Symbolic Incentives and the Recruitment of Volunteers for Citizen Science Projects

Simona Cicognani[a], Sebastian Stein[b], Mirco Tonin[c], and Michael Vlassopoulos[d, *]

[a]Leiden University (Department of Economics), Steenschuur 25 2311 ES Leiden The Netherlands, E-mail: s.cicognani@law.leidenuniv.nl
[b]University of Southampton (Department of Electronics and Computer Science), B32 East Highfield Campus University Road SO17 1BJ Southampton UK, E-mail: ss2@ecs.soton.ac.uk
[c]Free University of Bozen-Bolzano (Department of Economics), FBK-IRVAPP, CESifo, and IZA, Facoltá di Economia piazza Universitá 1 39100 Bolzano Italy, E-mail: mirco.tonin@unibz.it
[d]University of Southampton (Department of Economics) and IZA, B58 West Highfield Campus University Road SO17 1BJ Southampton UK, E-mail: M.Vlassopoulos@soton.ac.uk
[*]Corresponding author

Abstract

The provision of activities with external benefits that relies on voluntary contributions may often fall below societal needs. In this paper we focus on such contributions to a citizen science project (the World Community Grid) in which members of the general public are asked to offer unused computer power to advance cutting-edge scientific research. We investigate the role played by symbolic awards in stimulating existing contributors to recruit new contributors for this project. The recruitment campaign we study introduces badges for referrals (visible on each user's public profile page) varying, across randomized treatment groups, the threshold of successful referrals needed to receive these badges. We find that these symbolic incentives are effective in boosting referrals, and more so when the minimum threshold for achieving symbolic awards is higher. However, the overall effect of the incentives is quite modest, highlighting the challenges of running referral campaigns for the recruitment of volunteers.

JEL classification: C93; D64; H41.

1 Introduction

Referrals or word-of-mouth communications convey information through personal networks and can have an impact on important social and economic dimensions such as, for instance, technology adoption (Conley & Udry 2010), financial decisions (Duflo & Saez 2003, Banerjee et al. 2013), and brand choice by consumers (Godes & Mayzlin 2009). Organizations may actively deploy incentives to encourage referrals, for instance when hiring through referral bonuses to employees (Burks et al. 2015), or in their marketing, through, for example, discounts, free merchandise or other rewards to existing customers who engage in these activities (Aral & Walker 2011, Bapna & Umyarov 2015, Berman 2016, Wolters et al. 2020).[1]

In this paper, we study the role of symbolic incentives to refer acquaintances (known as refer-a-friend schemes) to contribute to the provision of a public good. In particular, we study recruitment of volunteers through referrals for a citizen science platform: the IBM World Community Grid (WCG), 'a platform that enables anyone with a computer, smartphone or tablet to donate their unused computing power to advance cutting-edge scientific research on topics related to health, poverty, and sustainability'.[2] WCG is one example of a large number of citizen science projects that are enabling researchers to draw on the contributions of millions of volunteers across the globe (including data, time or computational resources). Apart from WCG, other prominent citizen science projects include Zooniverse (2.2 million volunteers), iNaturalist (3.4 million volunteers), and NASA's Citizen Science Program (1.5 million volunteers).[3]

Recruiting volunteers for a citizen science project is akin to recruiting contributors to a global public good with a very large number of potential volunteers. Existing volunteers can be instrumental toward this end by drawing on their own social network and harnessing the power of social media. In the case of WCG, donating idle computing capacity has a negligible running cost after the initial effort of downloading and installing the software, but can come with a high perceived uncertainty, as the software could be a disguised virus or other malware. Therefore, being referred from a trusted user like a colleague or a friend could be a particularly effective mechanism to expand the network of contributors.

We use data from a randomized experiment involving almost 400,000 members of WCG implemented by the administrators of the platform to study the effectiveness of symbolic awards, which

[1] For a theoretical analysis of optimal referral channels, see Biyalogorsky et al. (2001), Lobel et al. (2016), and Carroni et al. (2020).

[2] See https://www.worldcommunitygrid.org/ (last accessed: 2023-02-09).

[3] See https://www.zooniverse.org/, https://www.inaturalist.org/, and https://science.nasa.gov/citizenscience, respectively. Figures correct as of January 2021.

take the form of badges that are visible on each user's public profile page, in encouraging the recruitment of new members by existing members. We also examine whether the specific form of the incentive scheme matters: the entry level and steepness, that is, the minimum number of new recruitees that is needed to receive the first award (a 'bronze badge'), and the additional number of recruitees that is needed to receive further awards ('silver badge', ...). The shape of the incentive scheme may affect both how many members become recruiters and how much effort they put into the recruitment activity. Thus, as we will explain in the experimental design section, a high initial threshold may discourage people to make any attempt at recruiting new members, but may encourage some to exert more effort.

It has to be kept in mind that, despite the large number of existing members receiving an invitation to participate in the referral campaign across the various treatments, the total number of new recruitees generated by the campaign is, including the recruitees stemming from the control group, rather modest, just 157. This low number is not so surprising given the channel of communication (an email) and the required action, as recruiters need to convince another person to register on a website and install a piece of software on one of his or her own computing devices. This low conversion rate is not unusual. For example, in the experiment by Lacetera et al. (2016), a charitable campaign reaching more than 6 million users generated only 30 donations. Also, evidence from previous studies analysing referral reward programs in other contexts suggests that the conversion rate is low. For example, in the context of a referral reward program aimed at attracting new bank customers, Wolters et al. (2020) report that the referral rate when a small monetary award is offered is 0.042%. This of course reduces the statistical power of our analysis to differentiate between the different incentive schemes.

What we find is that symbolic incentives more than double the number of recruiters and recruitees compared to the control group, but the specific form of the incentive scheme does not matter. When considering the number of recruitees, we find that the incentive scheme with the highest initial threshold achieves the best results, while the steepness does not matter. In our context, it thus appears that, at least for the entry level, discouragement effects are not important, while providing an ambitious initial threshold induces people to exert more effort.

This paper contributes to a growing literature on referrals. As mentioned earlier, referral incentives have been extensively studied in the field of marketing, as well as in labor economics for what concerns the recruitment of new employees. There are also contributions in the literature on fundraising. For instance, Meer (2011) studies the role of personal solicitations from a peer on

charitable giving, Castillo et al. (2014) study the effect of incentives on donors' willingness to share having donated by posting on their Facebook wall or by sending a private message to a friend on Facebook, Scharf & Smith (2016) and Payne et al. (2017) study the role of one's social network in individual online fundraisers, while Naroditskiy et al. (2014) study referral incentives in crowdfunding, where incentives are additional points to reach a target triggering a donation, thus not symbolic incentives as we study here.[4] What we study is not referrals for fundraising, but rather for volunteering to contribute to a global public good, which to the best of our knowledge has not been studied before.

Indeed, in the field of economics, the body of literature on the incentives for volunteering and the recruitment of volunteers is sparse compared to the literature on monetary donations.[5] Our study considers the recruitment of volunteers by existing ones for a type of volunteering that is not very time-consuming, as it entails using currently available resources with only minimal costs in terms of software installation. This is an instance of 'digital volunteering' (other examples include signing online petitions or providing online reviews without any rewards) the relevance of which is growing with the increasing importance of online interactions.

We also contribute to a recent literature on the effectiveness of symbolic awards that leverage on the self-esteem and social image concerns of the receiver (Bénabou & Tirole 2006, Frey & Neckermann 2008).[6] In a labour context, Kosfeld & Neckermann (2011) and Bradler et al. (2016) show that a congratulatory card increases performance (see also Neckermann & Yang 2017; Bradler & Neckermann 2019; Bruni et al. 2020), while in the context of blood donations Lacetera & Macis (2010) show that the prospect of a medal is instrumental in increasing the frequency of donations when the award is publicly announced and there is a public ceremony. Along this line, and more related to our online setting in which users' anonymity is preserved by pseudonyms, Gallus (2016) investigates the retention of new contributors to Wikipedia. The award in this case is randomly selected among new contributors and consists of a virtual medal visible only to online community members. Results show that receiving the award boosts the retention rate by 20%. This approach of awarding contributors with medals or badges is a common technique employed in *gamification*,

[4]More broadly, a number of studies have demonstrated the importance of peer effects in charitable giving and fundraising (e.g. Shang & Croson 2009; Smith et al. 2015).

[5]Lacetera et al. (2014), for instance, document that economic rewards boost blood donations, and the effect is stronger for higher incentives. Carpenter & Myers (2010) study volunteer firefighters and find that altruism as well as social image concerns matter. Brown et al. (2019) run a series of experiments to explain why, despite the high opportunity cost, some high earners prefer to donate their time (by volunteering) rather than donating money (see also Lilley & Slonim (2014)). Linardi & McConnell (2011) study the impact of the social environment on volunteering, showing the importance of the presence of peers.

[6]In a recent article, Adena & Huck (2020) disentangle the concerns for self-image from those for social image. Furthermore, they show how subjects embark on self-deception to keep their self-image at high levels.

the use of design elements from games in a non-game context (Deterding et al. 2011). There are numerous studies that show that such badges can increase the engagement of participants in online activities (Hamari et al. 2014, Seaborn & Fels 2015). However, work in this area typically focuses on encouraging existing participants to increase their contributions to a given activity, while our work specifically considers incentives to refer new members to a network. This is a particularly interesting distinction, in that one can be cooperative and helpful in the public good generation not only through his/her own direct contribution, but also through the additional channel we study in this paper: recruiting further contributors.

Since volunteers in the context that we study face several thresholds in terms of number of recruitees to receive the awards, our study also relates to the literature on the effect of thresholds in charitable giving. Besides being instrumental in receiving a badge, the thresholds we provide might also act as reference points or suggestions and, consequently, influence users' prosocial behaviour through this channel. In the context of charitable giving, according to Reiley & Samek (2019), donors might incur cognitive costs any time they deviate from the suggested amount, and people may interpret thresholds as a sort of guidance in a situation in which they do not know how much to donate, see for instance Adena et al. (2014). In terms of the effect on donations, the existing body of literature gives a mixed picture. Depending on the study, providing defaults/suggestions boosts charitable giving (Adena et al. 2014), has no effects (Altmann et al. 2019), or adverse effects (Reiley & Samek 2019, Adena & Huck 2020). This suggests that it is not just the presence of suggestions, but also their exact level that may matter. More focused on thresholds, Adena & Huck (2022) in their recent contribution on matching of donations underline that too low or too high thresholds may both be detrimental. Inspired by these findings, in our study we do not only explore the impact of having a threshold per se, but also have treatments in which we vary their levels. As we previously highlighted, in our setting thresholds have a different purpose compared to the above-mentioned studies on charitable giving: they are not suggestions for donations or triggers of matching donations, but minimum attainments for symbolic awards.

Finally, we study symbolic awards aimed at recruiting new members in a particularly interesting context, citizen science. This is a movement to engage the public in scientific research that, thanks to increased connectivity, is becoming increasingly relevant for the pursuit of scientific knowledge (Bonney et al. 2014). Contributors to citizen science projects often have a strong intrinsic motivation for their participation (Tinati et al. 2017), but a number of studies have looked at the use of extrinsic incentives to increase participation. Cappa et al. (2018) consider the effects of financial and symbolic

(acknowledgement on a public website) awards on the participants of a citizen science project. They show that both types of awards increase the participants' quantity of work, their enjoyment, and their willingness to refer others to the project. In contrast to our work, however, they do not track actual referrals and the study was performed in a laboratory setting. Badges are also commonly used in citizen science projects (Eveleigh et al. 2013, Tinati et al. 2017, Simperl et al. 2018), and there is evidence that they motivate participants to contribute (Bowser et al. 2013), but no existing work has looked at their potential to motivate referrals, which is the focus of our study.

The rest of the paper is organized as follows. Section 2 introduces the institutional context in which the field experiment was developed. Section 3 describes the experimental design and procedures, as well as the hypotheses. Section 4 presents the main results and section 5 provides the conclusions.

2 Institutional context

In this section, we describe the context in which the referral recruitment campaign took place and the pool of participants.

World Community Grid. WCG aims to use the Internet to foster collaborative scientific research. The goal is to combine computer resources and the shared knowledge of researchers to accelerate the pace of scientific discovery. WCG, a philanthropic initiative of IBM Corporate Citizenship, operates as a virtual supercomputer, harnessing spare computing power from volunteers' devices (computers, smartphones or tablets) to advance scientific research on topics related to health, poverty, and sustainability.

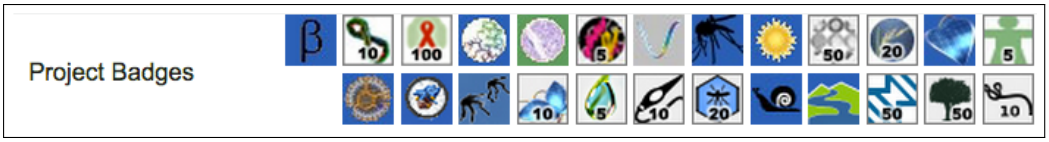
What a volunteer needs to do is to install an application on her device. The device then connects to WCG when not using its full capacity and performs small computational tasks, feeding the answers back to WCG. A volunteer can customize the computing preferences, limiting for instance the time period when to perform computations, or setting the maximum disk space and processor power available for computations. There also exists an option on whether or not to work while on battery and whether to start working only after the computer has been idle for a certain number of minutes.[7]

Members of the WCG receive badges for their contribution (see Figure 1). For instance, there is a badge named 'Gold Badge (90 days) for Help Defeat Cancer' or one called 'FightAIDS@Home'.

[7]The default settings are for the program to run always when not on battery, use no more than 10 GB of disk space and no more than 60% of the processor.

These badges appear on the public profile of each member. WCG has more than 30,000 teams or

Figure 1: Badges for contribution



groups which users can join. For instance, there are groups for universities or companies.

Although we do not have demographics of the users that were part of this study (apart from the country of origin of the device with WCG installed on), it is interesting to have some insights on the personal characteristics of WCG users. Data in this respect come from a general survey based on approximately 1,800 observations, which was carried out by WCG exactly one year before the intervention took place. Almost 30% of respondents come from the US, with the second and third most represented countries of origin being France and the UK (with 8.5% and 7%, respectively).[8] The most prevalent occupation of respondents is by far in the field of information technology (36%), followed by engineering (10.8%), pensioners (10.2%), and students (7.3%). In terms of education attainments, 45% of respondents have a college degree and 25% a post-college degree (Master's or Doctorate). About 67% of respondents are in an age range between 25 and 54 and the majority of participants are males (87.3%).

Our sample. Members of the WCG include both individuals and organizations. For instance, the largest contributor to the WCG is IBM itself, with an account with over 300,000 devices installed. As we are interested in the behaviour of individuals rather than large organizations, we restrict our sample to those with at most 10 devices installed.[9] We also exclude those without any device installed, as these signed up to the WCG, but have not actually installed the software and thus contributed to it. These excluded groups represent 3.7% and 15.3% of the 930,376 original total users, respectively. Given that the intervention we study involves sending emails, this also excludes those who did not provide consent to receive emails from WCG, representing around half of the sample. After this exclusion, we are left with 384,168 users.

Within the final pool of 384,168 users, the mode is to have just one device installed (59%), while 16.7% of users have two and 8.5% have three. Users have been part of WCG for an average of 5.7 years. On average, they are members of 0.45 groups, with 58.8% of them not being members

[8] Of the 160,509 users with non-missing values for country of origin in our sample (over a total of 384,168 users), 22.4% come from the US, which is broadly aligned with the data from the general survey. The proportions are also aligned for the other countries of origin.

[9] All the results that follow are robust to including in our sample only users with at most 3 devices installed.

of any group, and 38% being part of just one group. In terms of contributions, members have contributed on average the equivalent of 231 CPU days overall, of which 40 days in the year prior to the intervention, 2.8 and 0.63 days in the month and week prior to the referral campaign, respectively. Participants contributing a positive amount are 87% overall, 23% in the year prior to the intervention, 11% and 8.5% in the month and week prior to the intervention, respectively. Thus, active participation by users is not universal. Some users have never contributed, despite having registered and installed the software.

3 Experimental design and hypotheses

In this section, we detail the experimental design and procedures, providing also an overview of the timing of referral engagement. We then discuss our hypotheses.

3.1 The Experiment

The recruitment campaign consisted of introducing badges for referrals to WCG contributors. This was the first intervention of this type ever implemented on the platform. In particular, WCG created five different badges (Bronze, Silver, Gold, Ruby, and Emerald) and varied the thresholds needed to obtain each badge. In addition to a control group that received no symbolic awards for referrals, the design involved three treatments that are characterized by a different number of referrals needed to achieve the five badges, as displayed in Table 1.

Table 1: Treatments - Thresholds to receive badges

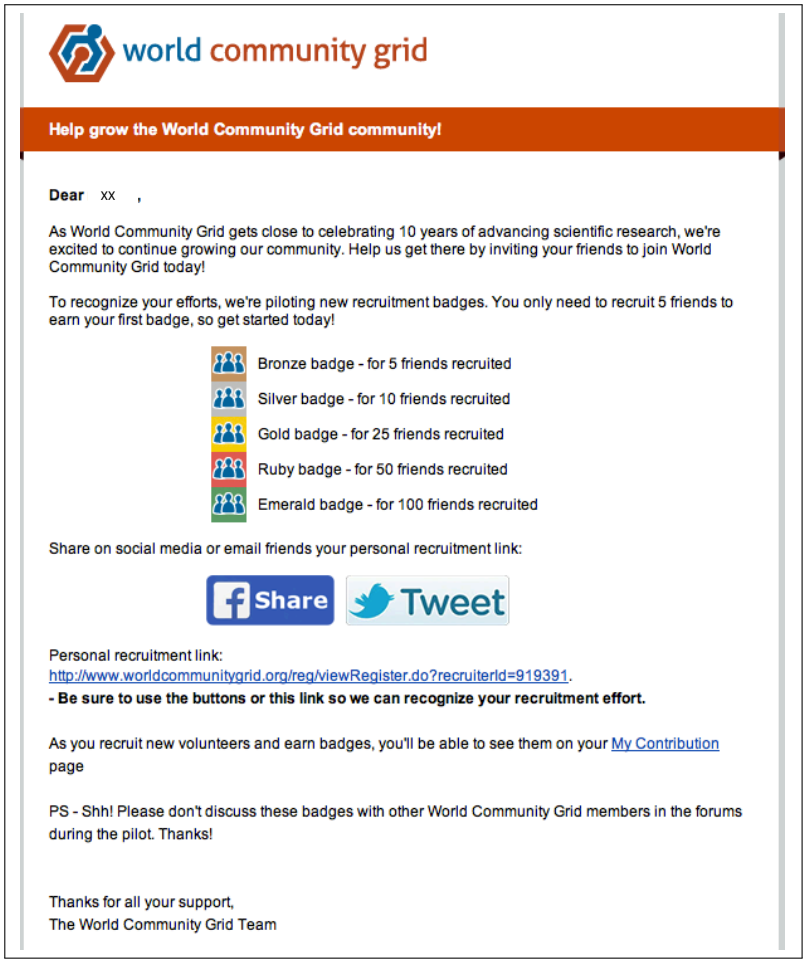
	Bronze	Silver	Gold	Ruby	Emerald
T1	1	5	10	25	100
T2	5	10	25	50	100
T3	1	10	25	50	100

In terms of the implementation of the experiment, users who had subscribed to receive emails from WCG were randomized into four groups, the three treatments plus a control group without badges. These users received an email inviting them to 'Recruit your friends to power scientific research' and detailing, for the treatment groups, the badge system (see Figure 2).

Users in the control group received the same recruitment email, without the references to the badge system.[10] Importantly, the subject of the email was the same across treatments, therefore the temptation to open the invitation email is orthogonal to treatments. The email included a personal

[10] Also, while in the control group we provide no minimum threshold for new recruitees, there is a 'logical' minimum threshold of 1 that coincides with what is provided in T1 and T3.

Figure 2: Recruitment Email



Notes: the recruitment email for the control group was the same as the one in this figure, except for the references to badges.

recruitment link through which WCG could track who recruited whom, and Facebook and Twitter buttons to spread the message through social media (see Figure 2).

In our final sample, about 15% of new recruitees were recruited via the Facebook link, 1.9% via the Twitter link, and the remaining via other means we cannot trace, e.g., emails. It is worth noting that it was not sufficient for a user to send an invitation to potential volunteers in order to earn a badge. In fact, the only referrals that count toward achieving the recruitment badge are those who signed up and contributed, and recruiters received an email notifying them about a successful referral (see Figure 3).

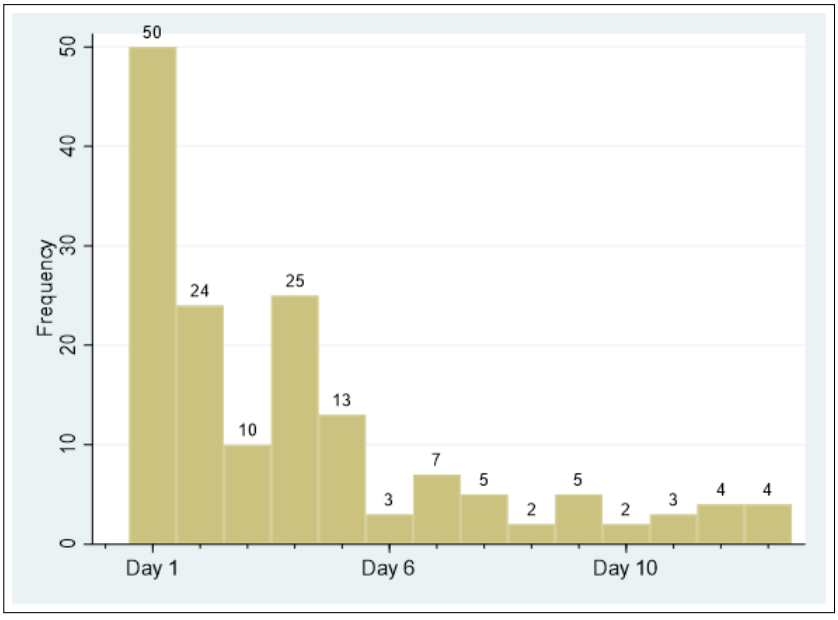
Users' referral behaviour was followed for three weeks. In terms of engagement, most of the recruitment activity took place on the day the invitation email was sent or soon after (see Figure 4). There is a peak in the number of new recruitees on the same day that the email invitation was sent, followed by a halving of the number of recruitees in the following day and a further decay in the

Figure 3: Message to successful recruiter

Notes: users' nicknames have been blurred for privacy reasons.

next days. Beyond 6 days post-invitation, the average number of recruitees per day is single-digit. Overall, the final sample includes 95,787 users for T1, 96,449 for T2, 96,330 for T3, and 95,602 for Control.

Figure 4: Timing new recruitees



3.2 Hypotheses

As explained above, users are awarded badges as a function of the number of recruitees they secure. We assume that the number of recruitees attained by a user is a non-decreasing function of the effort they exerted. In addition, inspired by the seminal work by Bénabou & Tirole (2006), we assume that the final utility that the user obtains hinges overall on three factors. The first one is an altruism component (e.g., a warm glow, as in Andreoni 1990) that would induce people to exert some effort even in the absence of badges. The second one is an image component, which we assume is increasing with the number of attained badges. Accumulating and displaying badges on one's WCG profile can indeed be desirable for improving the self-image of the user and his/her social image, to the extent to which s/he thinks to be recognizable on the platform. A last component of the utility function is a cost that is increasing in the amount of effort exerted in trying to attain recruitees. How effort translates into recruitees and, therefore, badges may differ (or be perceived to differ) from person to person. We assume that the first two components positively affect utility, whereas the cost of effort naturally exerts a negative effect.

The introduction of badges through the treatments activates the image component of the utility function, implying a higher effort compared to the control group. Thus, our first hypothesis is that the introduction of symbolic awards boosts the exertion of effort in recruiting. To test it, we compare the number of recruitees in the control group to the number of recruitees in the three treatments together.

Hypothesis 1 (Effect of symbolic awards). Introducing symbolic awards for referring new users increases the recruitment of volunteers: the number of recruitees in the treatment groups is higher compared to the control group.

A further issue concerns the specific design of the symbolic incentives. Across the three treatments, we varied the initial threshold needed to receive the first badge and the steepness of the incentive structure to receive the following badges. When designing the treatments, we (correctly) anticipated that users would hardly reach a threshold of 25 recruitees. However, we wanted to provide users in each treatment enough room for their ambitions. Importantly, the final threshold is the same in every treatment, so that users face the same expectations in terms of the final target.

Focusing on the initial threshold, a high one may appear unachievable, so that some users

may be discouraged to put in any effort at all.[11] We label this a discouragement effect. On the contrary, a high threshold may encourage people that decide to engage, to exert more effort to be able to achieve the ambitious target. We label this an ambition effect. Of course, these two effects are present also at the subsequent thresholds, but there may be differences in how they operate between different 'levels', for instance because there are decreasing returns to earning badges. The following two hypotheses provide a test.

In particular, to assess whether having a high initial threshold decreases or not the effort exerted in recruiting volunteers, we devised T2 such that the initial threshold is equal to 5, compared to T1 and T3 where it is equal to 1. Therefore, by comparing T2 *vs.* T1 and T3 we can test whether having a higher minimum threshold, as in T2, is detrimental or not for incentivizing recruitment effort. This leads to our second hypothesis:

Hypothesis 2 (Higher minimum threshold). If the ambition effect is stronger than the discouragement effect, a higher minimum threshold to achieve the first symbolic award for referring new users increases the effort exerted in recruiting volunteers: the number of recruitees in T2 is higher compared to T1 and T3.

To see the impact of higher targets beyond the first threshold, i.e., whether a higher steepness of the incentive structure discourages the exertion of effort in recruiting new volunteers, we devised T1 and T3 such that they start with the same initial (and low) threshold of 1, but then the threshold for the second level badge in T3 is double compared to T1, and also the following thresholds are higher, except the final one, which, as mentioned, is equal to 100 in every treatment. By comparing T1 *vs.* T3 we can assess whether a higher steepness of the incentive structure, as in T3, discourages subjects from exerting effort or not. This leads us to our third hypothesis:

Hypothesis 3 (Steepness of the thresholds). If the ambition effect is stronger than the discouragement effect, a higher steepness of the thresholds to achieve the symbolic awards for referring new users increases the effort exerted in recruiting volunteers: the number of recruitees in T3 is higher compared to T1.

[11] One potential mechanism is represented by the so-called 'goal-gradient hypothesis', according to which the farther away the goal is to being attained, the less effort is allocated to its achievement (Heath et al. 1999, Bonezzi et al. 2011). A possible channel in this theory of how goals affect motivation is a possible frustration that arises if the goal is perceived as unattainable or too difficult to attain (Louro et al. 2007, Huang et al. 2012).

4 Results

Overall effect of the campaign. Our first aim is to assess whether the intervention was overall effective in boosting participation in the WCG. To this end, we take into account the impact of the campaign on both the number of recruiters and the number of recruitees. Looking at the number of recruiters, we assess how many users on the WCG platform were influenced by the campaign, i.e., the number of users that successfully recruited at least one new volunteer. Recruiters of course can be more or less productive in their recruitment effort. For this reason, another output measure worth investigating is the number of recruitees, i.e., the total number of new users that decided to join the WCG thanks to the campaign. Given that the main aim of the campaign was to expand the network, this is also a relevant metric. Even if not all recruitees are immediately active, all have installed the software and therefore may become active later on. Moreover, existing members need to put kn effort to recruit new members, irrespective of their activity status within the three-weeks observation period. Therefore, the number of recruitees is a good indicator of effort. For these reasons, we focus our attention on all recruitees rather than on the subset of active recruitees. As a descriptive measure of recruiter productivity, we also computed the number of recruitees per recruiter.

Table 2 displays the number of recruiters and recruitees across treatments, as well as some other relevant output measures, including the referral rate, which is defined as the percentage of users who were successful in referring a new user. As the number of users in each condition is very similar but not identical, this metric is useful to correct for these differences. We also report the number of bronze badges actually obtained in the various treatments (remember that badges are assigned on the basis of active recruitees). No one managed to go beyond the first level. To help the reader remember the differences across treatments, we also indicate in the name the first three thresholds, so that T1 becomes T1 1-5-10.

Among the 384,168 users that constitute our sample, 112 users recruited at least one subject, for a total of 157 recruitees (among recruiters, the average number of subjects recruited per recruiter is 1.4, with a maximum of 13). Figure 5 shows the number of recruitees per recruiter across the four treatments. The mode in every treatment is to recruit only one new user. Figure 6 zooms in on the productivity of the recruitment and shows the subset of recruitees who signed up and are active in terms of CPU donation to WCG in the three weeks of observation.

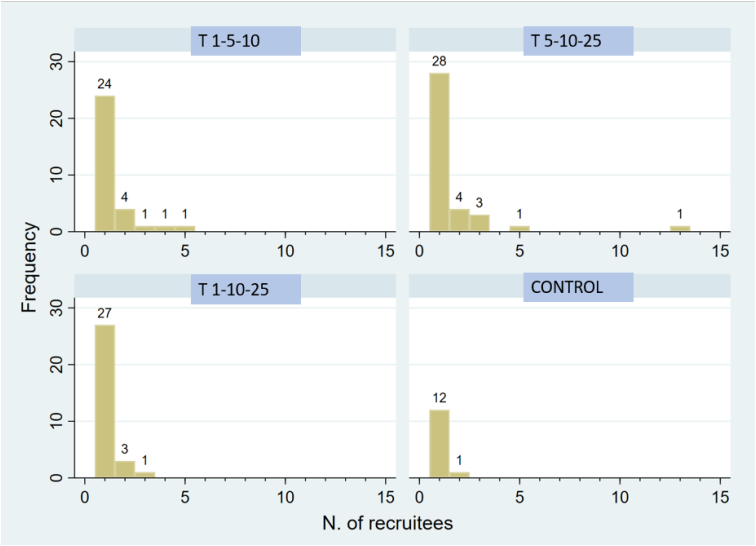
When looking at the number of recruitees per user, we find a clear boost comparing the intervention treatments pooled together (0.05%) with the *Control* group (0.015%). A test of proportions

Table 2: Output measures

	T1 1-5-10	T2 5-10-25	T3 1-10-25	Control
N. of recruiters	31	37	31	13
N. of recruitees	44	63	36	14
N. of active recruitees	21	17	10	9
Total points per recruitee	39,602	24,574.94	9,993	9,863.33
Total points	831,642	417,774	99,930	88,770
N. recruitees per recruiter	1.42	1.70	1.16	1.08
N. active recruitees per recruiter	.68	.46	.32	.69
N. users	95,787	96,449	96,330	95,602
Referral rate	0.032%	0.038%	0.031%	0.014%
N. bronze badges	14	1	10	-

Notes: Referral rate is the percentage of users who were successful in referring a new user.
Source: Authors' calculations.

Figure 5: Recruitees per recruiter



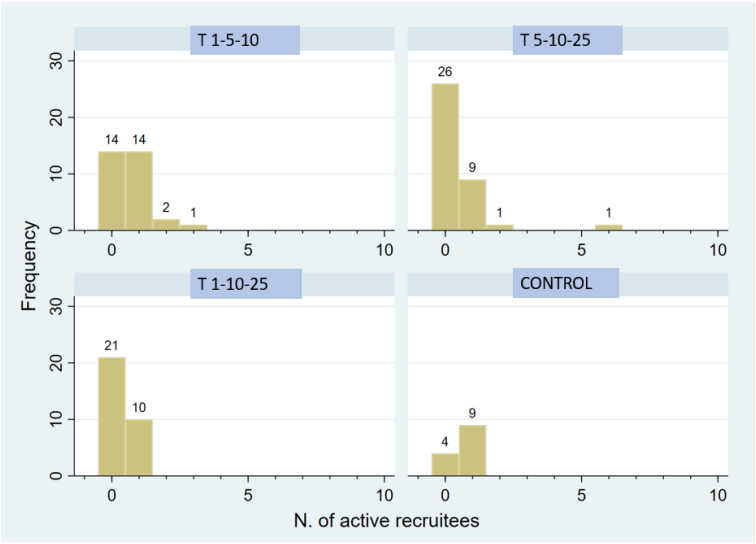
indicates the difference to be statistically significant (z = -4.63, p = 0.000). We can thus conclude that the symbolic incentives for referrals did work. In Figure 7, we report these metrics for each treatment separately, as well as the test result of pair-wise comparisons. As we can see, each single treatment delivers higher recruitees per user than the control treatment.

Looking at the other metric, the number of recruiters per user, we see, again, that the referral rate in the treatment groups (0.034%) is higher than in the control group (0.014%). A test of proportions indicates the difference to be statistically significant (z = -3.2507, p = 0.0012). Figure 8 reporting this metric for each treatment separately is in this regard very similar to Figure 7.

To summarize, in line with hypothesis 1, we find:

Result 1 *The existence of symbolic awards for referring new users has a positive overall effect on*

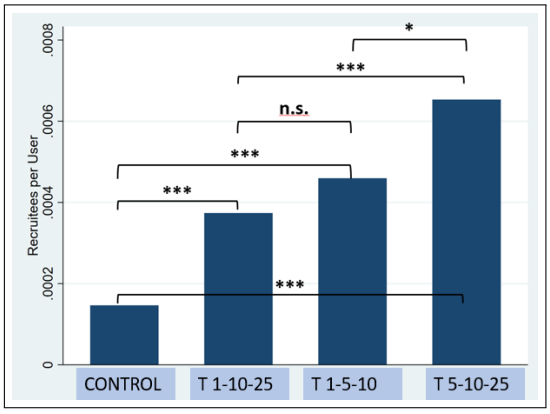
Figure 6: Active recruitees per recruiter



the number of recruitees, as well as on the number of recruiters.

Treatment differences. The analysis conducted so far also allows us to compare differences across the intervention treatments, in order to shed light on whether the design of symbolic incentives matters.

Figure 7: Recruitees per User across treatments



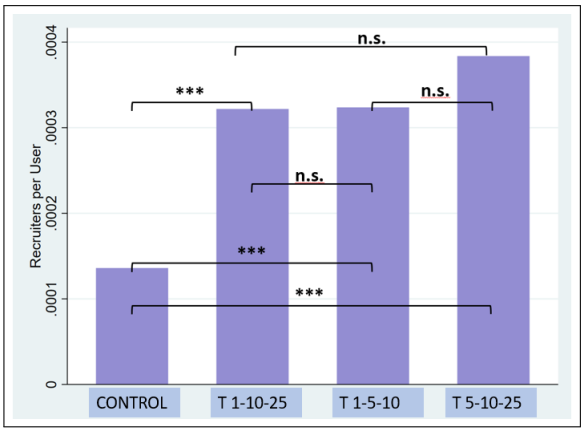
Notes: ***, **, * indicate the significance at the 1%, 5% and 10% respectively from a test of proportions, whereas n.s. indicates not significant.

With regards to the overall number of recruitees per user (see Figure 7), we find that T2 5-10-25 is the treatment with the highest number of recruitees per user, which means that having a high minimum threshold is the most effective channel for stimulating referrals.[12] This is in line with hypothesis 2. Instead, related to hypothesis 3, we do not find any effect of the steepness of the

[12] A pairwise test of proportions between T2 5-10-25 and the other two treatments yields p=0.0716 vs T1 1-5-10, and p=0.0068 vs T3 1-10-25.

incentive structure (T1 1-5-10 does not statistically differ from T3 1-10-25; test of proportions, z = 0.9199, p = 0.3576).

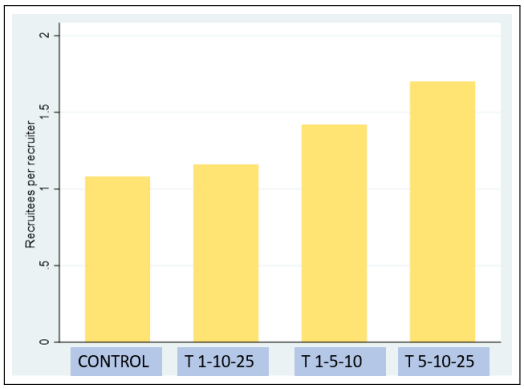
Figure 8: Recruiters per User across treatments



Notes: ***, **, * indicate the significance at the 1%, 5% and 10% respectively from a test of proportions, whereas n.s. indicates not significant.

When we move to the analysis of the number of recruiters (see Figure 8), we do not find evidence of any between-treatment differences in any of the pairwise comparisons.

Figure 9: Recruitees per recruiter across treatments



Notes: a Mann-Whitney test indicates no statistically significant differences in the pairwise comparisons.

To summarize, the comparison across treatments provides the following result related to hypotheses 2 and 3:

Result 2 *A minimum threshold equal to 5 in the incentive structure leads to a larger number of recruitees per user relative to a minimum threshold equal to 1, while the steepness of incentives does not matter.*

As a descriptive exercise, bearing in mind that to do this we restrict the analysis to the small sample of active recruiters, in Figure 9 we show the number of recruitees per recruiter. We see that

T2 5-10-25 is still the treatment with the highest outcome, though a Mann-Whitney test fails to reject equality against the control or any of the other two treatment groups (T2 5-10-25 vs Control, p=0.178).

Another outcome that is interesting to assess is the quality of recruitees across treatments. We can in fact distinguish active recruitees (those who donated a positive amount of CPU to the WCG) from non-active ones, and more specifically we can record the total points each recruitee accumulated for the hours of CPU donated to WCG. In principle, recruitees are not incentivized by the campaign, therefore their behaviour should not hinge on treatments. However, it could be that recruiters themselves implement recruitment strategies that are influenced by treatments, for instance by contacting potential recruitees that they expect are more likely to contribute to the WCG. We find that total points in T2 5-10-25 are higher than in T3 1-10-25 at the 5% significance level (Wilcoxon rank-sum test, z = 2.109, p = 0.0350), and that total points in T2 5-10-25 are higher than in C at the 10% significance level (Wilcoxon rank-sum test, z = 1.913, p = 0.0557), whereas we do not find any statistically significant differences across the other treatments.

5 Conclusions

Research on voluntary provision of public goods has been growing fast, but much still has to be learnt about the role of symbolic awards in this domain. Our research contributes to the literature by studying whether and to what extent symbolic awards play a role in enhancing the recruitment effort of contributors in an online community devoted to citizen science. The relevance of our approach is illustrated by the pervasiveness of refer-a-friend campaigns in several and diverse domains. We assess the effectiveness of a randomized referral campaign on the World Community Grid, an online community which accounts for more than 700,000 registered users. In this campaign, the thresholds for receiving referral badges, which are visible on each user's public profile page, were exogenously varied across treatment groups. We believe that the study of online communities is extremely relevant in itself, especially after the Covid-19 pandemic, during which there has been a sharp acceleration of digital technology solutions. Indeed, according to a recent report (GlobalWebIndex 2020), about 76% of internet users are also participants in online communities. In this sense, we contribute toward a better understanding of the effectiveness of refer-a-friend programs in volunteering activities in an emerging setting.

We find evidence that symbolic incentives are effective in increasing the number of referrals among users. Specifically, the most powerful incentive for triggering referrals is one with a high initial

threshold for achieving symbolic awards. In terms of interpretation of the results, as mentioned in the introduction, badges - and more generally symbolic awards - can exert their effect through self-esteem and social image concerns of the receiver. It would be interesting to disentangle these two channels, like in Adena & Huck (2020), for instance by running the very same intervention making the badges only privately available. This is, however, not practically feasible in the field context we study, because the common practice within the World Community Grid is that badges appear on the public profile of each member. Considering previous studies on online communities, it is usually the case that removing anonymity enhances donations and charitable behaviour (Soetevent 2005, Lacetera & Macis 2010). The issue in our context is whether the use of pseudonyms - as it is customary in the WCG - is enough to induce a behaviour similar to anonymity, where social image concerns are not at work and, therefore, only self-image motivations are present. Evidence from the social identification theory on online communities shows that members of online communities may feel a sense of belonging and self-identification with the community, the more so if their contributions are publicly acknowledged as valuable (Ren et al. 2012, Gallus 2016). If this line of reasoning prevails, the sense of belonging to the online community may trigger social image concerns even in absence of a direct identification of users. In this case, then both self-esteem and self-image concerns would be active.

Even though we find a positive effect of the campaign, the overall effect of incentives is, however, quite modest. One of the campaign's drawbacks that we document is the rapid decay in the response rate: we start with a massive number of users operating on the WCG, yet half of these users did not provide consent to receive email notifications (the channel through which the campaign takes place). Moreover, among the users reached by the intervention, it seems very hard to motivate them to refer a friend: 112 users over a total of 384,168 referred at least a friend who actually joined the WCG. As a consequence, this reduces our ability to discriminate the effect of the different treatments, which constitutes the main limitation of our study. One could speculate that referrals may be more effective in other (online) contexts. The ability to invite others to join critically hinges on the extent of the participants' social network. It is possible that our sample of mainly males working in IT and engineering has, on average, a narrower social network of, say, a community of political activists. If this is indeed the case, we could expect a similar 'refer-a-friend' campaign in a platform devoted to political activism (e.g., change.org) to be more successful.

It is also of interest to compare referrals to other ways of recruiting volunteers for citizen science projects. One obvious candidate is the use of social media (see, for instance, Robson et al. (2013)

and Crall et al. (2017)). Oliveira et al. (2021) provide an overview of the social media use by inspecting 300 citizen science projects from around the world and find that more than one-third of the projects investigated used at least one social media platform, with Facebook and Twitter accounting for approximately 70% of the social media use. However, even these activities can have a comparatively low rate of success. Thomas et al. (2022) focused on a health intervention, investigating the response of participants who captured and classified images of alcohol advertising they saw online. They found that a Facebook recruitment post placed by the cancer association reached an engagement rate of 2.41% (i.e., the percentage of subjects reached, who then clicked on the Facebook button of the campaign). In contrast, two paid posts placed by the research study reached an engagement rate of 5.09%.

The finding of a low uptake rate relates our work to the recent literature on 'illusion of activism', which suggests that although many people show support or interest for a good cause, they don't follow up with real engagement for it (Lacetera et al. 2016). Further research could prove useful in investigating the reasons behind the low take up rate of programs like the one we studied here. For instance, in our case, participants contacted via email or Facebook by existing WCG users may be skeptical about the platform and fear that downloading an application that operates when their electronic device is idle may also give, to unknown entities, access to their own personal data. Investigating ways of overcoming such limitations in order to fully harness the benefits of referrals is left for further research.

Supplementary material

Supplementary material is available on the OUP website. These are the replication files. The data used in this paper can be requested by contacting the WCG at: `https://www.worldcommunitygrid.org/viewContactUs.do`

Acknowledgements

We would like to thank the WCG for sharing their data with us.

References

Adena, M. & Huck, S. (2020), 'Online fundraising, self-image, and the long-term impact of ask avoidance', *Management Science* **66**, 722–743.

Adena, M. & Huck, S. (2022), 'Personalized fundraising: A field experiment on threshold matching of donations', *Journal of Economic Behavior & Organization* **200**, 1–20.

Adena, M., Huck, S., Rasul, I. et al. (2014), 'Charitable giving and nonbinding contribution-level suggestions—evidence from a field experiment', *Review of Behavioral Economics* **1**, 275–293.

Altmann, S., Falk, A., Heidhues, P., Jayaraman, R. & Teirlinck, M. (2019), 'Defaults and donations: Evidence from a field experiment', *Review of Economics and Statistics* **101**, 808–826.

Andreoni, J. (1990), 'Impure altruism and donations to public goods: A theory of warm-glow giving', *The Economic Journal* **100**, 464–477.

Aral, S. & Walker, D. (2011), 'Creating social contagion through viral product design: A randomized trial of peer influence in networks', *Management Science* **57**, 1623–1639.

Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. (2013), 'The diffusion of microfinance', *Science* **341**, 1236498.

Bapna, R. & Umyarov, A. (2015), 'Do your online friends make you pay? A randomized field experiment on peer influence in online social networks', *Management Science* **61**(8), 1902–1920.

Bénabou, R. & Tirole, J. (2006), 'Incentives and prosocial behavior', *American Economic Review* **96**, 1652–1678.

Berman, B. (2016), 'Referral marketing: Harnessing the power of your customers', *Business Horizons* **59**, 19–28.

Biyalogorsky, E., Gerstner, E. & Libai, B. (2001), 'Customer referral management: Optimal reward programs', *Marketing Science* **20**, 82–95.

Bonezzi, A., Brendl, C. M. & De Angelis, M. (2011), 'Stuck in the middle: The psychophysics of goal pursuit', *Psychological Science* **22**, 607–612.

Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J. & Parrish, J. K. (2014), 'Next steps for citizen science', *Science* **343**, 1436–1437.

Bowser, A., Hansen, D., He, Y., Boston, C., Reid, M., Gunnell, L. & Preece, J. (2013), 'Using gamification to inspire new citizen science volunteers', `https://doi.org/10.1145/2583008.2583011`. Last accessed: 2023-02-09.

Bradler, C., Dur, R., Neckermann, S. & Non, A. (2016), 'Employee recognition and performance: A field experiment', *Management Science* **62**, 3085–3099.

Bradler, C. & Neckermann, S. (2019), 'The magic of the personal touch: Field experimental evidence on money and appreciation as gifts', *The Scandinavian Journal of Economics* **121**, 1189–1221.

Brown, A. L., Meer, J. & Williams, J. F. (2019), 'Why do people volunteer? an experimental analysis of preferences for time donations', *Management Science* **65**, 1455–1468.

Bruni, L., Pelligra, V., Reggiani, T. & Rizzolli, M. (2020), 'The pied piper: Prizes, incentives, and motivation crowding-in', *Journal of Business Ethics* **166**, 643–658.

Burks, S. V., Cowgill, B., Hoffman, M. & Housman, M. (2015), 'The value of hiring through employee referrals', *The Quarterly Journal of Economics* **130**, 805–839.

Cappa, F., Laut, J., Porfiri, M. & Giustiniano, L. (2018), 'Bring them aboard: Rewarding participation in technology-mediated citizen science projects', *Computers in Human Behavior* **89**, 246 – 257.

Carpenter, J. & Myers, C. K. (2010), 'Why volunteer? evidence on the role of altruism, image, and incentives', *Journal of Public Economics* **94**, 911–920.

Carroni, E., Pin, P. & Righi, S. (2020), 'Bring a friend! privately or publicly?', *Management Science* **66**, 2269–2290.

Castillo, M., Petrie, R. & Wardell, C. (2014), 'Fundraising through online social networks: A field experiment on peer-to-peer solicitation', *Journal of Public Economics* **114**, 29–35.

Conley, T. G. & Udry, C. R. (2010), 'Learning about a new technology: Pineapple in Ghana', *American Economic Review* **100**, 35–69.

Crall, A., Kosmala, M., Cheng, R., Brier, J., Cavalier, D., Henderson, S. & Richardson, A. (2017), 'Volunteer recruitment and retention in online citizen science projects using marketing strategies: lessons from season spotter', *Journal of Science Communication* **16**, A01.

Deterding, S., Dixon, D., Khaled, R. & Nacke, L. (2011), 'From game design elements to gamefulness: Defining 'gamification'', `https://doi.org/10.1145/2181037.2181040`. Last accessed: 2023-02-09.

Duflo, E. & Saez, E. (2003), 'The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment', *The Quarterly Journal of Economics* **118**, 815–842.

Eveleigh, A., Jennett, C., Lynn, S. & Cox, A. L. (2013), "I want to be a captain! I want to be a captain!': Gamification in the old weather citizen science project', `https://doi.org/10.1145/2583008.2583019`. Last accessed: 2023-02-09.

Frey, B. S. & Neckermann, S. (2008), 'A view from psychological economics', *Zeitschrift für Psychologie/Journal of Psychology* **216**, 198–208.

Gallus, J. (2016), 'Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia', *Management Science* **63**(12), 3999–4015.

Godes, D. & Mayzlin, D. (2009), 'Firm-created word-of-mouth communication: Evidence from a field test', *Marketing Science* **28**, 721–739.

Hamari, J., Koivisto, J. & Sarsa, H. (2014), 'Does gamification work?–a literature review of empirical studies on gamification', `https://doi.org/10.1109/HICSS.2014.377`. Last accessed: 2023-02-09.

Heath, C., Larrick, R. P. & Wu, G. (1999), 'Goals as reference points', *Cognitive Psychology* **38**, 79–109.

Huang, S.-c., Zhang, Y. & Broniarczyk, S. M. (2012), 'So near and yet so far: The mental representation of goal progress.', *Journal of Personality and Social Psychology* **103**, 225.

Kosfeld, M. & Neckermann, S. (2011), 'Getting more work for nothing? symbolic awards and worker performance', *American Economic Journal: Microeconomics* **3**, 86–99.

Lacetera, N. & Macis, M. (2010), 'Social image concerns and prosocial behavior: Field evidence from a nonlinear incentive scheme', *Journal of Economic Behavior & Organization* **76**, 225–237.

Lacetera, N., Macis, M. & Mele, A. (2016), 'Viral altruism? Charitable giving and social contagion in online networks', *Sociological Science* **3**, 202–238.

Lacetera, N., Macis, M. & Slonim, R. (2014), 'Rewarding volunteers: a field experiment', *Management Science* **60**, 1107–1129.

Lilley, A. & Slonim, R. (2014), 'The price of warm glow', *Journal of Public Economics* **114**, 58–74.

Linardi, S. & McConnell, M. A. (2011), 'No excuses for good behavior: Volunteering and the social environment', *Journal of Public Economics* **95**, 445–454.

Lobel, I., Sadler, E. & Varshney, L. R. (2016), 'Customer referral incentives and social media', *Management Science* **63**, 3514–3529.

Louro, M. J., Pieters, R. & Zeelenberg, M. (2007), 'Dynamics of multiple-goal pursuit', *Journal of Personality and Social Psychology* **93**, 174.

Meer, J. (2011), 'Brother, can you spare a dime? Peer pressure in charitable solicitation', *Journal of Public Economics* **95**, 926–941.

Naroditskiy, V., Stein, S., Tonin, M., Tran-Thanh, L., Vlassopoulos, M. & Jennings, N. R. (2014), 'Referral incentives in crowdfunding', `https://doi.org/10.1609/hcomp.v2i1.13147`. Last accessed: 2023-02-09.

Neckermann, S. & Yang, X. (2017), 'Understanding the (unexpected) consequences of unexpected recognition', *Journal of Economic Behavior & Organization* **135**, 131–142.

Oliveira, S. S., Barros, B., Pereira, J. L., Santos, P. T. & Pereira, R. (2021), 'Social media use by citizen science projects: Characterization and recommendations', *Frontiers in Environmental Science* **9**, 715319.

Payne, A., Scharf, K. & Smith, S. (2017), 'Individual fundraising: The power of the personal', *Social economics: Current and emerging avenues* p. 305.

Reiley, D. & Samek, A. (2019), 'Round giving: a field experiment on suggested donation amounts in public-television fundraising', *Economic Inquiry* **57**, 876–889.

Ren, Y., Kraut, R., Kiesler, S. & Resnick, P. (2012), 'Encouraging commitment in online communities', *Building successful online communities: Evidence-based social design* pp. 77–124.

Robson, C., Hearst, M., Kau, C. & Pierce, J. (2013), Comparing the use of social networking and traditional media channels for promoting citizen science, *in* 'Proceedings of the 2013 conference on Computer supported cooperative work', pp. 1463–1468.

Scharf, K. & Smith, S. (2016), 'Relational altruism and giving in social groups', *Journal of Public Economics* **141**, 1–10.

Seaborn, K. & Fels, D. I. (2015), 'Gamification in theory and action: A survey', *International Journal of human-computer studies* **74**, 14–31.

Shang, J. & Croson, R. (2009), 'A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods', *The Economic Journal* **119**(540), 1422–1439.

Simperl, E., Reeves, N., Phethean, C., Lynes, T. & Tinati, R. (2018), 'Is virtual citizen science a game?', *ACM Transactions on Social Computing* **1**, 1–39.

Smith, S., Windmeijer, F. & Wright, E. (2015), 'Peer effects in charitable giving: Evidence from the (running) field', *The Economic Journal* **125**, 1053–1071.

Soetevent, A. R. (2005), 'Anonymity in giving in a natural context—a field experiment in 30 churches', *Journal of Public Economics* **89**, 2301–2323.

Thomas, J. A., Trigg, J., Morris, J., Miller, E. & Ward, P. R. (2022), 'Exploring the potential of citizen science for public health through an alcohol advertising case study', *Health promotion international* **37**, 1–10.

Tinati, R., Luczak-Roesch, M., Simperl, E. & Hall, W. (2017), 'An investigation of player motivations in eyewire, a gamified citizen science project', *Computers in Human Behavior* **73**, 527–540.

Wolters, H. M., Schulze, C. & Gedenk, K. (2020), 'Referral reward size and new customer profitability', *Marketing Science* **39**, 1166–1180.