

Motivating Contribution in a Participatory Sensing System via Quid-Pro-Quo

Anthony Tomasic

John Zimmerman

Aaron Steinfeld

Yun Huang

{tomasic,johnz,astein,yunhuang}@cs.cmu.edu
Carnegie Mellon University

ABSTRACT

Participatory sensing systems (PSS) require frequent injection of information that has a short shelf-life. The use of crowds to gather information for PSS is therefore particularly challenging. In this study, we explore the impact of two policies on user contributions. A quid-pro-quo policy exchanges contributions from users for access to critical information in the system. A request policy simply reminds the user that information is needed to make the system function well. Prior research has shown that request for help in crowdsourced system is an effective mechanism to increase contributions. During a large-scale experimental study within a publicly deployed, crowdsourced, transit information system, we analyzed metrics associated with frequency of contribution and commitment to long-term use over a 10-month period. Our results confirmed that quid-pro-quo led to more contribution, but at a cost of faster departure from the study. When a participant was simply requested to contribute, but could still access community-generated data if they ignored a request, was largely ineffective and was statistically similar to the control condition where no request for contribution occurred. Thus crowdsource system designers should consider imposing quid-pro-quo type policies for PSS that concentrate on fewer users, but makes them more productive.

Author Keywords

Crowdsourced transit information systems; quid-pro-quo; participatory sensing

ACM Classification Keywords

H.5.3.

INTRODUCTION

The recent emergence of social computing along with the meteoric adoption of smartphones has given rise to the concept of participatory sensing [3]. This is a new embodiment for socio-technical systems that blends crowdsourcing and the ubiquity of mobile phones, positioning this as a sensing platform that can generate valuable information about the world. Examples of

participatory sensing systems (PSS) include the commercial service Waze [22], a mobile app that motivates car drivers to report traffic conditions that benefit others. On the research side, a great example can be found in Tiramisu [26], a smartphone service that allows transit riders to generate their own real-time arrival information system by sharing location traces as they commute across a city. One critical challenge PSS like Waze and Tiramisu face is how to motivate users to continually observe and contribute information. Part of the challenge comes from the fact that in the case of many PSS, such as Waze and Tiramisu, the contributed information does not directly benefit the person who makes the contribution. For example, by the time a user can report a traffic breakdown, it is generally too late for them to avoid the problem spot. Additionally, observations such as emergent traffic conditions and transit vehicle locations are time and place sensitive. They have a very short shelf life, and they only benefit a small segment of a larger user community who need the information in the next few minutes.

Understanding why people contribute to online communities has been an ongoing topic of interest in CSCW. A recently book [15] details many strategies for motivating contributions, including asking people as well as intrinsic and extrinsic approaches. In addition, research on free-riding — people who use a service but do not contribute — in peer-to-peer systems [8] and with recommenders [7] have investigated the idea reciprocity, of requiring users to make a contribution in order to gain access to the service or information they desire. A question we chose to investigate is if this “quid pro quo” approach could be successful in the context of participatory sensing.

To investigate this question and to better understand the unrealized potential of PSS, we conducted a ten-month long field study. We deployed a PSS version of the Tiramisu app with three conditions. One condition followed a *quid pro quo* approach, requiring participants to share location traces in order to access arrival time information; one condition simply made a *request* for contributions to participants who engaged in free-riding; and the final condition functioned as a *control*, allowing participants to access arrival information regardless of whether they shared location traces when commuting. Users were randomly assigned to one of the three conditions upon download of the app.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CSCW'14, February 15–19, 2014, Baltimore, Maryland, USA.

Copyright © 2014 ACM 978-1-4503-2540-0/14/02...\$15.00.

[HTTP://DX.DOI.ORG/10.1145/2531602.2531705](http://dx.doi.org/10.1145/2531602.2531705)

Participants in the *quid pro quo* condition made significantly more contributions than the other participants. However, they also abandoned their use of the app at a higher rate than participants in the other conditions. On average, they used the app for fewer days. In addition, a few even publically complained about the quid-pro-quo design in reviews they made in the app stores. We observed no differences in contributions between the control and request conditions. Interestingly, we did observe that participants who shared location traces, regardless of condition, were more likely to share reports about their transit experience. This observation indicates a possible relationship between contribution and engagement with the service.

In this paper, we first summarize related research that motivated and informed the design of our study. Next, we describe the Tiramisu, the context for our study, and the three experimental conditions. We then report our findings and discuss what they mean and how they might aid in the design of a participatory sensing system.

RELATED WORK

Related work falls into two distinct categories: previous research on participatory sensing and crowdsourcing systems, and previous work on motivating contributions in online communities.

PPS and crowdsourcing Systems

Many smart city concepts have been developed based on the idea of urban infrastructures that are instrumented with robust sensors. While many of these concepts have yet to be realized due to challenges of costs and machine-learning, recently *participatory sensing* has emerged as an alternative way to developing these kinds of systems. Participatory sensing recognizes the many people out in the world with mobile phones as a new type of sensing infrastructure [3]. It combines human observers and their phones' sensing and communication abilities to create new services.

Waze [22], a system that crowdsources traffic information provides a great commercial example. This service encourages users to contribute observations of traffic as they drive. Users can attach clarifying details to specific geographic points so fellow users have insight on sources of congestion (e.g., "Construction in right lane"). Like many PSS, Waze employs a gamification approach. They motivate people to share reports by making the interaction feel like a game. Users earn points for each observation. The potential success of this participatory approach can partially be seen in Google's \$966 million purchase of Waze.

Researchers and developers have both recognized motivating participation as a key challenge to developing participatory sensing systems. One of the main challenges with PSS in an urban sensing context is that it requires very time-sensitive and place-sensitive contributions.

Often an observation has a short shelf life and benefits only a tiny segment of other users. Consider Tiramisu, the crowdsourced transit information system, as an example. A Tiramisu user shares a location trace when commuting. This live information is only good for the remainder of the trip this user is on. In addition, it only benefits other users who are considering riding this specific bus in the next few minutes. Most other crowdsourced information systems do not face this challenge. For example, if a user adds or updates a Wikipedia entry, the benefits of their effort can persist for days, weeks, months, or even years.

Researchers have suggested that one approach to addressing this problem might be to transition from a participatory sensing to opportunistic sensing framing [16]. The difference is that, with participatory sensing, the human explicitly makes an observation. With opportunistic sensing, the mobile phone acts without requiring user input. Researchers have developed systems that employ this approach to generate transit arrival information similar to Tiramisu [20, 25]. This approach, however, raises two issues. One is privacy when a phone shares potentially personal information like location without explicit consent. Another is in the robustness of the information. For example, a phone can report its location on the bus, but it cannot easily report the route number when multiple routes share the same roadways. In this case the human observers ability to add information beyond what a phone can sense helps to demonstrate the potential of PSS.

Recent work in time-sensitive crowdsourcing has focused on the use of multiple, remote contributors (workers) who work together in support of an individual user (requestor). For example, Soylent utilized Mechanical Turk to provide on-demand, real-time document editing [2]. Another example, SCRIBE, also uses Mechanical Turk as a labor force and brings four Turkers together to transcribe live events in place of a traditional transcriptionist [17]. While these systems are successful and work well for their target applications, it is difficult to tell if they can scale to support millions of requestors. It is difficult to tell if there are enough workers to do all of the work people might request. Another key difference in the embodiment of these other time-sensitive crowdsourcing services is that they make a clear distinction between requestors and workers. In PSS systems the users that want the sensed information are often the same users who are in a position to make and share the observations.

Motivating contribution

Encouraging users to contribute to online systems has been an ongoing research topic within CSCW going back to early investigations of why workers failed to enter data into shared calendars [11]. Over the years, many studies have built up a rich set of theories designers can draw on in order to envision more effective, collaborative systems. A recent book by Kraut and Resnick does a wonderful job

of digesting this research and putting it into language that is actionable by designers [15]. Kraut and Resnick note that people are more likely to contribute when they are asked, and when they are asked personally, based on their individual knowledge and abilities. In addition, they detail a number of approaches that use intrinsic and extrinsic motivations to drive contribution. On the intrinsic side, systems can be designed to be fun or social, both of which has been found to increase contribution. On the extrinsic side, systems can offer rewards such as money, game points, or rewards such as reputation or status. They connect the results of many different studies around a model of how a user perceives the individual value they receive and the effort being asked of them. When the benefit (individual value) outweighs the effort, the user will contribute.

One ongoing challenge with social systems that require user contribution is free-riding, people that use a system but who never make a contribution. Adar and Huberman conducted one of the first studies on free-riding. They investigated usages patterns in the peer-to-peer (P2P) file sharing system Gnutella and found that 70% of users free-ride. They also reported that 25% of users provided 99% of the shared content and the top 1% provided almost 40% of the shared content [1]. More recent work looking at Gnutella reported even higher levels of free-riding, with more 85% of users exhibiting this behavior over different durations [12].

Some P2P systems have attempted to address the problem of free-riding by employing a reciprocity based incentive model where users are required to contribute in order to access the information they desire [14]. Many researchers have investigated BitTorrent's use of this incentive model [6, 14]. BitTorrent employs a "tit for tat" approach where the speed with which a user can download a file is controlled by their willingness to contribute, to allow others to download information from their own machine. When users do not share, they are choked out. This research generally describes BitTorrent's approach as effective, though a number of papers detail how users can cheat the system and engage in free-riding and propose variations to refine the incentive design, i.e. [13]. While "tit for tat" is an abstractly similar policy to quid-pro-quo, the BitTorrent user experience is completely different from a PSS. The quid-pro-quo policy of a PSS requires human effort to generate contributions with every use of the system. BitTorrent simply requires a user to agree (by default) to expend some computational resources.

Researchers working on recommender systems have also investigated reciprocity as a way to motivate contribution. It is important to note that the usage model of recommenders is a different than for PSS and P2P. With most recommenders, users need to enter information (make a contribution) before the recommender can make a recommendation. Many other social systems, such as

online dating services, follow this same usage model. With PSS and P2P, the data any user might want is often already available, so it is possible to gain the benefit of the system prior to making any contribution. While recommenders are different from PSS, they have a shared challenge of time-sensitive information. They need to generate ratings for new content in order to be able to recommend this content to others. Some of the recommender users must rate a new thing before any other ratings have been made. However, the shelf-life of these contributions made to recommenders are generally much longer than for PSS.

The need to make an initial contribution for receiving a service has been viewed as a barrier to adoption. Researchers looking at recommenders have investigated this issue. One very relevant study of the movie recommender MovieLens made a great advance on understanding the use of reciprocity during the initial use [7]. In this study, researchers created several versions of MovieLens: one required participants to enter in 25 tags in addition to an initial set of ratings; one required them to enter 5 additional tags; and one required no tagging, but allowed participants to tag if they desired. Participants who were required to enter more tags generated more tag information than participants in the other conditions. However, they also had a higher rate of abandoning the system during their initial use than participants in the other conditions. Nevertheless, even with the higher rate of abandonment, the high contribution requirement generated the most contributions even taking abandonment into account. Most interesting was that fact that participants who were required to enter the most information during their initial use of the system on average continued to enter more information during extended use of the system.

Our study advances understanding on contribution by applying current theories around motivating online contribution to a new domain: participatory sensing systems. We wanted to understand if these theories generated in experiments with systems like recommenders and wikis could work on this different kind of social computing system. In particular, because of the requirements of PSS, quid-pro-quo continuously monitors user behavior, unlike previous approaches that only set an initial barrier. In addition, we wanted to push of the ecological validity of these theories. While many of them had been developed and refined of lab studies and short-term field deployments, we intended to test them with a long-term (10-month) field study with a relatively large number of participants (8000+). A long-term field study exposes the research to additional effects, such as the impact of media announcements and product reviews.

STUDY

We wanted to investigate if a quid-pro-quo approach, which has been successful for driving contribution in P2P

services and for recommenders, could be effective at motivating high and sustained levels of contribution in the context of a PSS.

Study Context

For the context of our study, we selected the Tiramisu system, which allows transit riders to crowdsource arrival information by sharing location traces on a smart phone. In the current implementation, users begin by selecting a bus stop from a map. The service then displays a list of arrival times for all buses within a 90-minute window (those that should have passed 30 minutes ago to those that should arrive within 60 minutes) (Figure 1a). When another user is on a bus and contributing a location trace, the system provides a “real-time” arrival estimate to other users. When there is no real-time, but enough traces for this specific trip have been contributed in the past, the system provides a “historical” estimate of the arrival time. When there is neither real-time nor historical data of sufficient quality, the system provides “scheduled” arrival information from the transit agency’s static schedule.

To contribute a trace, users select the route number of the vehicle they board (Figure 1a) and identify their destination stop. This transitions them to the “Record” screen (Figure 1b). Here they indicate how full the bus is and then select the green “Start Recording” button to begin contributing a location trace (Figure 1b). When they are near their destination, the trace is automatically stopped, and the user is provided with a summary of their current and cumulative contributions. For more details on the design and rationale, please see [26].

While real-time arrival information systems are quite common in most European and many Asian cities, they are less available in the US. In the US only 5% percent of commuters use public transit to get to work [21]. Access

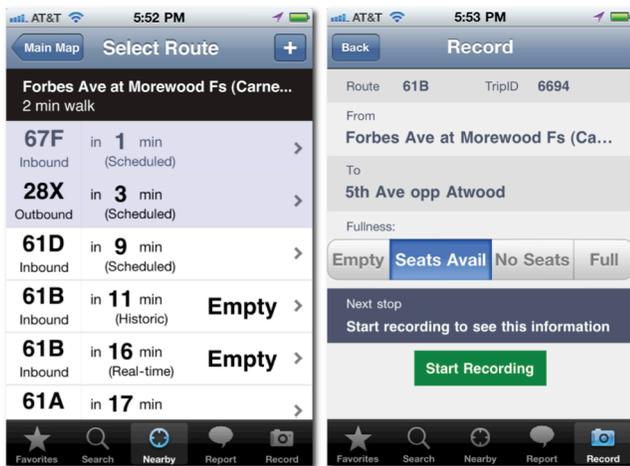


Figure 1. (a) Select Route Screen: users see three kinds of arrival information and select the vehicle they board. (b) Record Screen: users to report fullness and start sharing a location trace.

to arrival time estimates can increase ridership [5], particularly when delivered to mobile devices [9]. However, an embedded arrival information system including networked GPS units installed on all transit vehicles and electronic signage for display of the arrival information at the stops is expensive to purchase, install, and maintain. We suspect that the relatively small number of citizens who use public transit systems in the US make installation of an embedded real-time information system a challenging prospect for the taxpayers to accept. Given this situation, a PSS, where transit users co-produce the arrival information they desire, seems a possible solution.

In addition to being able to crowdsource arrival information, the Tiramisu app allows users to submit notes: questions, compliments, or complaints about the transit service. During a field trial of Tiramisu [26], researchers observed that usage of the app seemed to drive use of the notes feature. Most participants in the field trial reported that they had never contacted the transit service to share a compliment or complaint, yet when given this mobile service, more than half of the participants submitted a note about the transit service within the three-week trial. The research team did not know if this behavior was driven by simply having the phone in hand, with a transit app in focus, when experiencing a problem worthy of reporting, or if the act of sharing location traces (co-producing this new service) actually increased participants’ engagement with the transit service.

Conditions and Hypotheses

To run our study, we configured Tiramisu to run three experimental conditions parallel to its normal interaction design. Below we describe the conditions:

Request Condition (Req): When participants are observed to be free-riding, they are prompted via a popup message to contribute a location trace (Figure 2a).

Quid-pro-quo Condition (QPQ): When participants are observed to be free-riding, they are informed by a popup message that access to arrival information has been blocked (Figure 2b), and when they close the popup, they see only a list of upcoming buses and the word “Blocked” in place of the arrival time information (Figure 2c).

Control Condition: Follows the normal interaction design previously described where the system makes no attempt to detect free-riding behavior.

Our system detected free-riding by observing the number of sessions where participants accessed arrival information and the number of sessions where they shared a location trace. We used a point system. For each day a participant accessed the arrival information, they were awarded one *usage-point*. Regardless on how many times they accessed the arrival information, the most they could earn was one usage-point for a day. For each day where

they shared a location trace, they were awarded one *contribution-point*. Again, it did not matter how long they shared a location trace or how many times they shared location traces in a day. The maximum they could earn was one *contribution-point* per day. We used the ratio *contribution-point/usage-point* to infer free-riding behavior. When a new participant joined the study by downloading the app, they began with a default score of 1 *contribution-point* and 2 *usage-points*; a starting ratio of $\frac{1}{2}$.

For the QPQ and Req conditions, we added a high and low threshold for the free-riding detector. Participants in the high-level (QPQ-High and Req-High) were considered to be free-riding if they earned three usage-points and no contribution-points (three days of accessing information without contributing a location trace). Participants in the low-level (QPQ-Low and Req-Low) were considered to be free-riding if they earned six usage-points and no contribution-points (six days of accessing information without contributing a location trace).

After the system detected free-riding, it continued to accrue usage-points for each day participants accessed the system and did not share a location trace. When free-riding, if they shared a location trace, the points reset. This reset turned off the popups for both QPQ and Req participants and turned on the arrival information to participants in the QPQ condition. For participants in QPQ-High and Req-High, a reset changed *contribution-point* to 1 and *usage-point* to 3. For participants in QPQ-Low and Req-Low, a reset changed the contribution-point to 1 and usage-point to 6. These reset values meant that if the following day, the participant accessed the arrival information without sharing a location trace, the system would once again consider them to be free-riding. With this reset behavior, participants needed to only contribute once to be seen as not free-riding, regardless of the

number of days they had exhibited free-riding behavior. The actual total number of contributions and usage sessions for a device were recorded separately from this scoring system to aid in our analysis.

We had several hypotheses drawn from the literature:

- H1: Participants in QPQ will contribute more location traces than participants in the other two conditions. The success of BitTorrent [6] and the work on recommenders [7] made QPQ appear to be a more successful strategy that simply asking participants to make a contribution.
- H1.1: Participants in QPQ-High will submit more location traces than participants in QPQ-Low. The Drenner study [7] indicated that requiring more contribution should result in more total contributions, even after people quit using the service.
- H2: Participants in Req will share more location traces than participants in Control. Requesting contribution has been shown to be more effective than not requesting [15].
- H2.1: Participants in Req-High will share more location traces than participants in Req-Low. It seems logical that simply asking more frequently should generate more action from participants.
- H3: Participants in QPQ will stop using the app after fewer days than participants in Control and Req. The Drenner et al. study [7] observed that for their recommenders, the requirement to contribute caused people to abandon the service a higher rate during initial use than without the requirement.
- H3.1: Participants in QPQ-High will quit using the app at a higher rate than participants in QPQ-Low. Again,



Figure 2a. Popup message participant sees in the request condition.



Figure 2b. Popup message participant sees in the quid-pro-quo condition.



Figure 2c. Arrival information the seen in the quid-pro-quo condition.

the Drenner et al. study [7] noted that requiring more contribution cause more people to quit using their recommender during initial use.

- H4: Participants that share location traces will contribute more notes than participants that submit fewer location traces. We have two rationales for this. First, the Drenner et al. recommender study [7] showed that people who were required to contribute more during their initial sessions continued to contribute more. Second, results from the field trial of Tiramisu hinted that there might be a connection between sharing location traces and generation of notes.

Measures

Our study had 3 conditions and five possible configurations (Control, Req-High, Req-Low, QPQ-High, QPQ-Low). In the Control condition, there was no level as the system never attempted to detect free-riding and change its behavior.

For all configurations, we logged a series of device interaction measures. We used Device ID as a proxy for participant. When participants installed the app, the app would assign them to one of the conditions based on the Device ID. If they deleted and reinstalled the app, they were reenrolled in the same condition and recognized as the same participant. If they upgraded to a new phone during the course of this study, they appeared to be a new participant using the service for the first time. They could be assigned to any of the configurations.

Each device log record contained a GPS location and counts for the following: *Usage*: the number of days where the participant accessed information at least once. *Contribution*: the number of days where the participant shared at least on location trace. *Impressions*: the total count of schedule, historical and real-time *impressions* delivered. An impression is generated when a particular bus arrival time prediction, i.e. trip row, is displayed on the device (horizontal panels in Figure 1a). *Notes*: the total number of notes submitted from a device.

The log record also contained data about each device's configuration, the current contribution count, and a boolean value indicating if the device was considered to be free-riding. Additional data on other interactions within the app were also collected.

This data was used to compute a variety of summary statistics of each participant's use of the system. For example, the contribution percentage is the fraction of days that the participant shared a location traced out of all days the participant used the system.

Study Timeline

A deliberate effort was made to manage the mass media influence on this study so that external factors associated with public awareness would be somewhat controlled.

While never a problem in lab studies, these real-world factors can influence how people use a system in the wild. We deployed our system via Apple's App Store as a free app and held a press event on July 27, 2011 for the local papers, television news channels, and radio stations. The event was successful and resulted in very high saturation throughout the Pittsburgh media market. All major, local news outlets mentioned the application service. As desired, this saturation led to a dramatic surge in participant enrollment. We intentionally did not run a marketing campaign to increase participation and contribution rates. One exception to this was a press release we issued when an Android version became available (October 19, 2011).

External events occurred during the deployment that we assume could impact usage of the app. As with most agencies, Port Authority of Allegheny County (PAAC) changes their schedule periodically to adjust route timings, route offerings, and provide drivers a chance to change routes (i.e., a new "bid"). During the ten months of this deployment (August through May), the transit service released three new schedules. These occurred three months apart and are important in the system. The representation of the transit schedule complicates efforts to link data across bid changes. Each time a new schedule was released, the system discarded the historical estimates and used new location traces to generate new estimates.

In addition, during the initial release we encountered complaints that sharing a location trace led to unwanted battery drain. The software was update to reduce the location sampling rate, making use of the app much less battery draining. This change occurred during the second month of the deployment.

The study ended on May 15, 2012 when a new version of the software that removed the two experimental conditions, Req and QPQ. However, the system continued to operate, thus permitting tracking of the devices that continued to use the system (either under the old or new release). We continued to track devices after the study. This additional data permits a precise designation of the participants that continued to use the system after the new version of the software was deployed. This designation (called "right-censoring") allows us to properly account for participants who continued to use the service beyond the end of our data sample in survivability analyses.

RESULTS

8447 participants (smartphone device IDs) took part in our study over the 10-month duration. Results are organized below around the four hypotheses. Table 1 shows a summary of the results including the number of participants in each condition and level. Condition and level nested within condition were tested for significance. Participants in the QPQ condition had a contribution rate of 19.1%. This rate is the number of days they shared at

Table 1. Measurement of participant behavior by condition.

Condition	Level+	Total devices (<i>n</i>)	Total device days	Total trace days	Contribution %	Total trace duration (min)	Average trace duration (min)
Condition main effect results:			$F(2, 8444) = 7.1$ $p = 0.0008$	$F(2, 8444) = 8.5$ $p = 0.0002$	$F(2, 8444) = 10.7$ $p < 0.0001$	$F(2, 8444) = 4.7$ $p = 0.009$	$F(2, 8444) = 2.6$ $p = 0.075$ n.s.
Control			A++	A	A	A	
	Subtotal	2825	25124	6,017	15.5%	67,614	2.67
Request			AB	A	A	AB	
	High	1352	10,744	3,231	17.2%	40,654	3.78
	Low	1467	12,044	3,514	17.0%	42,721	3.55
	Subtotal	2819	22,788	6,745	17.1%	83,375	3.66
Quid-pro-quo			B	B	B	B	
	High	1379	9,858	4,938	19.9%	54,979	5.58
	Low	1424	11,076	4,101	18.2%	46,806	4.23
	Subtotal	2803	20,934	9,039	19.1%	101,785	4.86
Total		8447	68,846	21,801	17.2%	252,774	

+ There were no significant main effects for Level nested within Condition on the measures in this table.

++ Conditions not connected by letters (A, B, etc) are significantly different from each other for Tukey HSD post hoc test. For example, Control has significantly more Device Days than QPQ, but Req is not significantly different from either Control or QPQ.

least one location trace over the total number of days they used the app.

Participants in the Req condition had a contribution rate of 17.1% and Control had contribution of 15.5%. Tukey HSD post hoc analyses show that the QPQ mode was significantly different from the Control and from Req for contribution percentage and also for the total number of days where at least one location trace was shared. H1 was confirmed.

Participants in QPQ-High had a contribution rate of 19.9% while participants in QPQ-Low had a contribution rate of 18.2%. However, there was no significant effect for Level. Therefore, H1.1 was not confirmed. Requiring more contribution for participants in the QPQ-High configuration at twice the rate of contribution for participants in the QPQ-Low configuration did not result in significantly more contributions. H1.1 was not confirmed.

Participants in the Req condition had a higher rate of contribution than participants in Control; however, the effect was not significant. Asking participants to

contribute did not result in significantly more contributions than not asking. H2 was not confirmed.

Participants in Req-High had a 17.2% contribution rate and those in Req-Low had a 17.0% contribution rate. This effect was not significant. Prompting participants more frequently to contribute did not result in significantly more contributions. Therefore H2.1 was not confirmed.

Participants in QPQ used the app for an average duration 41.4 days (Table 2). Participants in Req had an average duration of 46.9 days and those in Control a duration of 50.3. To better tease out the details of duration we performed a parametric survival analysis using a Weibull distribution (Figure 3). This test showed significance for condition ($X^2 = 31.5$, $DF = 2$, $p < 0.0001$) and condition*Level ($X^2 = 2.69$, $DF = 1$, $p < 0.05$), but not Level ($X^2 = 7.8$, $DF = 2$, $p = 0.101$). Participants in QPQ dropped out 18% faster than participants in Control and 12% faster than Req. Therefore, QPQ participants left the study faster than those in the other conditions. This trend can also be seen in the number of participants who continued to use the system after the new release (“right-censored”), with QPQ retaining 176 censored devices while Req and Control retaining 233 and 259, respectively. H3 was confirmed.

Participants in QPQ-High stopped using the app after an average of 38.3 days while those in QPQ-Low stopped after an average of 44.3 days. Due to the mixed results for the parametric survival analyses and difficulty in running post hocs for this analysis method, we performed a regular survival analysis on Level for only the QPQ data. There was no difference between QPQ-Low and QPQ-High for a Log-Rank test ($X^2 = 2.4$, $DF = 1$, $p = 0.12$) but there was for a Wilcoxon test ($X^2 = 3.8$, $DF = 1$, $p < 0.05$). Given these mixed results and the large sample size in this study, H3.1 was not confirmed.

Table 2. Summary statistics for days in study.

Condition	Level	Days In Study		
		Mean	Median	St Dev
Control		50.3	23	62.1
Req	High	47.1	21	59.5
	Low	46.6	23	57.6
	Subtotal	46.9	22	58.5
QPQ	High	38.3	14	52.0
	Low	44.3	22	55.5
	Subtotal	41.4	18	53.9
Total		46.2	21	58.4

We conducted analysis of note submission and contribution of location traces. We wanted to understand if submission of notes was more related to simply having the phone in hand, such as accessing the arrival information, or if it was related to contribution and the co-producing of a service with a service provider. We discovered that the quantity of notes submitted was significantly correlated with contribution rate ($0.150, p < 0.0001$) and total number of share location traces ($0.353, p < 0.0001$). Participants who shared more location traces also submitted more notes about their transit experience. H4 was supported.

The mobile client application is available on iTunes and Google Play. These systems provide an opportunity for users to rate and comment on the application. We evaluated all 60 reviews for some mention of the fact that providing traces helps the crowd. For example, a 5-star review on July 28, 2011 begins *“Using the app seems straightforward. It gets me everything my usual approach does- with the promise of real time info when there are enough users. So c'mon everybody! ...”*. Of all the reviews, 24 contained some reference to the crowdsourced nature of the application. Some reviews however viewed the crowdsourced nature of the application as a negative feature, for example, a 3-star review on August 7, 2012 begins *“Because this app relies on riders using it to track the busses, it only works well when lots of people use it. On the lines that I ride here in Pittsburgh, it appears that no one else is using this app ..”*

Reviewers also commented on the quid-pro-quo policy. Two reviews explicitly reference this policy and complain about it, giving a very low rating to the mobile application. For example, a 1-star review dated April 7, 2012, *“... if I don't record my trip, times are blocked. Thanks for ruining my night out (and ride home). Guess what? I don't need ... App deleted.”*

The quantity of popup messages received by participants in the request and QPQ conditions varied. Analysis of the number of popup messages received by participants showed significant differences for Condition ($F(5620,1) = 142.3, p < 0.0001$) and Level ($F(5620,1) = 35.3, p < 0.0001$). Participants in Req received three times as many popup messages as participants in QPQ (means of 1.1 and 3.1, respectively). Participants in the High level received 1.6x more popup messages than those in the Low level (2.6 and 1.6, respectively).

DISCUSSION

The main question of our research was to investigate if reciprocity, which has to encourage contribution for recommenders and for P2P systems, could be used to effectively motivate contribution for a participatory sensing system. Contribution in a PSS is different than in these other social systems because of the time-place sensitive. PSS need near constant sense input from all points along the network because the shelf-life of any

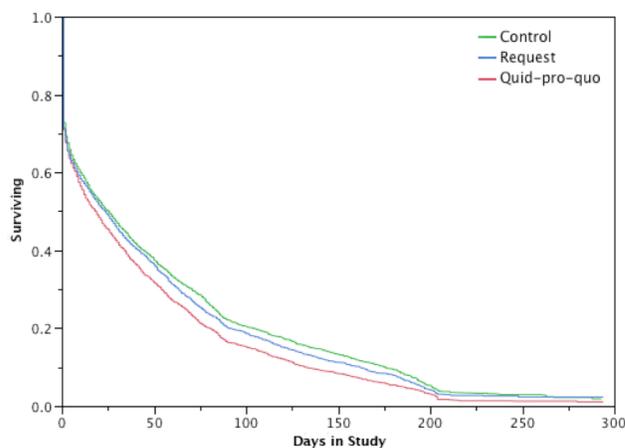


Figure 3. Survival by Condition.

contribution is often very, very short. Our first hypothesis, that participants in the QPQ condition would contribute more was confirmed. However, H3, our hypothesis that participants in the QPQ condition would abandon the app faster than participants in the other conditions also proved true. Together these make the choice to use a quid-pro-quo approach for a PSS more complicated.

If the goal of the system is simply to produce as much sense data as possible, then reciprocity seems like a good choice. Note that H1 automatically accounts for abandonment by participants, so QPQ produces more data despite the fact that participants quit using the system after fewer days. If a good user experience is also a critical goal, then reciprocity might not be the answer. This might particularly be the case where a PSS is simply one of many apps within a single product brand. The reviews participants submitted to the app stores also cloud this issue. The poor reviews clearly related to the reciprocity design might drive future users away, especially in cases where there are competing apps that follow a different motivational approach.

We caution PSS designers to carefully consider the specific goals of their system before committing to a specific motivational approach. Reciprocity seems a viable choice for generating more contribution, and in the case of Tiramisu, more data from more users produces more real-time arrival information, making the overall system more valuable to all users. At the same time, blocking negatively impacted user experience and might drive future users away.

Level had no effect in terms of contribution. Our hypothesis that requiring more contributions would increase contributions was not confirmed. One possible reason we observed no effect while [7] saw an effect could be a result of the different demands between QPQ-High and Low. In [7], participants had to contribute five-times as much information during their initial use, while in our study, participants in QPQ-High were only required

to provide twice as much as those in QPQ-Low. This issue deserves more investigation. We suggest a study with several different levels in order to better understand where the impact thresholds reside.

Req had no effect when compared to control. This result is counter-intuitive since previous research has shown that asking people to contribute will make a difference. We suspect this behavior depends on the specific embodiment of the request in our design. Participants in our study experienced this condition as a popup pushed to their phone. Participants may have been or became quite savvy at dismissing the popup without actually reading the text or without experiencing the popup as an actual request. This unexpected outcome also deserves more investigation. It would be helpful for designers to know if specific designs are more appropriate to specific motivational approaches. Other automated request forms, such as email, SMS, robocall, etc. may have had a very different effect than our design. This open research issue is part of the challenge designers face when attempting to move from theory to practice. They often have many different options within a design and no way to estimating the impact of a choice. There currently is no literature they can turn to that describes the best design patterns for working with the different motivational approaches.

In our study, participants in the Req condition received a lot of popups — three times as many as those in QPQ — yet they did not contribute more traces. This fact raises the issue that a Req approach used for a PSS could potentially have a negative impact on the user experience, but without the positive impact of producing more data. The lack of negative evidence experience in terms of higher abandonment between Req and Control does not mean that the experience of the two was equal.

The findings supporting H4 show a link between contribution of location traces and submission of notes about the transit service. The finding hints that there may indeed be a connection between the co-production of a service, in this case the co-production of the real-time information system, and increased engagement that drives people to submit more notes and to take more care and concern for their transit agency. From this study, it is impossible to tell if people who already have more engagement simply share traces and notes more frequently than other users or if the action of sharing traces or the action of submitting notes works to increase engagement. The issue requires a more carefully crafted study to tease out the underlying mechanism. The results should be of particular interest to the service research and service design communities who claim that co-production of service can be linked to customer engagement.

The reviews Tiramisu proved to be an unanticipated source of information on user experience. We had not considered these at the time we were designing the experiment. While an interesting source of qualitative

information, the reviews raise some concerns for conducting behavioral experiments in the wild. In our case, the reviews could easily reveal that the app had several conditions. For example, two of the reviews clearly talked about the quid-pro-quo design, and participants assigned to other conditions might realize that they are using a different version of the app. It is hard to tell what specific impact this might have, but it simply indicates another challenge of running studies in the wild that focus on ecological validity.

CONCLUSION

Changes in technology raise the possibility of creating participatory sensing systems as an alternative to fully embedded sensing systems. With participatory sensing, a community of people use their mobile phones to report conditions, thus producing a sensor network made up of people and phones. One key challenge in designing a PSS is motivating people to make contributions. A second challenge is the high rate of contribution that is needed. PSS are time and place sensitive. Generally each contribution only impacts a tiny section of the network for a very short amount of time. To build an effective PSS, designers need to figure out how to motivate users across the network to continually make contributions.

Reciprocity, where systems require users to make contributions in order to access the information they desire, has been shown to be effective for recommenders and for peer-to-peer file sharing systems. An open research question is if reciprocity might work for PSS. We conducted a 10-month field study to investigate this approach, by instrumenting Tiramisu, a participatory sensing transit information system. Our results show that reciprocity can help to increase contribution. However, the choice to use it is complicated by the fact that reciprocity also negatively impacts the user experience. Designers need to carefully consider the harmony they seek between collecting data and user experience when designing their systems.

This study also raised issues with the challenge of moving from theory to practice. Previous research has shown that asking people to contribute can increase contributions; however, in our study we observed no effect when we attempted to operationalize this previous finding. Additional research is clearly needed to better match specific interaction design patterns to behavioral theory so that designers can more effectively apply these theories and get the results they expect. Finally, the study showed an interesting interaction between the co-production of the real-time arrival service and participants' engagement with the transit service as evidenced through the increased submission of notes about the service. Again, more work is needed to better explicate this relationship.

ACKNOWLEDGEMENTS

This research and development are activities of the Rehabilitation Engineering Research Center on

Accessible Public Transportation (RERC-APT). The RERC-APT is funded by grant number H133E080019 from the United States Department of Education through the National Institute on Disability and Rehabilitation Research. Additional support was provided by Traffic21 at Carnegie Mellon University, a program developed with the support of the Hillman Foundation; a University Transportation Center grant (DTRT12-G-UTC11) from the US Department of Transportation.

REFERENCES

- Adar, E. and Huberman, B. A. Free Riding on Gnutella. *First Monday*, 5, 10 (2000).
- Bernstein, M., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D., & Panovich, K. Soylent: a word processor with a crowd inside. In *Proceedings of UIST*, (2010) ACM Press, 313-322.
- Burke, J. A., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S. & Srivastava, M. B. Participatory Sensing. In *Workshop on World-Sensor-Web*, 2006
- Burton, M. A., Brady, E., Brewer, R., Neylan, C., Bigham, J. P., & Hurst, A. Crowdsourcing subjective fashion advice using VizWiz: challenges and opportunities. In *Proceedings of ASSETS*, (2012) ACM Press, 135-142.
- Casey, C. April 2003. Real-Time Information: Now Arriving. *Metro Magazine*.
- Cohen, B. Incentives Build Robustness in Bittorrent. In *Proceedings of Workshop on Economics of Peer-to-Peer systems*, (2003), 68-72.
- Drenner, S., Sen, S., & Terveen, L. Crafting the initial user experience to achieve community goals. In *Proceedings RecSys*, (2008) ACM Press.
- Feldman, M. and Chuang, J. Overcoming Free-Riding Behavior in Peer-to-Peer Systems. *ACM SIGecom Exchanges*, 5, 4 (2005), 41-50.
- Ferris, B., Watkins, K. & Borning, A. One bus away: results from providing real-time arrival information for public transit. In *Proceedings of CHI*, (2010) ACM Press, 1807-1816
- Fuglestad, P., Dwyer, P., Filson Moses, J., Kim, J., Anna Mannino, C., Terveen, L., & Snyder, M. What makes users rate (share, tag, edit...)? Predicting patterns of participation in online communities, In *Proceedings of CSCW*, (2012) ACM Press.
- Grudin, J. Why Groupware Applications Fail: Problems in Design and Evaluation. *Information Technology & People*, 4, 3 (1989).
- Hughes, D., Coulson, G. and Walkerdine, J. Free Riding on Gnutella Revisited: The Bell Tolls? *IEEE Distributed Systems Online*, 6, 6 (2005).
- Jun, S. and Ahamad, M. Incentives in BitTorrent Induce Free Riding. In *Proceedings of ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems*, (2005) ACM Press, 116-121.
- Karakaya, M., Korpeoglu, I. and Ulusoy, O. Free Riding in Peer-to-Peer Networks. *IEEE Internet Computing*, 13, 2 (2009), 92-98.
- Kraut, R. E. & Resnick, P. 2012. *Building successful online communities: Evidence-based social design*. Cambridge, MA: MIT Press.
- Lane, N. D., Eisenman, S. B., Musolesi, M., Miluzzo, E. and Campbell, A. T. Urban Sensing Systems: Opportunistic or Participatory? In *Proceedings of HotMobile*, (2008) 11-16.
- Lasecki, W., Miller, C., Sadilek, A., Abumoussa, A., Borrello, D., Kushalnagar, R., & Bigham, J. 2012. Real-time captioning by groups of non-experts. In *Proceedings of UIST*, (2012) ACM Press, 23-34.
- Lasecki, W. S., Song, Y. C., Kautz, H., & Bigham, J. P. Real-time crowd labeling for deployable activity recognition. In *Proceedings of CSCW*, (2013) ACM Press, 1203-1212.
- Steinfeld, A., Zimmerman, J, Tomasic, A., Yoo, D., Aziz, R.D. Mobile transit rider information via universal design and crowdsourcing, *Transportation Research Record - Journal of the Transportation Research Board* 2217, (2012) 95-102.
- Thiagarajan, A., Biagioni, J., Gerlich, T. and Eriksson, J. Cooperative transit tracking using smart-phones. In *Proceedings of SenSys*, (2010) ACM Press.
- US Census Bureau (2010) <http://www.census.gov/prod/2010pubs/acsbr09-5.pdf> Accessed: September 1, 2012.
- Waze (2013) <http://www.waze.com> Accessed: May 30, 2013.
- Wilson, J. Volunteering. *Annual Review of Sociology*, 26 (2000) 215-240.
- Yoo, D., Zimmerman, J., Steinfeld, A., Tomasic, A. Understanding the space for co-design in riders' interactions with a transit service. In *Proceedings of CHI* (2010), ACM Press.
- Zhou, P., Zheng, Y., Li, M. How long to wait? Predicting bus arrival time with mobile phone based participatory sensing. In *Proceedings of MobiSys*, (2012) ACM Press.
- Zimmerman, J., Tomasic, A., Garrod, C., Yoo, D., Hiruncharoenvate, C., Aziz, R., Thiruvengadam, N.R., Huang, Y., Steinfeld, A. Field trial of Tiramisu: Crowdsourcing bus arrival times to spur co-design. In *Proceedings of CHI*, (2011) ACM Press.