

PledgeWork: Online Volunteering through Crowdwork

Keiko Katsuragawa
School of Computer Science,
University of Waterloo
National Research Council
Waterloo, Canada
kkatsura@uwaterloo.ca

Qi Shu
School of Computer Science,
University of Waterloo
Waterloo, Canada
q3shu@edu.uwaterloo.ca

Edward Lank
School of Computer Science,
University of Waterloo
Waterloo, Canada
lank@uwaterloo.ca

ABSTRACT

In this paper, we explore an alternative form of volunteer work, PledgeWork, where individuals, rather than working directly for a charity, make indirect donations by completing tasks provided by a third party task provider. PledgeWork poses novel research questions on issues of user acceptance of on-line volunteerism, on quality and quantity of work performed as a volunteer, and on the benefits low-barrier volunteerism might provide to charities. To evaluate these questions, we conduct a mixed methods study that compares the quality and quantity of work between volunteer workers and paid workers and user attitudes toward PledgeWork, including perceived benefits and drawbacks. We find that PledgeWork can improve the quality of simple tasks and that the vast majority of our participants expressed interest in using our PledgeWork platform to contribute to a charity. Our interview also reveals current problems with volunteering and online donations, thus highlighting additional strengths of PledgeWork.

KEYWORDS

crowdsourcing, volunteering, controlled experiments

ACM Reference Format:

Keiko Katsuragawa, Qi Shu, and Edward Lank. 2019. PledgeWork: Online Volunteering through Crowdwork. In *CHI Conference on*

This article was authored by employees of the Government of Canada. As such, the Canadian government retains all interest in the copyright to this work and grants to ACM a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, provided that clear attribution is given both to the authors and the Canadian government agency employing them. Permission to make digital or hard copies for personal or classroom use is granted. Copies must bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the Canadian Government must be honored. To copy otherwise, distribute, republish, or post, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

© 2019 Crown in Right of Canada. Publication rights licensed to ACM.

ACM ISBN 978-1-4503-5970-2/19/05...\$15.00

<https://doi.org/10.1145/3290605.3300541>

Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK. ACM, New York, NY, USA, 11 pages.
<https://doi.org/10.1145/3290605.3300541>

1 INTRODUCTION

People support charitable causes via donations of their time and/or money. Volunteering – either through necessity or preference – is typically done in the real world, and, as a result, requires a significant time commitment. In contrast, the advantage of donations is that they require less time commitment from donors in exchange for a financial and privacy cost to one’s support of the charity, a factor that can be potentially challenging for people who wish to support causes that they find important.

In this paper, we propose *PledgeWork* an alternative form of on-line volunteering. PledgeWork involves three parties; charitable organizations, task requesters and volunteer workers. The task requesters post their task requests to an on-line PledgeWork system and deposits the cost for the task. Then the volunteers complete the tasks to support a charity of their choice. Once the volunteers complete the task, the deposit will be donated to the charities the volunteers specified.

The premise of this work is that crowdwork could prove to be a valuable potential avenue of support for charitable causes. Specifically, because crowdwork is done solo in a room, typically at home, and on any schedule and for any period from a short period of time to a longer period, potential donors are free to perform tasks at their convenience. Furthermore, because individuals can specifically perform tasks to support the charity without needing to make a financial contribution, the barriers to financially supporting charitable causes an individual values – the need for a credit card, high costs of small donations, loss of privacy – may appear lower. For example, an individual may feel silly filling out a lengthy online form to donate a few dollars and cannot spend an afternoon performing a task with a group of volunteers, but perhaps they can do a series of tasks – as many as they choose and on a flexible schedule – that would result in a donation of the few dollars to a charity they value.

PledgeWork also provides benefits to charities as a result of a diversified revenue stream. The charity can recruit the volunteers both through the PledgeWork crowdsourcing site and from their own connections. For example, charities can provide the link to PledgeWork on their website as an alternative donation method along with credit card and other payment options.

As we began to unpack the process of PledgeWork, several questions regarding the process of donation via PledgeWork arose. For example, how will potential crowdworkers feel about givings going to charity versus to themselves? Will donors perform tasks to the same high (or low) quality and quantity as individuals who do crowdwork for pay? More generally, how does online donations of time to a charity compare to the act of physically volunteering or directly giving money, i.e. are there other unanticipated benefits or drawbacks to PledgeWork?

To address these questions, we develop a simulated online volunteering platform that allows the volunteer workers to donate their time for a charity they choose. As a control condition, we also develop an identical simulated online crowdsourcing platform where paid workers make money by completing same tasks as the volunteer workers. Given the simulated platforms, we evaluate the feasibility and user acceptance of PledgeWork via a two-part – task followed by interview – 2×2 mixed design study that contrasts PledgeWorkers with paid workers (a between subjects factor) for simple and more complex online tasks (audio annotation and text editing, a within subjects factor).

Overall, our user study validates the potential for online PledgeWork as a form of charitable support. In particular, the convenience of short, low-direct-cost charitable support was perceived of as valuable to our participants. More interestingly, while both pledge and paid workers worked for statistically similar time periods on a task, paid workers did significantly more tasks overall. However, detailed analysis found some evidence that PledgeWork may limit the number of crowdworkers who do low quality crowdwork.

The remainder of this paper is organized as follows. We first review related work, including work on crowdsourcing, altruism, and current opportunities for crowd-based volunteering through initiatives like citizen science. We then describe the three-party conceptualization of PledgeWork. Finally, we describe our user study and present data from both user behavior logging and from interviews, synthesizing these into a cohesive view of the potential benefits of online PledgeWork.

2 RELATED WORK

Crowdsourcing has been used as a volunteering platform for various causes. Citizen science is, perhaps, the most prominent example of successfully using crowdsourcing for scientific research [22]. Tinati [24] categorized citizen science tasks into three categories: data collection/curation [9, 18], data analysis [7, 8, 25] and problem solving (e.g. algorithm discovery) [13].

One challenge of crowdsourcing is quality control. To improve the quality of crowdsourcing work, Mason and Watts [17] found that monetary incentives increase the quantity of the work but not the quality, a result that has been confirmed by multiple additional studies [3, 20]. Ho et al. [10] found higher quality of work results if they incentivize a worker based on the quality of their work using a comparison with the work done by other workers. One challenge with this approach is cost, both because of the incentives and because tasks must be done by multiple users. Alongside variations in incentive, researchers have also found that expert feedback, self assessment [6], and explanations of task purpose [5] all improve task quality.

One challenge with evaluating worker quality on a crowdsourcing platform is user group control [19, 20]. Rogstadius et al. [20] addressed this issue by recruiting all participants on a paid crowdsourcing platform (Amazon Mechanical Turk) and then dividing participants into two groups. Using deception, one group was told the task was requested by a non-profit organization to help others and the other group was told it was requested by a for-profit company. They found the group that believed they worked for the non-profit organization did higher quality work.

The most relevant related work to our current examination of PledgeWork involve contrasting volunteer workers with paid workers. In this area, Mao et al. [16] re-post a task completed on a citizen science platform (Planet Hunters) to a paid crowdsourcing platform (Amazon Mechanical Turk) and found, anecdotally, that the work of the two groups was of similar quality. In this same domain, Borromeo et al. [2] conducted a controlled study contrasting quality. Specifically, they created two sets of tasks on a paid crowdsourcing platform (CrowdFlower) and an open source crowdsourcing platform (PyBossa). Although they found higher quality from volunteers, they note that the paid platform data included SPAM responses. It is unclear if the quality difference would exist if they used a crowdsourcing platform with quality control.

Our work differs from the above work because the task workers do via PledgeWork may be completely unrelated to the charity being supported. Consider Rogstadius et al. [20], who examine whether workers do better work for a non-profit or for a for-profit enterprise when given a task

requested by the organization, or Mao et al. [16] who analyze citizen scientist versus paid crowdworkers. In each case, there exist only two parties, the task requester and the crowdworker. The crowdworker is either paid or unpaid (in the case of Rogstadius et al.) and the requester influences quality of the work. However, it is not clear if the goal of the requester (scientific research / saving lives by improving health) or the task itself (involvement in the process of achieving the goal) drove the volunteer workers. Overall, issues of quality, quantity, and perception of workers who work to donate time to others on tasks unrelated to the cause they support is something unexplored in the current literature.

3 PLEDGEWORK LOGISTICS

The PledgeWork platform is conceptualized as an on-line system similar to a conventional crowdsourcing system. First, charitable organizations register themselves with PledgeWork. Then the task requester posts their task request and deposits the cost for the task. Task requesters can choose the charities to donate their task cost to, or leave this choice to PledgeWorkers. It might often be the case that requesters are willing to accept both PledgeWork and paid work; in this case, requesters may also choose to pay more than the normal cost of the task for PledgeWork, similar to a matching donation scheme. The volunteers choose tasks and, if not pre-specified, the charity they wish to support. Volunteers have the option to either reveal their information to the platform and/or charity or to contribute to a charity anonymously. As with typical crowdwork, once the volunteer completes the task, the requester checks the task results, confirms the task results match their criteria, and, instead of paying the workers, approves a donation to the charity specified.

PledgeWork results in some (relatively minor) changes compared to conventional crowdsourcing platforms. First, because the volunteer workers do not receive monetary compensation, no online transaction is required between the workers and the crowdsourcing platform. This makes the registration process for the volunteer workers simpler, potentially even non-existent if the workers choose anonymity. For task requesters, the complexity of registering a task should be equivalent to conventional crowdsourcing. For charitable organizations, registration with PledgeWork is similar to the current registration process for existing online fundraising platforms. A PledgeWork platform handles the transaction and provides tax information for the charities who receive money. Again, as with real world contributions, the charities then handle relevant tax-exempt issues through their usual channels.

4 UNDERSTANDING PLEDGEWORK

While PledgeWork has the potential to benefit charities, PledgeWork raises questions regarding crowdworker behavior. These questions include questions on the quality of work (e.g., Is there any discrepancy between the care taken by paid and charity workers?), the quantity of work (Is one group or another more motivated to do more work?, and the overall perspective of workers when working for a cause (Is PledgeWork something they would consider doing regularly?). We believe that these questions on worker behavior are critical questions to answer when studying the potential for PledgeWork, and, as noted above, we are unaware of any work that speaks specifically to this domain of ‘donation-by-proxy’.

To evaluate the acceptability towards our PledgeWork framework and understand the behavior of the potential PledgeWorker, we conducted a mixed methods user study. The first part of the study simulates PledgeWork versus paid crowdwork (henceforth charity versus paid conditions) in-lab using two online environments simulating charity versus paid crowdsourcing. Participants completed two crowdwork simulated tasks; audio annotation and text editing using a 2×2 mixed experimental design with TASK (AUDIO / TEXT) as within subjects factors and CAUSE (CHARITY vs. PAID) as a between subjects factor. This lab-based study was followed by an interview and questionnaires capturing the PledgeWork experience. The order of audio annotation task and text editing task was counter-balanced. As quantitative measurements, we extracted *total task time* and *completed task count*. We used the same questionnaires and interview topics for both participant groups, charity and paid.

Ethical Considerations and Deception

Ideally, to evaluate PledgeWork, one would conduct a study ‘in-the-wild’, but there are pragmatic and ethical considerations that limit our ability to do so. First of all, unfortunately, we as a task requester in our experiment could not directly contribute to charity due to the status of our institution as a non-profit. Taxation requirements within our jurisdiction prevent non-profit charities from making charitable contributions to a third party. Given these constraints, we were forced to deceive participants into believing that they were working for charities. The need for deception further restricted us to lab-based studies due to ethical requirements for debriefing post-study when deception is revealed and explained to participants. Our approach to deception is similar to that used in the Rogstadius et al.’ study [20], and this requires in-person follow up with the participants to explain the rationale for deception alongside the existence of deception.

Study Design

Our study design is similar to Rogstadius et al.'s [20] with some modification. First, the primary difference is the separation of the motivation (for charity/non-profit) and the task requester. In Rogstadius et al.'s study, participants completed a task requested by the charity. In our study, the task is independent from the charity, and only the payment goes to the charity.

Alongside this difference, Rogstadius et al. assigned participants to non-profit and for-profit conditions randomly. In contrast, as we will note below, we gave participants a choice to either contribute to a charity or receive money themselves. This experimental design more closely mimics the overall conceptualization of PledgeWork, i.e. we believe it is more ecological.

Tasks

To partially control for the nature of the crowdsourcing task, we designed two tasks with different complexity; an audio annotation task (low complexity) and a text editing task (high complexity).

Audio Annotation. Audio annotation is a task where participants label areas of an audio recording. For this task, we used audio recordings from a conversation in an English lesson. The conversation is between one male teacher and six female students in an English class, and we cut the audio record into 50 short clips. Each audio clip is between 10 - 30 seconds long and includes more than 1 utterance.

The participants were asked to select both the speaker (teacher / student) label and conversation between (teacher-student / student-student / unknown) label for all of the areas they specified. To allow labeling, we used the open-source audio annotation interface [4] with small modifications. In our audio annotation interface, the audio spectrum was visualized as time series data and the user selected an area by dragging a mouse in the visualization window. The participants could play the audio clip from any time and as many times as they wished.

Although the task itself is simple and does not require any skill other than normal computer mouse operation, to label the area precisely the workers must pay attention to the audio and must play the audio multiple times.

Once the participants finished annotating each audio clip, they were instructed to hit the “submit” button to confirm they finished annotating the file. Every time they submitted the annotated file, the system asked if they wanted to continue to the next audio annotation task or not. Participants could continue the audio annotation task as long as they wished to a maximum of 45 minutes total or until they finished annotating 50 audio clips. They were told they could

stop the task any time they wished. At the top of the experiment window, the amount of bonus / donation was shown.

Text editing task. For the text editing task, we used the Cambridge Learner Corpus (CLC) [26]. CLC is a collection of short essays written by English language learners. We chose 25 essays that contained at least 12 errors ($M = 20.5$, $SD = 6.5$). The length of the essays ranged from 139 to 240 words ($M = 189.7$, $SD = 28.6$).

For the text editing task interface, we developed a web interface using the online text editor TinyMCE¹. The interface had a main editing window on the left and comment window on the right. We asked the participants to find all grammatical errors in the essay shown in the main editing window and fix the errors directly in the essay. Participants were also instructed that they could make comments in the comment window which might be useful for the students, but comments were optional.

Similar to the audio annotation task, the participants could continue the task as long as they wished to a maximum of 45 minutes total or until they finished editing 25 essays.

Participants

We recruited 28 participants (18 female and 10 male) ages between 18-32 ($M = 22.0$, $SD = 3.1$). All participants were recruited through our institutional participants pool or posters on campus. All participants were affiliated with our institution and all but one were students. The lone exception was a recent graduate.

We divided participants into charity and paid groups as follows: In the recruitment process, the participants were given three options; 1. Crowdfund design study 2. Crowdfund for the United Nations Children's Fund (UNICEF) or 3. Crowdfund for the Office of the United Nations High Commissioner for Refugees, i.e. the United Nations Refugees' Agency (UNHCR). With all three options, participants were informed that tasks included audio annotation and text editing. If they were interested in contributing to UNICEF or UNHCR, they could choose option 2 or 3; otherwise, they could choose option 1. These options were confirmed at the beginning of the study. Fourteen participants chose option 1, nine participants chose option 2 and five participants chose option 3. Option 1 is considered the PAID group, and options 2 and 3 are considered the CHARITY group.

Rewards and Deception Logistics

All participants received \$15 as a base remuneration for their participation. This fee was to incentivize recruitment.

To simulate paid and charity crowdsourcing, we used slightly different reward strategies for the PAID versus CHARITY groups. For the PAID group, we explained that they would

¹<https://www.tiny.cloud/>

receive a monetary bonus as they complete the tasks. For the CHARITY group, we explained that we would make a donation to the charity they chose (UNICEF or UNHCR) based on the number of tasks they completed. Rewards were described after participants chose their respective group. In each case, the reward rate was 20 cents per audio file for the audio annotation task and 30 cents per short essay for the text editing task.

Although we told participants that we would donate to a charity, because our institution is not allowed to make donations to other charitable organizations, at the end of the study, all participants received a monetary bonus based on the number of tasks completed. Deception was not revealed until after all tasks and data collection were completed to avoid biasing results. Participants did not react negatively to the presence of deception: once the purpose was explained to them, they understood the rationale for deception and remained supportive of the premise of PledgeWork.

Apparatus

Our experiment room was a small office with a desk and a chair. It contained a 27-inch computer monitor and USB keyboard and mouse connected to a notebook computer placed on the desk. This environment simulates a general home office environment.

We used a web interface to support the paid and charity crowdwork environment. Our experimental system ran on our web server and participants accessed the web server using a Chrome browser on the computer. All user activities were logged for subsequent analysis.

Study Procedure

The study proceeded as follows.

- Participants were greeted and consent was obtained. Participants were told we were studying crowdwork interface design, and that they would be paid \$15 for their participation in the study. They will edit text, annotate audio and fill a questionnaire followed by an interview. The study takes up to 2 hours with duration varying depending how long they decided to spend on the task part. As an option, they could choose to make a contribution to a charity through the study. We did not explain the bonus payment nor the detail of the contribution they would make before they chose the option.
- After the participants chose the charity or paid option, the experimenter explained the bonus payment for the PAID group and the donation payment for the CHARITY group.
- An interface, designed to deceive the participant into believing they were really performing paid or charitable crowdwork, was presented to and explained to participants.

- During this presentation of the interface, the experimenter demonstrated all of the operations with the task interface. For example, for the audio annotation task, this demonstration included selecting an area, area modification, playing and stopping the audio, playing the audio from a given point, deleting the selected area, selecting the label, changing the label, selecting and deselecting the area. For the audio annotation task, we also explained that, if one person was continuously talking in the audio clip, each sentence was suppose to be labeled individually. If multiple people were talking at the same time, including acknowledgements (phatic utterances) such as “Uh-huh” or “yes”, those multiple utterances should be labeled individually. We also demonstrated creating over-lapped labels for these utterances. We asked the participants to annotate the audio clip as precisely as possible. Similar to the audio annotation task, we demonstrated the text editing task using our editor interface. We asked the participants to fix all of the grammar errors in the essays.
- At the end of each task explanation, we emphasized the donation to the CHARITY group and the bonus rewards to the PAID group. We also noted that they could continue as long as they wished, and that they could stop any time they wanted.
- After the explanation, the experimenter left the room and the participants were instructed to notify the experimenter outside of the room once they finished the first assigned task. They then switched to the second task. As noted above, these were counterbalanced across participants.
- After the audio annotation and text editing task, we administered a questionnaire and conducted a semi-structured interview.
- At the end of the study, we conducted a post-deception debriefing. Regardless of the condition, all participants received the bonus based on the number of tasks they completed. All participants agreed to allow their data to be used and post-deception consent was obtained.

Data Collection

All user activities in our prototype interface were logged. Log data included output from the task, timing information, and task completion information.

Our questionnaire, administered post-experiment includes three 7-point Likert questions asking:

1. Do you wish you could do more volunteering than you do now?
2. Do you wish you could donate more to charities than you do now?
3. If the PledgeWork platform was broadly available, would you use it to contribute the charity of your choice?

Following these three Likert questions, the questionnaire included three additional open-ended questions:

- What kind of tasks are you willing to do as a volunteer?
- How long are you willing to spend on a task resulting in a \$5 donation to a charity?
- How long would you be willing to spend on PledgeWork per month?

Following the questionnaire, we conducted a semi-structured interview. In the interview, we asked participants about their current volunteering and monetary donation experience as well as any barriers they experienced for volunteering/donation with/to charities.

5 QUANTITATIVE RESULTS

Three participants had software problems in the audio annotation task and one participant reported he fell asleep during the audio annotation session, so we excluded these 4 participants' data from our quantitative analysis. As a result, we have 24 participants, 12 participants each for charity and paid groups.

Quantity of work. Recall that participants could continue the task as long as they wished. To analyze the quantity of work performed for each participant group, we analyzed *completed task count* and *total task time* with CAUSE and TASK as independent variables.

A repeated measured two-way ANOVA of TASK and CAUSE for the *completed task count* and *total task time* found a main effect for CAUSE ($F_{1,22} = 9.487, p < .01$) indicating a significant difference between the CHARITY group ($M = 9.6, SD = 7.4$) and PAID group ($M = 21.3, SD = 15.5$) for the *completed task count*. The main effect for TASK ($F_{1,22} = 32.599, p < .001$) was significant on AUDIO ($M = 21.0, SD = 15.8$) and TEXT ($M = 9.9, SD = 7.3$). The interaction between CAUSE and TASK was also significant ($F_{1,22} = 11.853, p < .01$). Post hoc t-tests indicate the PAID group had significantly higher *completed task count* on the AUDIO task ($M = 30.2, SD = 15.6$) than the CHARITY group ($M = 11.8, SD = 8.0$) ($F_{1,44} = 18.6, p < .001$) and the PAID group had higher *completed task count* on the AUDIO task than the TEXT task ($M = 12.3, SD = 7.6$) ($F_{1,22} = 41.9, p < .001$). Figure 1 shows *completed task count* for both TASK s. Synthesizing the results, overall PAID participants completed more tasks than CHARITY participants.

Considering overall task time and completion rate, 7 participants (5 from PAID group) continued the text editing task up to the maximum time allocation (45 minutes) and 2 participants (both from the PAID group) spent 45 minutes on the audio annotation task. 4 participants in the PAID group completed all 50 audio tasks, and no one in the CHARITY group completed all 50 audio tasks. One each in both groups completed the maximum number of text edit tasks (25 files).

In terms of time, the average total task time for the audio annotation task was 1345.8 sec ($SD = 751.8$) for the CHARITY group and 1772.0 sec ($SD = 448.2$) for the PAID group. Text editing times were 1601.8 sec ($SD = 701.0$) for the CHARITY group and 2021.3 sec ($SD = 603.0$) for the PAID group. Although the PAID group completed significantly more tasks than the CHARITY group, there was no statistically significant difference in the total task time. More specifically, while the PAID group completed more tasks, the CHARITY group spent more time for individual tasks compared to the PAID group.

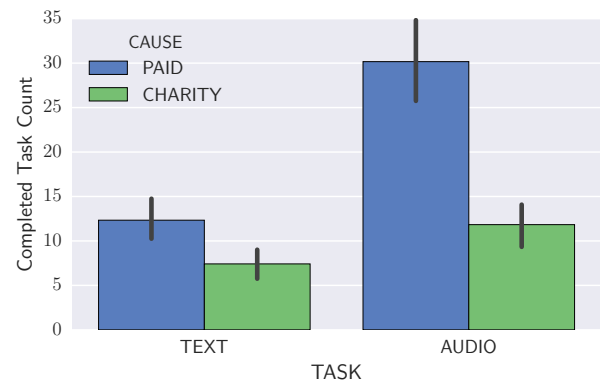


Figure 1: Average completed task count per participant. Error bars show standard error.

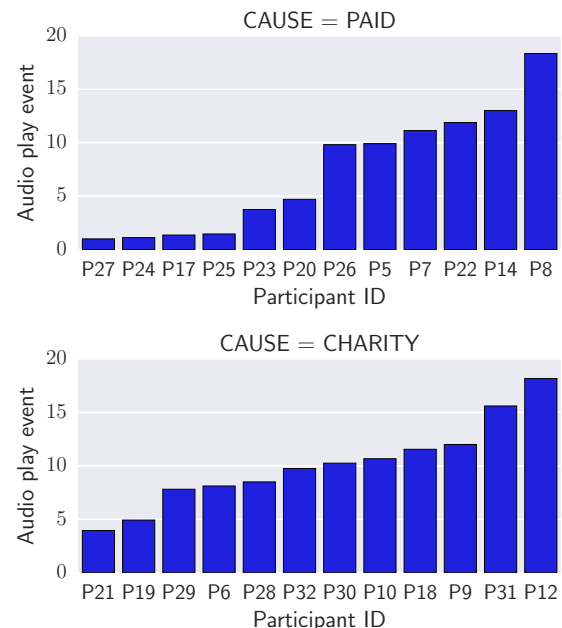


Figure 2: Average number of audio play events per audio file for each participants. Top: PAID group, Bottom: CHARITY group

Quality of Work. There are two possible reasons for a discrepancy in task count but not in task time. It may be that PAID participants were more efficient than CHARITY participants, completing more work in the allocated time. It may also be the case that PAID participants exhibit less care than CHARITY participants.

Unfortunately, it is difficult to extract direct measures of work quality. However, to indirectly compare the quality of the work, we examined the task log in detail. Our audio annotation interface allows users to play the audio file repeatedly or to start and stop during playback to position audio. In a pilot study, we observed both of these strategies for the audio annotation task, i.e. “play all and modify later through playing”, or “Step by step”. With the first strategy, participants played through the file selecting areas, then replayed to place labels and modify the selection, potentially playing a third time to verify and add additional labels. With the step-by-step approach, we also saw multiple “play” events but fewer “modification” events: participants repeatedly “play” and “stop” to finish individual labeling before moving to the next utterance. Since all of our audio files include more than one utterance, it is necessary to press the “play” button more than once per file to select and label the proper area. As a result, if participants only press play once and attempt to label the entire file without stopping the audio, labeling will, of necessity, be poorly done.

To examine the level of the care towards the audio annotation task, we calculated the proportion of tasks with 1 or less audio play events in the completed task count (*low quality task rate*) as well as the *number of annotations* made by participants. The *low quality task rate* was not normally distributed, so a Kruskal-Wallis test was performed to test the effect of CHARITY on *low quality task rate*. Although no significant difference on the average *number of annotations* was found, we found the PAID group had a significantly higher *low quality task rate* ($M = 0.335$, $SD = 0.392$) than the CHARITY group ($M = 0.029$, $SD = 0.069$) ($\chi^2(1) = 4.12$, $p < .05$).

Some paid crowdworkers do high quality work, so, to examine this factor in more detail, we examine individual participant data. Figure 2 shows the average count of audio play events per participant for both groups. As we can see from the graph (top), four participants from the PAID group played the audio file only (on average) once per file whereas all of the participants from the CHARITY group played the audio files at least 4 times in average.

In any realistic crowdsourcing environment, there are workers who perform poorly by trying to perform as many tasks as possible regardless of quality (hence the significant body of work cited earlier that seeks to improve crowdwork quality). With PledgeWork, while high quality crowdwork was possible from individuals in both groups, what we saw was that our CHARITY group had fewer poor performing

crowdworkers than our paid group. Overall 52.2% (189 out of 360) audio tasks done by the PAID group were of low quality whereas only 4.9% (7 out of 142) audio tasks done by the CHARITY group were of low quality.

We performed a similar analysis for the text editing task, counting the number of changed words in the modified essay. We did not find a significant difference between CHARITY and PAID groups. However, as noted above, participants completed fewer of the text editing tasks, making it less likely we would see significant effects ($M = 9.9$ vs 21 for audio).

6 FREQUENCY AND QUALITATIVE DATA

As noted in the experimental design, after completing the audio annotation and text editing task, the participants completed a questionnaire and we conducted a semi-structured interview. Our questionnaire included both frequency and qualitative data that allows us to more fully characterize participants. Alongside this, we report on themes extracted from semi-structured interviews.

Questionnaire

Exploring participants attitudes toward charitable giving, we found that our participants were positively inclined toward the idea of charitable crowdwork. The majority of our participants wished to contribute more to charity than they currently contributed both as a volunteer donating time (24 out of 28) and financially (27 out of 28). The vast majority of our participants (25 out of 28) responded positively to PledgeWork, and 3 participants said they would definitely use PledgeWork to engage with a charity of their choice.

We also explored types of tasks that participants would be willing to perform via PledgeWork. The following percentage of our participants answered that they are willing to do each task as a volunteer through PledgeWork: proof reading: 75.0%, translation: 53.6%, image annotation: 50%, audio annotation: 46.4%, audio transcription: 39.3%, programming: 25%, and writing: 21.4%. Only one participant answered she is willing to do any task as a volunteer; all other participants had task biases, i.e. they limited the types of tasks they would do.

In our questionnaire, we selected an arbitrary donation level of \$5, and asked participants how long they would be willing to spend working for a \$5 donation to a charity of their choice. The median answer in terms of time versus a \$5 donation was 30 minutes (13 out of 28 participants answered 30 minutes for volunteering; 7 were willing to spend less than or equal to 20 minutes; 8 participants answered 40 minutes or more for a \$5 donation). The amount indicated, median \$5 per half hour, is below the prevailing minimum wage (\$14 per hour) in our jurisdiction. Our questionnaire also asked the amount of time per month participants were willing to spend on PledgeWork. Table 1 shows participant responses: The average time they answered was 245 minutes,

Table 1: Duration of the time the participants were willing to spend on PledgeWork per month

Preferred time (min)	Num of Participants
-30	2
30-60	3
60-120	9
120-180	5
180-240	2
240-300	4
300-	3

i.e. approximately 4 hours per month ($SD = 353$ minutes) with range [0.5, 6] hours.

Interview Data

Given the results of the questionnaire where most participants noted a desire to volunteer more – both in time and financially – in our interview, we examined participants' volunteering backgrounds in more depth. Our goal was to understand both the volunteering experience and the barriers to volunteering.

All of our participants had been involved in volunteering activities in the last 5 years. Many of the challenges noted by our participants revolved (as expected) around the time commitment of in-person volunteering. Examples of this were varied: in some cases (e.g. Habitat for Humanity), because of the organization cost and training cost (*P5*), volunteering requires a minimum time commitment or a commitment for repetitive volunteering. As well, if the volunteering activity requires people to be at a certain place at a certain time, this time might conflict with their own working / study schedule (all participants had concrete instances of this). Alongside timing, other logistics also limit participants' ability to volunteer. For example, the lack of a car (*P10*) can make it impossible to perform volunteer tasks. *P10* also mentioned that he/she is not always aware of volunteering opportunities available nearby.

Beyond time, the activity the charity needs may not always match the skill or interest of the volunteers. *P12* and *P13* mentioned that they do not participate in volunteering activities if the task itself is not interesting to them, even if they care about the cause. However, it is also the case that many participants noted that the experience (for resumé building) and the skills acquired as a volunteer can be valuable: *P5*, *P15* and *P21* mentioned that their volunteer experience enhanced their career opportunities and *P13* commented on skills acquired through volunteering as being something of value.

Alongside giving of time to a charity, we also explored donations, i.e. financial support of charitable causes. One aspect of financial donations is that they require less time.

Most of our participants (71%) answered that they had made a monetary donation in the last 5 years.

While several participants noted that one limitation on additional charitable donation was simply availability of funds, we identified four additional barriers to additional financial giving. These included:

1. The need for a credit card. While it is common to have a credit card, some participants were simply not comfortable using a credit card for online donation to a charity. This might be a result of a general reluctance to use a credit card online (*P22*), or a simple reluctance to not engage with new or unknown financial websites.
2. Lack of trust online. Alongside a reluctance to provide financial information online, some participants expressed distrust of charities that ask for online contributions. As one example, *P30* was suspicious that the donation might not be used as described by the charity; recent instances of charity fraud arouse suspicion, and the ease of setting up misleading online sites and the impersonal nature of online donation increases this distrust. *P24* stated “*For online donation, I don't see the person. I must know the person who I'm giving my money to.*”
3. Lack of knowledge. Similar to the barrier for volunteering, participants pointed to a lack of awareness about charities. Some participants were aware of the existence of a charity but did not feel they had enough information about the charity (e.g., *P16*: “*I know the name of the charity but don't know what they are actually doing on a regular basis.*” Physically volunteering for a charity addresses this challenge because volunteers experience first-hand the charitable activities. Coupled with this lack of knowledge is a general suspicion of online sources, making it difficult for participants to inform themselves to a level they are comfortable with.
4. Lack of anonymity. One disadvantage of online payment is the lack of anonymity. Credit card information typically includes name, billing address, and billing phone number to allow processing of a transaction. While online donation may cause distrust that this information will be traded, even with assurance that it will not be, the charity will have your information. One participant noted that Wikipedia asks for donations every year and, after donating once, they remind users that: “*On [DATE], you donated \$[AMOUNT] to keep Wikipedia online for hundreds of millions of readers. I'm surprised by and deeply grateful for your continued support. We need your help again this year.*” Finally, even if one eliminates the risk of information sharing and the follow-up demands for support, it remains the case that charities will store this information. Whether charities are investing sufficiently in information security (and whether they should be investing heavily in security

given their desire to maximize financial support for their cause) is unclear.

7 ADVANTAGES AND DRAWBACKS OF PLEDGEWORK

PledgeWork exists at the boundary between volunteering time for a charity and donating money to a charity. Considering each of the stakeholders, the PledgeWorkers, the task requesters, and the charitable organization, PledgeWork provides important advantages for each.

For the volunteers

Many participants listed the lack of time as one barrier for engaging in volunteer work. Specifically, multiple participants pointed out that the time commitment can be a burden, especially because charitable causes often want repeat volunteering. PledgeWork increases flexibility and efficiency of the charitable work. In terms of flexibility, PledgeWorkers can work when they are available and at any location, rather than on a schedule and at a specific location. Additionally, the time required to travel to a location and to receive instructions to organize tasks increases the time taken by in-person volunteering.

Temporal cost and efficiency also play out in terms of simplifying on-boarding for volunteers. Volunteers can quickly contribute to a charity without the need for extensive training or commitment, meaning that volunteers can potentially be recruited more quickly and in larger numbers.

PledgeWork also addresses barriers to financial support of a charity. Through a study on medical crowdfunding, Kim et al. [14] emphasize the importance of varied volunteering options to contribute to charity by pointing out beneficiaries' close friends can feel pressure to donate more money than they afford. PledgeWork can provide an additional option to potential supporters rather than identifying money in their budget. It also can preserve their anonymity from the charity if they so choose.

The flexibility of task choice is another advantage of PledgeWork for volunteer workers, especially over offline volunteerism and online volunteerism (e.g. citizen science). With other types of volunteering, volunteer workers' tasks are specified by and then performed for the charity. However, the charity the volunteer worker wishes to support might not be able to best utilize their skills. With PledgeWork, the task can be either relevant to or completely removed from the charity. Volunteer workers can choose tasks that suit their skill set, thus allowing them to contribute more efficiently to the charity.

One challenge with PledgeWork for volunteers is that the social aspect of interaction with a charity is more limited. While many people become involved in charitable causes because those causes are important to them, it is also the

case that the social aspects of volunteering – the meeting of like-minded people and the camaraderie that develops around volunteering activities – are absent in the PledgeWork domain.

Advantages for Task Requesters

One factor we considered in analyzing the effect of PledgeWork was the overall behaviour of crowdworkers. While earlier work indicates that paid workers performed similarly to volunteer workers [16], the earlier contrast was between unpaid volunteers and workers working for money. Here, our comparison is slightly different – between paid workers and volunteers working for a monetary donation to be directed to a charity of their choosing – and our experiment suggests PledgeWorkers spend more time for each task and conduct the task with more care, by a statistically significant margin for our simple task.

Alongside benefits for work quality, many of the tasks on existing crowdsourcing platforms are relatively simple and can be done in a short period of time, typically three minutes with a rate of pay of several cents per task. Because of this, crowd workers, especially people who perform crowdwork for money, prefer task sets with a large number of tasks; as one masters the task, speed increases, meaning that large task sets are more efficient[15]. As a result, for a smaller set of tasks, the requester often has to pay more per task to recruit workers. PledgeWorkers, in contrast, might be willing to do smaller sets of tasks than traditional crowdworkers because the calculus of pay and loss differs. This, in turn, might make PledgeWork a more efficient option for a subset of crowdwork tasks that are more challenging to staff.

Advantages for Charities

In our interview, many of the participants mentioned they do volunteering to give back to their own community and to develop friendships and connections. Because volunteering is a social activity, the volunteer workers connect with people in the community, not just the cause. Although interpersonal bonding can improve retention rate, it can also unintentionally exclude new volunteers. With PledgeWork, the volunteer workers can help the charity they care about without the need to integrate into a new, potentially closed, community. PledgeWork can also recruit new volunteers or donor groups. Alongside this, we also note that PledgeWork can draw people to charitable contribution, serving as a gateway to more involvement, a gradual immersion within the community.

Through PledgeWork, charities receive a monetary donation. Monetary donations can allow charities to hire professionals as opposed to inexperienced volunteers for skilled tasks (e.g. Habitat for Humanity). As well, some charities

can buy more efficiently from wholesalers if provided with monetary donations [23] (e.g. Food Banks).

PledgeWork and Ambiguous Altruism

While PledgeWork has many positive attributes, there are also potential negatives. Many researchers have studied the economics of crowdwork [11, 12, 21]. The take-away from many of these studies is that pay is low [1, 11], but, despite low pay, there exist a group of crowdworkers who can be characterized as “professional” crowdworkers [1, 21], leveraging crowdwork as a means of supporting themselves. PledgeWork confounds this calculus. While the benefit is that workers can donate to charities, the liability is that PledgeWorkers will begin to consume work that would otherwise go to economically struggling individuals who rely on crowdwork to supplement their income. If a new source of crowdworkers allows task requesters to avoid paying workers living wages through crowdwork, the balance between altruism in support of worthy causes and the misanthropy of downward wage pressure or lower availability of paid work becomes unclear.

8 LIMITATIONS

Due to ethical requirements for the deception in our study, we were required to leverage our own, simulated platform, and our study was, of necessity, localized to our geographic region and had to be conducted in person with a limited number of participants compared to typical crowdsourcing studies. The limitations associated with our in-lab study are following:

- While it is true that inability to deploy on a crowdsourcing platform is a liability, it is also the case that direct contact with our participants allowed us to conduct semi-structured interviews that further triangulated our data. In other words, despite loss in ecological fidelity, we gain in data saturation and triangulation.
- While geographical limitations on our participants means that we cannot – unfortunately – generalize our results to all crowdworkers everywhere, that does not negate the fact that there exists a group of potential charitable volunteers who would be interested in supporting worthwhile causes through charities in, at the very least, our geographic region. More importantly, even though our work is not universally generalizable, it is probable that there exists demographics that mimic the demographics we leverage in our work. Hence, this work retains its overall utility.
- In-person contact with participants might have made them more amenable to charitable giving, whereas doing crowdwork out-of-sight might have made them more likely to

seek remuneration. Countering this, we deceived the participants by stating that we were examining interfaces for crowdwork and that they had a choice of whether they wished to work on the crowdwork platform or the PledgeWork platform. We noted the interfaces were similar. The concept of PledgeWork was not explored until after participants had completed their task.

9 CONCLUSION

In this paper, we explored the concept of PledgeWork: the online donation to charitable causes through crowdwork. We discussed the differences between PledgeWork and other volunteer crowdwork previously explored in the literature. In particular, we note that PledgeWork, which involves a donation via a proxy task unrelated to the charity, is a unique space. Given this unexplored style of crowdwork, we evaluated the behaviour and acceptability of PledgeWork as a form of charitable giving. Overall, our data showed that PledgeWork can address many difficulties people have supporting worthwhile causes. As well, it may improve the quality of crowdwork for task requesters, thus serving as another mechanism for guarding against poor-performing crowdworkers.

ACKNOWLEDGEMENTS

We thank participants in our study and the Office of Research Ethics at the University of Waterloo for their assistance with study design. This work was made possible by funding from NSERC Discovery Grant (#2018-05269).

REFERENCES

- [1] Janine Berg. 2015. Income support for the unemployed and the poor. In *Labour Markets, Institutions and Inequality*. Edward Elgar Publishing, Chapter 10, 263–286. https://ideas.repec.org/h/elg/eechap/16143_10.html
- [2] Ria Mae Borromeo and Motomichi Toyama. 2016. An Investigation of Unpaid Crowdsourcing. *Hum.-centric Comput. Inf. Sci.* 6, 1 (12 2016), 68:1–68:19. <https://doi.org/10.1186/s13673-016-0068-z>
- [3] Michael Buhrmester, Tracy Kwang, and Samuel D. Gosling. 2011. Amazon’s Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science* 6, 1 (1 2011), 3–5. <https://doi.org/10.1177/1745691610393980>
- [4] Mark Cartwright, Ayanna Seals, Justin Salamon, Alex Williams, Stefanie Mikloska, Duncan Macconnell, Edith Law, Juan P Bello, and Oded Nov. 2017. Seeing Sound: Investigating the Effects of Visualizations and Complexity on Crowdsourced Audio Annotations. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017). <https://doi.org/10.1145/3134664>
- [5] Dana Chandler and Adam Kapelner. 2013. Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization* 90 (6 2013), 123–133. <https://doi.org/10.1016/j.jebo.2013.03.003>
- [6] Steven Dow, Anand Kulkarni, Scott Klemmer, and Björn Hartmann. 2012. Shepherding the Crowd Yields Better Work. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW ’12)*. ACM, New York, NY, USA, 1013–1022. <https://doi.org/10.1145/2145204.2145355>

- [7] Debra A Fischer, Megan E Schwamb, Kevin Schawinski, Chris Lintott, John Brewer, Matt Giguere, Stuart Lynn, Michael Parrish, Thibault Sartori, Robert Simpson, Arfon Smith, Julien Spronck, Natalie Batalha, Jason Rowe, Jon Jenkins, Steve Bryson, Andrej Prsa, Peter Tenenbaum, Justin Crepp, Tim Morton, Andrew Howard, Michele Belevu, Zachary Kaplan, Nick VanNispen, Charlie Sharzer, Justin DeFouw, Agnieszka Hajduk, Joe P Neal, Adam Nemecek, Nadine Schuepbach, and Valerij Zimmermann. 2012. Planet Hunters: the first two planet candidates identified by the public using the Kepler public archive data. *Monthly Notices of the Royal Astronomical Society* 419, 4 (2012), 2900–2911. <https://doi.org/10.1111/j.1365-2966.2011.19932.x>
- [8] Jessica Heinzelman and Carol Waters. 2010. *Crowdsourcing Crisis Information in Disaster-affected Haiti*. U.S. Institute of Peace. <https://books.google.ca/books?id=DlyhnQEACAAJ>
- [9] Christian Heipke. 2010. Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 6 (11 2010), 550–557. <https://doi.org/10.1016/j.isprsjprs.2010.06.005>
- [10] Chien-Ju Ho, Aleksandrs Slivkins, Siddharth Suri, and Jennifer Wortman Vaughan. 2015. Incentivizing High Quality Crowdwork. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 419–429. <https://doi.org/10.1145/2736277.2741102>
- [11] John Joseph Horton and Lydia B Chilton. 2010. The Labor Economics of Paid Crowdsourcing. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC '10)*. ACM, New York, NY, USA, 209–218. <https://doi.org/10.1145/1807342.1807376>
- [12] Nicolas Kaufmann, Thimo Schulze, and Daniel Veit. 2011. More than fun and money. Worker Motivation in Crowdsourcing-A Study on Mechanical Turk. In *Americas Conference on Information Systems*.
- [13] Firas Khatib, Seth Cooper, Michael D Tyka, Kefan Xu, Ilya Makedon, Zoran Popović, David Baker, and Foldit Players. 2011. Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences* 108, 47 (2011), 18949–18953. <https://doi.org/10.1073/pnas.1115898108>
- [14] Jennifer G Kim, Kristen Vaccaro, Karrie Karahalios, and Hwajung Hong. 2017. "Not by Money Alone": Social Support Opportunities in Medical Crowdfunding Campaigns. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1997–2009. <https://doi.org/10.1145/2998181.2998245>
- [15] Walter S Lasecki, Jeffrey M Rzeszutarski, Adam Marcus, and Jeffrey P Bigham. 2015. The Effects of Sequence and Delay on Crowd Work. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1375–1378. <https://doi.org/10.1145/2702123.2702594>
- [16] Andrew Mao, Ece Kamar, Yiling Chen, Eric Horvitz, Megan E Schwamb, Chris J Lintott, and Arfon M Smith. 2013. Volunteering versus work for pay: Incentives and tradeoffs in crowdsourcing. In *In Proceedings of the First AAAI Conference on Human Computation and Crowdsourcing (HCOMP '13)*.
- [17] Winter Mason and Duncan J Watts. 2009. Financial Incentives and the "Performance of Crowds". In *Proceedings of the ACM SIGKDD Workshop on Human Computation (HCOMP '09)*. ACM, New York, NY, USA, 77–85. <https://doi.org/10.1145/1600150.1600175>
- [18] Aditi Misra, Aaron Gooze, Kari Watkins, Mariam Asad, and Christopher A. Le Dantec. 2014. Crowdsourcing and Its Application to Transportation Data Collection and Management. *Transportation Research Record: Journal of the Transportation Research Board* 2414, 1 (1 2014), 1–8. <https://doi.org/10.3141/2414-01>
- [19] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. 2010. Running Experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5, 5 (2010), 411–419.
- [20] Jakob Rogstadius, Vassilis Kostakos, Aniket Kittur, Boris Smus, Jim Laredo, and Maja Vukovic. 2011. An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets. In *ICWSM*.
- [21] Joel Ross, Lilly Irani, M Six Silberman, Andrew Zaldivar, and Bill Tomlinson. 2010. Who Are the Crowdworkers?: Shifting Demographics in Mechanical Turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM, New York, NY, USA, 2863–2872. <https://doi.org/10.1145/1753846.1753873>
- [22] Robert Simpson, Kevin R Page, and David De Roure. 2014. Zooniverse: Observing the World's Largest Citizen Science Platform. In *Proceedings of the 23rd International Conference on World Wide Web (WWW '14 Companion)*. ACM, New York, NY, USA, 1049–1054. <https://doi.org/10.1145/2567948.2579215>
- [23] Valerie Tarasuk and Joan M Eakin. 2005. Food assistance through "surplus" food: Insights from an ethnographic study of food bank work. *Agriculture and Human Values* 22, 2 (6 2005), 177–186. <https://doi.org/10.1007/s10460-004-8277-x>
- [24] Ramine Tinati, Max Van Kleek, Elena Simperl, Markus Luczak-Rösch, Robert Simpson, and Nigel Shadbolt. 2015. Designing for Citizen Data Analysis: A Cross-Sectional Case Study of a Multi-Domain Citizen Science Platform. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 4069–4078. <https://doi.org/10.1145/2702123.2702420>
- [25] Kyle W. Willett, Chris J. Lintott, Steven P. Bamford, Karen L. Masters, Brooke D. Simmons, Kevin R. V. Casteels, Edward M. Edmondson, Lucy F. Fortson, Sugata Kaviraj, William C. Keel, Thomas Melvin, Robert C. Nichol, M. Jordan Raddick, Kevin Schawinski, Robert J. Simpson, Ramin A. Skibba, Arfon M. Smith, and Daniel Thomas. 2013. Galaxy Zoo 2: detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society* 435, 4 (11 2013), 2835–2860. <https://doi.org/10.1093/mnras/stt1458>
- [26] Helen Yannakoudakis, Ted Briscoe, and Ben Medlock. 2011. A New Dataset and Method for Automatically Grading ESOL Texts. In *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (HLT '11)*. Association for Computational Linguistics, 180–189.