significant when we reassign subjects, because this reassignment was made on *a priori* rather than statistical grounds.

Whereas statistical significance reveals whether results are likely to be due to chance, effect size helps understand the magnitude of the difference between groups. Panel D of table 3 reports the effect size (Cohen's *d*) of the differences between conditions. Cohen (1988) interprets the effect size as small ($d = 0.2$), medium ($d = 0.5$), and large ($d \geq 0.8$). The Cohen's *d* effect size suggests a small practical difference between service and socializing ($d = 0.426$) and between socializing and control ($d = 0.386$). The practical difference between service and control was miniscule ($d = 0.059$).

Panel B of table 3 reports results for gender subsamples. Because these groups are smaller (there are 59 females and 32 males), the statistical power of hypothesis tests is reduced. It takes a bigger estimated effect size to reject the null hypotheses, particularly in the male subsample. Women in the control group donated an average of 372.73 tokens (median = 300). Average donations increased to 423.68 (median = 500) in the service condition and decreased to 268.42 (median=200) in the socializing condition. The difference in mean giving between the service and socializing conditions is significant at the 5% level, and the difference between socializing and control is marginally significant at the 10% level. Again, if we vary the treatment of the four problematic subjects to determine the robustness of these results, we find an increase in standard errors. When we drop the subjects from the sample, the difference between service and socializing becomes only marginally significant, as does the difference between socializing and control. According to Panel D of table 3, for women, there is a medium-sized effect for the differences between service and socializing ($d = 0.726$) and between socializing and control ($d =$

0.506). The effect size is small for the difference between service and control ($d = 0.255$). The pattern for men was starkly different, with the largest mean donations for the socializing treatment (309.09), followed by the control group (300). Those in the service treatment donated the lowest amount (280.77). There are no statistically significant or practical differences between the groups.

Although random assignment obviates the need for covariates, we estimated regressions to control for sampling variations that lead to departures from “all else held equal” and report results in Table 4. We used Tobit to account for the censoring of the dependent variable at 0 and 600 tokens. The baseline specification includes treatment dummies for socializing and control (service was the omitted category), a female dummy, and empathy measured as the sum of the empathic concern and perspective taking scores.⁷ Columns 1 and 2 of Table 4 report the raw coefficients and unconditional average marginal effects in the baseline specification. Compared with the service treatment, the socializing treatment has a negative average marginal effect on donations (-88.5 tokens), which is statistically significant at the .10 level. This suggests that emphasizing personal benefits leads to lower donations than stressing charitable service. The difference in donations between the control and service treatments was not significant. The sum of empathic concern and perspective taking scores has a small positive effect: a one standard deviation increase in the combined score increases tokens donated by 37.90 (significant at the .10 level). This indicates that people with higher level of empathic concern and perspective taking scores tend to be more generous. The gender effect is not statistically significant in the baseline model.

Next, we allow treatment effects to vary with gender by adding interaction terms, reporting results in columns 3 and 4 of Table 4. A likelihood ratio test does not reject the nested

baseline model, but results of the less parsimonious specification are sufficiently interesting to warrant their display. The socializing condition still has a negative average marginal effect (-97.45 tokens) and is now significantly different from the service condition at the .05 level. The difference between the control and service condition remains statistically insignificant. Although the total effect of Female is not significant, the partial effect of gender being female is large and positive (201.32 tokens) and significant at the .10 level, whereas the interaction effect between Female and socializing is large and negative (-266.6 tokens) and significant at the .10 level. We can interpret this as saying that females are more generous than males in the service and control groups, but have such a large negative reaction to the socializing condition that they become less generous than corresponding males. The sum of the empathic concern and perspective taking scores is not significant in this specification, and when entered separately, neither empathy variable is significant.

[Table 4]

We also conducted corresponding OLS and median regressions as robustness checks and find similar patterns. The socializing condition was negative and significant in all three estimations. However, the point estimate of the average effect of the socializing treatment was smallest for the tobit estimator (-97.45), larger for OLS (-162.87) and largest in median regression (-195). The total effect of gender is significant for OLS, but not the other two.

Although there are no significant differences in empathic concern and perspective taking scores across conditions, we find that men ($\mu = 38.16$) and women ($\mu = 43.08$) in our sample have significantly different combined scores ($p = 0.0028$). This leads us to examine donations separately for the female and male subsamples respectively (Table 5). For women, the socializing treatment, compared with the service treatment, has a large negative average marginal

effect on donations (-157.44 tokens), which is statistically significant at the .05 level. There is still no significant difference between the control and service treatment. The sum of the empathic concern and perspective taking scores is not significant. For men, the treatment effects are not significant. These results are robust to dropping the four subjects who experienced the service condition when assigned to socializing. We performed an additional robustness test, adding dummy variables for the subject's major, as females in our sample are more likely than males to major in the helping fields of philanthropic studies, social work, and nursing and less likely than males to major in economics, business, and engineering. Results, not displayed here, show the same pattern. The socializing condition reduced female giving by an average of 140.82 tokens, significant at the 0.05 level. The conditions do not substantively or significantly affect males. Finally, against the possibility that because the assessment of empathic concern/perspective taking was post treatment, the conditions affected its value, we ran a tobit specification without this variable.⁸ There was no noticeable effect on the size and significance levels of the average treatment effects.

[Table 5]

Discussion

Our first null hypothesis, that donations by service club members in the socializing condition equal donations by control group members, is not rejected for the full sample ($p=0.13$) or for men but is rejected for women ($p=0.08$). Women give less under the socializing condition. The regression results are similar, but less informative – the difference is only nearly significant for women ($p=0.13$), although when we drop the empathic concern and perspective taking scores from the list of control variables, significance is restored.

Our second hypothesis, that donations in the service condition are the same as in the control, is not rejected by any of the Wilcoxon-Mann-Whitney tests or regressions even when the men and women are analyzed separately.

Our third hypothesis, that donations in the service condition are the same as in the socializing condition, is rejected at the 10% level for the full sample, at the 5% level for women, and is not rejected for men. Regressions obtain the same pattern of significance levels.

Collectively, these results support the conclusion that the socializing condition has a negative impact on giving, but only for women. The service condition has no statistically significant effects relative to the control group, but for both the sample as a whole and for women, donations are significantly higher under the service than the socializing condition. In this section, we first discuss the gender-specificity of our findings, then possible explanations for our non-finding on service vs. control along with a discussion of whether membership itself increases generosity. Finally, we discuss practical implications, highlighting the need for further research.

Gender Effects. The finding that all treatment effects are restricted to females in our sample was surprising and intriguing. Taken at face value, it has the interesting implication that the service club model, combining social and service benefits, works less well than it used to when service club membership was restricted to males. We would not go so far without additional experiments specifically designed to uncover gender interactions. Among other problems, we had only 32 males in our present sample, leading to larger standard errors for estimated effect size.

Studies of the effect of gender on giving obtain contradictory results, depending on the charitable cause and numerous other details. In our context, providing educational aid to the

needy, we are not surprised that male subjects generally gave less than female subjects did (although men gave slightly larger average gifts than women within the socializing condition). This could be due to the observed gender difference in empathic concern and perspective taking. Willer et al. (2015) studied men's and women's intent to give and intent to volunteer for a fictional poverty relief agency, arguably similar to our ESC/ESA in its goal to help the needy. Men gave significantly less than women did, but the difference shrank and lost statistical significance after controlling for dispositional empathy.

Our problem is to explain the difference in response to conditions, not the level of giving. Noting that results from some studies of gender-linked treatment effects contradict results from other studies, Croson and Gneezy (2009, p. 16) explain:

We believe that the cause of these conflicting results is that women are more sensitive to cues in the experimental context than are men. ... Small differences in experimental design and implementation will thus have larger impacts on female participants than on male participants.

Two studies suggest that there are gender differences in how extrinsic incentives affect the self-image motivation for giving. Mellström and Johannesson (2008) found that women, but not men, reacted negatively to payments for blood in their experiment. Willer et al. (2015) found that men and women reacted differently to a condition that framed the poverty problem as harmful to all Americans. Men intended to give or volunteer more in this "aligned self-interest" frame (significantly so for giving), while women intended to give or volunteer less (significantly so for volunteering). They speculate that women, and those with higher levels of dispositional empathy, value acts that reinforce their self-image as altruistic people. Reminding these people of their self-interest makes it more difficult to attribute charitable acts to altruism.

Service vs. Control. The second hypothesis is really a joint hypothesis. The control and service condition groups differ in two ways: the latter is welcomed into membership, and performed the service-priming task. We have discussed why the priming task might affect behavior, but membership status can have an independent effect. Subjects who are members presumably feel more responsible and committed to the success of the organization.

When nonmembers are asked to give (our control group), they learn of the organization's mission from the solicitation message they receive. This may activate the same cognitive networks of associations as the service condition. If so, it is not surprising that we found no difference in average donations across the two groups. In terms of social identity theory, our results suggest that there are not two distinct identities – “philanthropist” and “member philanthropist” in play.

Taken at face value, our result suggests that membership organizations are no better at securing donations than nonmember charities – at worst, the membership structure hurts donations (if the negative effect of socializing is present), at best it does not help (using service stimuli). However, our experimental setting removes some of the factors that may make service stimuli more effective for members than nonmembers. In our experiment, members interact through participation in a collective goods treatment only once, and there are no dues or other mechanisms for reducing free riding. Self-selection also matters in ways not captured by our protocol. Clark and Wilson (1961) discuss three broad categories of incentives used to increase donations – material, solidary, and purposive. Material incentives are tangible benefits with a monetary value, like savings on car rentals. Solidary incentives are intangible rewards “derived from the act of associating” (p.134), including socializing, having fun, providing a sense of belonging, and sometimes obtaining elite status. Purposive incentives are intangible rewards

associated with the goals and values of organizations. Although service clubs offer all the three types of incentives, Clark and Wilson argue that they primarily rely on solidary incentives. This attracts members more responsive to solidary than purposive and material benefits. In contrast, nonprofit organizations without members rely more on purposive incentives, attracting donors who are more motivated by service. Hence, membership has an effect mediated through selection of those more responsive to solidary benefits.

It would be useful to separate the effect of membership *per se* from the effect of the service prime. This would allow the analyst to answer the question “Which organizational form collects more donations: contributor- or membership-based? A purposive or a solidary group?” We could not do so in our experiment because of loss of experimental control. The membership letter contains the information used in both priming tasks. If we added a second control group, a possible that some members would pay attention to the socializing stimuli and some to the service stimuli contained in the letter. Then unidentifiable treatment effects would confound results from what was intended as a control group.

Practical Implications. At face value, our results suggest that service clubs and giving circles should be careful about marketing the social aspects of membership. However, before we would want to make claims about practice, we would want to know the mechanisms behind our effects. Suppose, for example, that a) socializing increases member commitment to the cause, and hence donations but b) socializing also sends the message that the club is not serious about service, reducing donations and c) the second effect is bigger than the first and d) neither the reputation and image motivations nor social identities affect donations. Then we would expect that donations are lower in the socializing condition. However, real-world service clubs can emphasize socializing benefits and prove that they are serious about providing charitable services

in other ways, through their actions and marketing. They would overcome the negative effect of socializing, allowing the positive effect to dominate. The implication is different if social identity or reputation/image/motivational-crowdout concerns were behind the negative effects of socializing. Real-world service clubs have no obvious strategies to overcome these effects.

Unfortunately, our experimental design does not allow us to test the relative importance of these (nonrival) mechanisms towards explaining our findings on the first and third hypotheses. In order to assess the importance of quality signaling, future researchers could include questions about the perceived effectiveness of the service effort, delivered post-treatment. That would enable them to tell whether the socializing condition affected beliefs about efficacy. To assess the importance of social identity mechanisms, future researchers would need to develop new psychometric scales designed to measure the relative salience of alternative hypothesized identities. To assess the importance of reputation and image effects, researchers would need to determine whether the social benefits of membership are perceived as intrinsic, extrinsic or altogether different from both. A better approach would be to develop a psychometric scale designed to assess the salience of these benefits. Many scales exist that get at particular benefits in particular contexts (e.g, Ko et al., 2014), but we have not located a suitable existing scale for the service-club setting.

Conclusion

Service clubs provide an alternative approach to supporting charitable causes, combining mission orientation with social interaction. In this way, service clubs have something in common with giving circles, special events fundraising, service fraternities, and some religious establishments. In this setting, we present results from a lab experiment that finds that when members are primed to consider the socializing benefits of membership, they give less than those

primed to consider the service benefits and also less than nonmembers asked to make a contribution. These effects are only present for female members of the sample. Our findings suggest that the decisionmaking context matters, although we are unable to determine the exact mechanism behind observed effects without further research. Future research would also need to look at how the treatment of current members affects the number and characteristics of those wishing to join the service club. Ultimately, this program of research can determine ways to make a service club more effective and answer the question of when an approach that combines social interaction with a request for funds is more effective than direct solicitation of nonmembers.

Appendix. Experimental Instructions*

[Treatment Groups]

About the Education Service Club (ESC). You have chosen the Education Service Club, and are now being granted the membership in the Education Service Club. In your Instructions envelope, there is a letter from the Education Service Club. Let us read the letter together.

In the letter, there is a section subtitled “Projects and Activities” (the third and fourth paragraphs). It provides detailed information about both the ESC’s service projects and social activities.

[Service Condition] In this section, please find and **highlight at least 3 phrases** that are directly related to the ESC’s **charitable mission**.

[Socializing Condition] In this section, please find and **highlight at least 3 phrases** that are directly related to the ESC’s **social activities**.

[Control Group]

About the Education Service Association (ESA). You have chosen the Education Service Association and here is more information about it. The Education Service Association is a charity whose mission is to enhance the educational experiences of students from elementary school to college.

[All Groups]

About the ESC (ESA) Scholarship Fund. It provides support for minority, disabled, as well as first-generation college students. Please make a donation to the ESC (ESA) scholarship fund.

Running Head: CHARITABLE GIVING IN NONPROFIT SERVICE ASSOCIATIONS

100% of your donation will benefit students because an anonymous donor is covering all of the administration and fundraising costs associated with this campaign.

The Allocation Problem. As in the practice round, you have been given 600 tokens to use in place of money. You can keep any share of these tokens for yourself or donate any share of the tokens to the ESC (ESA) Scholarship Fund. At the end of the study, tokens will be converted to real money at the rate of 100 tokens to \$1. Your token donations will become real donations to a scholarship fund you will choose from a list of real-world charities.

For example, if you keep 350 tokens for yourself and donate 250 tokens to the ESC (ESA) Scholarship Fund in this study, you will receive \$3.50 in addition to your \$5 show-up fee, and a real scholarship fund will receive \$2.50. More examples are specified in the following table:

	Tokens to Yourself	Your Additional Cash Payment	Tokens to the ESC (ESA) Scholarship Fund	Donations to a Real-World Charity
Example 1	0	0	600	\$6
Example 2	100	\$1	500	\$5
Example 3	200	\$2	400	\$4
Example 4	300	\$3	300	\$3
Example 5	400	\$4	200	\$2
Example 6	500	\$5	100	\$1
Example 7	550	\$5.5	50	\$0.5
Example 8	600	\$6	0	0

These are only examples. You may allocate any amount of the tokens between yourself and the ESC (ESA) Scholarship Fund. The only restriction is that the sum of the allocation to yourself plus the allocation to the ESC (ESA) Scholarship Fund must equal to 600 tokens.

Decision Making.

[*Treatments Groups*] Now decide, as an ESC member, how many tokens you wish to keep for yourself, and how many tokens you want to donate to the ESC (ESA) Scholarship Fund.

[*Control group*] Now decide how many tokens you wish to keep for yourself, and how many tokens you want to donate to the ESA Scholarship Fund.

* The full experimental materials are available upon request.

Notes

1. We obtained IRB approval before running our experiments.
2. Three subjects assigned to treatments did not complete the priming task. Five subjects assigned to treatments did so inappropriately, highlighting phrases that related to both the organization's mission and social activities despite being instructed to highlight only one of them.
3. The sessions were conducted in the summer and fall semesters of 2014. For each week of experiments, students could sign up for morning, noon, or afternoon sessions. Students were asked to list all their available times, and we later blindly assigned them to certain sessions based on their availability and session capacity. Sessions were cancelled if only one subject arrived as scheduled. That person was assigned to another session upon her agreement.
4. Subjects did not see the list of real-world charities until they completed their allocation tasks. This order ensures that subjects' allocation decisions were not influenced by their feelings about the particular charities on the list.
5. The average identity score is 4.88 (SD = 1.93) in the service treatment and 5.68 in the socializing treatment (SD = 2.17). There was not a significant difference in the scores between the two treatments ($t = -1.56$, $p = 0.13$). The scale is reliable (Cronbach's $\alpha = .96$) in measuring subjects' identity with the club in each setting, but it lacks construct validity for comparing our conditions and contains questions that are inapplicable to our control group. Thus, we did not end up using the scale in our analysis.
6. We ran a variety of alternative specifications, not reported here, with additional controls including the Unlikely Virtue score (as a level or as a dummy equaling 1 when the score was very high), a dummy equaling 1 for the majors of philanthropic studies, social work, and nursing,

and a dummy representing membership experience in any real-life service clubs. We also ran variants with interaction terms between these additional variables and the treatment variables.

None of these additional variables were robustly significant, and inclusion of these variables had small effects on our baseline coefficients, although the p values for the treatment effects sometimes varied by a percentage point or two.

7. We also ran a regression with empathic-concern and perspective-taking scores as two separate independent variables, and found that the latter was positive and significant at the 10% level.

8. We thank an anonymous referee for suggesting this robustness check.

References

- Akerlof, G. A., & Kranton, R. E. (2000). Economics and Identity. *Quarterly Journal of Economics* , 115 (3), 715-753.
- Andreoni, J. (1988). Privately Provided Public Goods in a Large Economy: The Limits of Altruism. *Journal of Public Economics* , 35, 57-73.
- Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. *Journal of Political Economy* , 97, 1447-1458.
- Ariely, Dan, Anat Bracha and Stephan Meier. (2009). Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, 99(1), 544-55.
- Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, 96 (5), 1652-1678.
- Charles, J. A. (1993). *Service Clubs in American Society: Rotary, Kiwanis, and Lions*. Urbana and Chicago: University of Illinois Press.
- Clark, P. B., & Wilson, J. Q. (1961). Incentive Systems: A Theory of Organizations. *Administrative Science Quarterly* , 6 (2), 129-166.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Auflage). Hillsdale, NJ: Erlbaum.
- Croson, R., and Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature*, 47(2), 1-27.
- Davis, M. H. 1983. "Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach." *Journal of Personality and Social Psychology*, 44, 113–126.

Deci, E. L. (1971). Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology*, 18 (1), 105-115.

Eikenberry, A. M. (2006). Giving Circles: Growing Grassroots Philanthropy. *Nonprofit and Voluntary Sector Quarterly*, 35, 517-532.

Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking Together or Falling Apart: In-group Identification as a Psychological Determinant of Group Commitment versus Individual Mobility. *Journal of Personality and Social Psychology*, 72, 617-626.

Kandori, M. (1992). Social Norms and Community Enforcement. *Review of Economic Studies*, 59, 63-80.

Karlan, D. & McConnell, M.A. (2014). Hey look at me: The effect of giving circles on giving. *Journal of Economic Behavior & Organization*, 106, 402-412.

Knoke, D. (1988). Incentives in Collective Action Organizations. *American Sociological Association*, 53#3, 311-329.

Ko, Y.J.; Rhee, Y.C.; Walker, M. and Lee, J.-H. (2014). What Motivates Donors to Athletic Programs: A New Model of Donor Behavior. *Nonprofit and Voluntary Sector Quarterly*, 43, 523-546.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28, 129-137.

Lions Clubs International. (n.d.). *Mission and History*. Retrieved August 22, 2015, from Lions Clubs International: <http://www.lionsclubs.org/EN/who-we-are/mission-and-history/index.php>

Mellström, C., & Johannesson, M. (2008). Crowding out in Blood Donation: Was Titmuss Right? *Journal of the European Economic Association*, 6 (4), 845-863 .

Molden, D.C. (2014). Understanding Priming Effects in Social Psychology: What is “Social Priming” and How does it Occur? *Social Cognition*, 32 (special issue), 1-11.

Morewedge, C.K. and Kahneman, D. (2010). Associative Processes in Intuitive Judgment. *Trends in Cognitive Sciences*, 14#10, 435-440.

Olson, M. (1965). *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Patrick, C. J.; Curtin, J. J., Tellegen, A. 2002. “Development and Validation of a Brief Form of the Multidimensional Personality Questionnaire.” *Psychological Assessment*, 14, 150-164.

Samson, A. and Miles, R. 2016. Selected Behavioral Science Concepts. In A. Samson (Ed.), *The Behavioral Economics Guide 2016*. Retrieved from <https://www.behavioraleconomics.com/mini-encyclopedia-of-be/priming-conceptual/>.

Samuelson, P. A. (1954). The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, 36 (4), 387-389.

Tajfel, H., & Turner, J. (1979). An Integrative Theory of Intergroup Conflict. In W. Austin, & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations* (pp. 33-47). Monterey, CA: Brooks/Cole.

Willer, R., Wimer, C., Owens, L.A.. (2015). What Drives the Gender Gap in Charitable Giving? Lower Empathy Leads Men to Give Less to Poverty Relief. *Social Science Research*. 52, 83-98.

Author Bibliography

Heng Qu is an assistant professor in the Bush School of Government and Public Service at Texas A&M University. Her research areas lie in nonprofit finance and economics. Her current research projects include experimental studies of charitable giving, nonprofit revenue management, and nonprofit financial reporting.

Steinberg is a professor of Economics and Philanthropic Studies at IUPUI. His research interests span the field, analyzing natural and experimental data on giving and volunteering, the behavior and role of nonprofit organizations, economics for nonprofit managers, and related governmental policies. He served as Co-President of ARNOVA from 1992-94.

Tables

Table 1 Experimental Design

Treatment and Condition	Group Assignment	Initial Selection	Organization Name	Membership Manipulation	Priming	Solicitation
Member, Service	Member	Choose an organization	Education Service Club (ESC)	New Member Welcome Letter	Service Dimension	Fundraising Message
Member, Socializing	Member	Choose an organization	Education Service Club (ESC)	New Member Welcome Letter	Socializing Dimension	Fundraising Message
Nonmember (Control)	Nonmember	Choose an organization	Education Service Association (ESA)	No	None	Fundraising Message

Table 2. Summary Statistics

Variables	All		Women		Men	
	N	Mean	N	Mean	N	Mean
Empathic Concern (EC)	91	22.02	59	23.56***	32	19.19***
Perspective Taking(PT)	91	19.33	59	19.53	32	18.97
EC and PT Combined	91	41.35	59	43.08**	32	38.16**
Unlikely Virtue (UV)	89	2.04	58	2.08	31	1.97

Variables		All		Women		Men	
		N	%	N	%	N	%
Major	Economics/ Business/ Engineering	30	32.97	13	22.03**	17	53.12**
	Philanthropic Studies/ Social Work/ Nursing	20	21.98	17	28.81*	3	9.38*
	Other	41	45.05	29	49.15	12	37.50
Membership in any service clubs	No	45	50.56	29	50.00	16	51.61
	Yes/Not now, but used to	44	49.44	29	50.00	15	48.39

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ for the test of equality of means or proportions between genders

Table 3. Treatment Effects

Panel A Whole Sample

	Control	Service	Socializing
Average donations	353.33	365.63	275.81
Std.Dev	199.54	218.29	202.43
Median	300	400	200
Bottom 25%	200	125	100
Top 25%	600	600	450
Number of 0 donations	1	2	2
Number of 600-token donations	8 (27%)	12 (38%)	6 (19%)
Number of subjects	30	32	31

Panel B Gender Difference

<i>Treatment</i>	<i>All</i>			<i>Women</i>			<i>Men</i>		
	Mean	Median	N	Mean	Median	N	Mean	Median	N
Control	353.33	300	30	372.73	300	22	300	300	8
Service	365.63	400	32	423.68	500	19	280.77	300	13
Socializing	275.81	200	31	268.42	200	19	309.09	300	11

Panel C Wilcoxon-Mann-Whitney tests (*p* values for the null of identical distributions)

<i>Treatment</i>	<i>All</i>	<i>Women</i>	<i>Men</i>
Service vs. Socializing	0.0982	0.0349	0.5928
Service vs. Control	0.7295	0.3723	0.941
Socializing vs. Control	0.1315	0.0844	0.8658

Panel D Effect Size

<i>All</i>	Effect Size (Cohen's d)	95% Confidence Interval	
Service vs. Socializing	0.426	-0.073	0.926
Service vs. Control	0.059	-0.44	0.557
Socializing vs. Control	0.386	-0.892	0.121
<i>Women</i>			
Service vs. Socializing	0.726	0.069	1.382
Service vs. Control	0.255	-0.361	0.871
Socializing vs. Control	0.506	-0.118	1.129
<i>Men</i>			
Service vs. Socializing	0.145	-0.659	0.949
Service vs. Control	0.089	-0.792	0.97
Socializing vs. Control	0.047	-0.958	0.864

Table 4. Tobit Regressions on Donations, Full Sample

	(1)	(2)	(3)	(4)
	Coefficients	Unconditional Average Marginal Effects	Coefficients	Unconditional Average Marginal Effects
Socializing	-128.537* (74.620)	-88.502* (50.821)	28.099 (113.604)	-97.452** (49.768)
Control	-38.829 (76.550)	-25.930 (51.067)	27.022 (126.373)	-33.401 (50.029)
Empathic Concern & Perspective Taking	7.823* (4.624)	5.316* (3.070)	6.565 (4.612)	4.468 (3.087)
Female	65.579 (67.584)	45.154 (46.824)	201.319* (108.404)	47.608 (46.975)
Female×Socializing			-266.605* (149.541)	
Female×Control			-125.962 (155.010)	
<i>N</i>	91	91	91	91
pseudo <i>R</i> ²	0.010		0.013	

Regressions estimated by double tobit with truncation at 0 and 600 tokens.

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Tobit Regressions on Donations, Female vs. Male Subsamples

	Female		Male	
	Coefficients	Unconditional Average Marginal Effects	Coefficients	Unconditional Average Marginal Effects
Socializing	-241.672** (103.992)	-157.441** (65.264)	28.256 (103.196)	21.787 (79.546)
Control	-99.600 (99.997)	-60.310 (59.722)	24.811 (116.536)	19.135 (89.774)
Empathic Concern & Perspective Taking	7.327 (6.526)	4.635 (4.071)	5.460 (6.302)	4.208 (4.767)
<i>N</i>	59	59	32	32
pseudo R^2	0.014		0.002	

Regressions estimated by double tobit with truncation at 0 and 600 tokens.

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$