Crowdsourcing GO: Effect of Worker Situation on Mobile Crowdsourcing Performance

Kazushi Ikeda
KDDI Research, Inc.
Fujimino, Japan
kz-ikeda@kddi-research.jp

Keiichiro Hoashi
KDDI Research, Inc.
Fujimino, Japan
hoashi@kddi-research.jp

ABSTRACT
The increasing popularity of mobile crowdsourcing platforms has enabled crowd workers to accept jobs wherever/whenever they are, and also provides opportunity for task requesters to order time/location specific tasks to workers. Since workers on mobile platforms are working on the go, the situation of the workers is expected to influence their performance. However, the effects of mobile worker situations to task performance is an uninvestigated area. In this paper, our research question is, "*do worker situations affect task completion, price and quality on mobile crowdsourcing platforms?"* We draw on economics and psychology research to examine whether worker situations such as busyness, fatigue and presence of companions affect their performance. Our three-week between-subjects field experiment revealed that worker busyness caused 30.1% relative decrease of task completion rate. Mean accepted task price increased by 7.6% when workers are with companions. Worker fatigue caused 37.4% relative decrease of task quality.

Author Keywords
Crowdsourcing; mobile crowdsourcing; situational effect; worker performance.

ACM Classification Keywords
H.5.3 Information interfaces and presentation (e.g., HCI): Group and Organization Interfaces.

INTRODUCTION

Mobile Crowdsourcing have gained popularity, where workers use their mobile phones to work on location and time specified tasks. Increasing number of research focus on mobile crowdsourcing such as restaurant recommendations for specific locations [4], and news article generation by requesting photos and videos from on-site workers [52]. Some commercial service platforms have emerged including Gigwalk [2] and Field Agent [1] that mainly deal with mobile crowdsourcing tasks. Push notification is a key function to actively ask workers for contribution who meet the task conditions such as location and time. Unlike conventional crowdsourcing platforms, where workers do tasks at their convenient time, workers on mobile crowdsourcing platforms may receive task requests via push notifications in situations when they are not suitable for doing tasks. However, the effect of worker situation on mobile crowdsourcing performance has not been investigated well.

In this paper, we set a research question: "*does worker situation affect task completion rate, accepted task price and task quality on mobile crowdsourcing platforms?*" Drawing behavioral economic research on crowdsourcing has successfully highlighted an effective incentive method from potential alternatives [29]. By following the idea, we draw on the economics and psychological literatures to derive critical worker situations that are expected to affect mobile crowdsourcing performance. First, the idea of *opportunity cost*, i.e., a concept where someone obtains something by giving up others [47], is referred to extract the busyness factor. When a worker feels busy, he has something more worth of doing than the crowdsourcing task, which may result in lower task completion rate or require higher price to complete the same task. Second, from biological psychology, we hypothesize the effect of mental fatigue. Namely, we expect that mental fatigue of workers decreases their motivation and affects their work quality [10]. Third, we focus on the existence of companions. In the presence of companions, e.g., family members, friends, etc., touching mobile phone or even putting it on the desk can increase mental barriers for task completion [42]. In other words, crowd workers may feel rude to do task in the presence of companions.

This study aims to examine whether these worker situations, i.e., degree of busyness, degree of fatigue, and presence of companions, affect worker performance on mobile crowdsourcing platforms. Verifying the validness of these three hypotheses for mobile crowdsourcing is not a trivial issue, since the existing studies in the economic and psychology domain do not assume the mobile environment. For example, mental fatigue has been reported to decrease worker performance in continuous tasks, but its applicability to a one-time task on mobile crowdsourcing must be investigated thoroughly. We conduct between-subjects field

CHI 2017, May 6–11, 2017, Denver, CO, USA.

ACM ISBN 978-1-4503-4655-9/17/05...$15.00.
DOI: http://dx.doi.org/10.1145/3025453.3025917

experiment with time-limited mobile crowdsourcing tasks. The task includes questionnaires that collect worker situation during their task completion. Content of tasks are unchanged throughout the experimental period. Task price is changed per task request, and controlled so that it would be uniformly distributed with a fixed median price. Workers tend to decline cheaper price tasks. If workers tend to decline more frequently in a specific situation, the mean task price for that situation increases. In order to evaluate task quality for each situation, we used gold standard, or known answer task.

Results of the experiment indicate that task completion rate was most affected by the degree of busyness. Compared to low busyness worker, high busyness worker showed 30.1 % relative decrease in task completion rate ($p < .001$). Mean accepted task price was most affected by the presence of companions. Workers with companions had a higher mean accepted task price by 7.6% ($p < .001$) compared to workers without companions. Task quality was most affected by the degree of fatigue. Compared to lowly fatigued workers, workers with high fatigue showed 37.4% relative decrease in task quality, measured by the accuracy (F-measure) of completed tasks ($p < .05$). Subjective opinions collected after the experiment also supported our hypotheses.

The rest of the paper is organized as follows. We first review related work on crowdsourcing performance, mobile crowdsourcing and notification to mobile devices. Second, we develop our research question and hypotheses by drawing theories in economics and psychology. We then introduce the experimental design of our field experiment and present quantitative and qualitative results. Finally, we describe limitations and future work.

RELATED WORK

Crowdsourcing Performance

Studies for improving performance on crowdsourcing platforms such as incentives and worker motivation have attracted much attention. Most of these studies focus on task completion, task price and task quality. Task completion and task price have strong relations with each other. Regardless of their motivations, microtask crowd workers seek to earn money [5]. They seek out tasks that maximize their expected earnings [26]. The simplest approach to increase task completion is to pay more: this was one of the first widely-cited results in crowdsourcing research [40]. Some existing studies realize higher task completion without task price increase. Participants increase their effort when the system reminds them of the uniqueness of their contribution [8]. Stimulating worker curiosity keeps longer participation on micro task crowdsourcing [36]. Introducing a method for shortening task completion time is effective to reduce task price. A framework for managing paid experts on the crowdsourcing platform has successfully reduced both the task completion time and the task price in relatively large projects, such as animation movie creation and development of software [46]. In order to minimize the cost of micro tasks, a new interface has been proposed to drastically accelerate labeling tasks by allowing workers to make errors [35].

Task quality is also an important measure to evaluate crowdsourcing platform performance. Accepting the task only when workers agree with each other improves the accuracy of the task [49]. Offerings to train workers and give feedback can improve quality [54, 17]. Including known answers as gold standard has also improved individual worker precision, and aggregated task accuracy [37]. Simply offering higher payment does not increase quality [40]. The order of consequence tasks affects task accuracy [12].

Mobile Crowdsourcing and Notification

Mobile crowdsourcing technologies are increasingly gaining attention [39, 15]. Various types of applications and tasks have been developed on mobile crowdsourcing platforms. Simple transcription task of hand written text is provided by utilizing mobile phones as portable interfaces [23]. Recent advances in mobile technologies have also allowed for more intricate and creative tasks. For instance, location-based distribution of crowdsourcing tasks has allowed workers to perform highly context specific tasks. Some examples of this include giving location-aware recommendations for restaurants [4], and authoring news articles by requesting photographs or videos of certain events from workers [52]. An investigations of situated crowdsourcing has been reported, where kiosk devices located on a university campus are used to reveal the relationships among task completion, location, time and incentives [27]. Some of these applications require to push tasks to workers on specific location and time. However, worker situation has not been investigated well on mobile crowdsourcing platforms.

In HCI research, various notification methods for mobile devices have been proposed, which consider user situation estimated from sensed information on their mobile devices. Several researchers have investigated the effect of interruption caused by notification under various situations [3, 7, 34]. Relevance of information to the workers' situation is a key factor to avoid notification from being noisy [16]. Even though many users are frustrated with notifications, they prefer to keep using notification systems for information delivery, instead of checking for new information manually [30]. Wearable devices have also been used to notify information in the timing of activity switch [25]. Psychological approaches help to understand human behavior towards notifications. Psycho-physiological states of users under interruptive notifications have been investigated [56]. The influence of task interruption on individual decision making has been investigated to quantitatively evaluate the effects of overloaded notifications [50].

SITUATIONAL EFFECT ON WORK PERFORMANCE

Research Question

Existing studies on notifications mainly focus on increasing response rate, and targeting notifications of new information

or advertisements on mobile devices. However, these techniques are not suitable for improving the performance of mobile crowdsourcing platform, such as task completion rate, task price and task quality. Existing work has shown that physical location of workers affects their decision of accepting task prices on mobile crowdsourcing [43]. Furthermore, qualitative studies based on worker interview for the motivation on the mobile crowdsourcing have also indicated that situational factors can affect worker motivation [51, 14]. Despite such efforts, effects of workers' situations on the task has not been quantitatively investigated well. This observation motivates our research question:

RQ: *Does worker situation affect task completion rate, accepted task price and task quality on mobile crowdsourcing platforms?*

Hypothesis

In order to develop reasonable hypotheses of worker situations that affect mobile crowdsourcing performance, we draw on economics and psychology literature. There is a rich history of research in economics and psychology that can help us to understand how human internal factors such as economic sense and mental status affect worker performance. This section presents a review of studies in economics and psychology, in order to establish the hypotheses of this research.

In economic theory, the opportunity cost of a resource is referred as the value of the next-highest-valued alternative use of that resource. Generally, the opportunity cost of time is measured by wages or income [38, 20]. The value of time for individuals without income, such as house workers, has been estimated by indirect approaches [22, 55]. Variety of human activities such as shopping [41] and leisure [32] have been explained by opportunity cost, and have been priced based on alternatives. So far, few studies have discussed the applicability of opportunity cost to the situation of workers in crowdsourcing. When a worker is busy, s/he is required to give up many alternatives to complete a mobile crowdsourcing task. We assume the opportunity cost of a task is relatively higher under busyness. Therefore, our first hypothesis is,

H1: Task completion rate on mobile crowdsourcing platforms decreases, and the accepted task price increases when workers are feeling busy.

Mental fatigue is a very common phenomenon that applies major consequences for everyday task performance. Fatigued people often experience difficulties in concentration and appear to be more easily distractible [11]. Mental fatigue also affects worker motivation and productivity [10]. Existing studies mainly focus on continuous tasks such as driving a car. For example, driving errors such as large speed variations and even running off the road have observed to be increasingly frequent in correlation to driver's fatigue [13]. Mental fatigue also affects more simple tasks such as recognizing letters on the monitor, where continuous work has increased error rate [18]. Based on such existing studies, we assume that fatigue caused by external situations may affect worker performance on mobile crowdsourcing tasks, which leads to our second hypothesis:

H2: Task quality on mobile crowdsourcing platform decreases when workers feel fatigue.

The presence of mobile devices can increase mental barriers for face-to-face communication. In a field experiment, 100 couples discuss a randomly assigned topic together. Some participants held their mobile device in their hand or placed it on the table. The conversations were self-rated by the participants. Conversations without mobile devices were rated as significantly superior compared with those in the presence of a mobile device [42]. Some workers may feel rude to work on their mobile devices while talking or being with companions. In addition, multi-tasking sometimes decreases productivity by causing more errors and slowing task completion. In a university class, the professor asked students to use laptops during the lecture, which resulted in lack of students' concentration and significant decrease in the examination score [24]. More simply, watching television while remembering sets of digits or while doing homework resulted in poorer performance than when focusing on a single task [6, 45]. Based on these literature, we derive our third hypothesis as follows. Note that the previous literature do not directly evaluate the effect of the presence of companions on task performance.

H3: Task completion rate on mobile crowdsourcing platforms decreases, task price increases, and task quality decreases when workers are with companions.

EXPERIMENTAL DESIGN

To investigate whether the worker situations of busyness, fatigue and the presence of companions affect worker performance on mobile crowdsourcing platforms, we have conducted a between-subjects experiment in August 2016. In the experiment, we recruit participants to install a mobile crowdsourcing application that sends tasks intermittently throughout the day. Workers are allowed to accept or ignore the tasks according to their situation. A different task price is set per each task request, and controlled so that it would be uniformly distributed with a fixed median price. A known answer task (gold standard) is used to evaluate task accuracy for each situation. Workers are asked to answer their situation of degree of busyness, degree of fatigue, presence of companions along with their task completion. We also periodically send additional questionnaire to collect worker situations when workers ignore the task. The task completion rate, mean accepted task price and task quality are measured for each situation.

Participants

We recruited participants (N=50, 58.0% female, aged 20–58, mean age=37.5, std. dev=9.1) through email advertisements sent to approximately 1,000 staff working for a recruiting company in Japan. Their living and working location

Task	Quality of service survey
Notification	27 times/day (7AM-8:30PM)
Required time	216 sec (median) / task
Incentive price	10~50 yen/task with 5-yen basis (uniform distribution with mean of 30 yen)
Measurement	Task completion rate, mean task acceptance price and task quality
Participants	50
Duration	3 weeks from Aug 8 to Aug 28, 2016

Table 1. Summary of experimental settings. We evaluated task completion rate, mean accepted task price and task quality in different worker conditions.

expands over the whole country of Japan. Participants must have an Android or iPhone smartphone to participate. Other than that, there were no restrictions applied to the participants by their demographics, and assume that no factor was concerned to cause bias in worker demographics.

Method

Since there is no mobile crowdsourcing platform service available in Japan (e.g., Gigwalk) which support push notification, the experiment was conducted on our own crowdsourcing platform. The platform is capable of providing tasks to workers based on specific location and time. All participants installed an application for this experiment on their phones, and the task was branded as a quality-of-service survey from the telecommunications company. It is notable that mobile network monitoring is common task on mobile crowdsourcing platforms [19, 33].

We periodically send task notifications to workers. Workers can select whether s/he accept the task or not after confirming the task price for each notification. Existing studies about interruption determine the availability of users based only on their response to notifications. In mobile crowdsourcing, we must consider not just the workers' availability, but also their current situations and task conditions (e.g. task price). Thus, we categorize workers' response to the following 3 categories: (1) Unavailable (No response to notification), (2) Decline (Confirm notification but choose not to complete task), and (3) Accept (Complete task). These categories can be distinguished from the experiment log data

Table 1 summarizes the experimental conditions. The application notifies tasks to workers approximately every 30 minutes, 27 times per day. Since tasks are location dependent, the invitation expires five minutes after notification. After confirming the task, the completion deadline is set to 15 minutes. The median of task completion time was 216 sec. The task price is randomly decided every time between ¥10 yen ($1.5/hr) and ¥50 ($7.5/hr) by ¥5 basis with a uniform distribution of ¥30 median ($4.5/hr). The experiment was conducted for a three-week period.

In the task, workers are asked to watch a three-minute streaming video and answer the questionnaire about the quality of the video and their situations. Evaluating Quality

Q. 1 Please answer the following questions after watching the video.
Q. 1-1 How many times did the video suspend?
(a) None (b) Once (c) Twice (d) Three times (e) Four times or more (f) Suspend and never resume (after 15 sec passed)
Q. 1-2 How many times did you find noise in the video?
(a) None (b) Once (c) Twice (d) Three times (e) Four times or more
Q. 2 Please answer about your situation.
Q. 2-1 What are you doing now? (Select one for each)
(a) Moving (b) Staying in one place
(a) on the Train (b) Outdoors (c) Indoors
Q. 2-2 Please describe details of your location (Free text)
e.g. "I'm at station [X] on the [Y] line"
Q. 2-3 How busy did you feel before starting this task?
(a) Very busy (b) Moderately busy
(c) Moderately not busy (d) Not busy at all
Q. 2-4 How much fatigue did you feel before starting this task?
(a) Very high (b) Moderately high
(c) Moderately low (d) Not at all
Q. 2-5 Describe currently present companions
(a) Alone (b) Family (c) Friends (d) Colleagues
(e) Others (free text)

Table 2. Task description (translated from Japanese).

of Experience (QoE) for video streaming by crowd workers under various network performance is a famous crowdsourcing application [28]. Table 2 describes the detail of the questionnaire. After watching the video, workers are asked to answer the number of suspended times of the streaming and the number of noise appeared on the screen due to low communication service quality (an example frame extracted from a typical noisy video is shown to workers). Workers are also asked to answer their situation that affects communication quality, such as their current location and moving status (Q. 2-1, 2-2). Finally, workers are asked about their degree of busyness, degree of fatigue, and presence of companions (Q. 2-3, 2-4, 2-5).

Although the content of streaming video is the same for every task, we randomly assign noisy video with a 25% probability to evaluate task accuracy. We prepare five types of noisy video, by inserting 15 sec of noise in different parts of the three-minute video. In order to avoid cases where the streaming quality actually degrades due to low quality communication service, the application buffers the entire streaming video. In case the playing time within the video catches up with the end of the buffered video, the streaming video is paused and never resumes (workers answer (f) in Q. 1-1 in such cases). This mechanism allows us to distinguish the gold standard from actual degradation of communication service quality.

Conditions

In the experiment, we evaluate three types of worker situations; busyness, fatigue, and presence of companions, with three measurements of crowdsourcing performance; task completion rate, mean accepted price and task quality. The three types of worker situations are collected by the questions from Q. 2-3 to Q. 2-5 in Table 2. We ask busyness and fatigue of worker with four graded questions and grouped into two categories, respectively. Namely, we asked workers to select one from the four degrees of busyness: "very busy", "moderately busy", "moderately not busy" and "not busy at all". The former two options are grouped into high busyness (HB), whereas the latter two options are grouped into low busyness (LB). In the same manner, the four degrees of fatigue are grouped into high fatigue (HF) and low fatigue (LF). The presence of companion is grouped into "with someone" (WS) and "alone" (AL).

In order to examine task completion rate, we also need to collect worker situations when workers do not complete tasks. For this purpose, we send additional (out-of-task) questionnaire every three hours, inquiring the situation of workers (Q 2-3 to Q 2-5) during the last 3 hours. This questionnaire enables exhaustive data collection in all situations, regardless of task completion.

Measures

We evaluate task completion rate, mean accepted task price and task quality across all situations. In order to collect worker situations, we used the out-of-task questionnaire to evaluate task completion rate, and the in-task questionnaire was used to evaluate mean accepted task price and task quality of completed tasks.

Task completion rate is defined for each situation as follows.

Task completion rate = number of completed tasks / number of notified tasks

Mean accepted task price is calculated for each situation only from the completed (accepted) tasks. The price of all notified tasks are uniformly distributed with the median of ¥30, thus, if workers tend to decline cheaper tasks (like ¥10) more frequently in a specific situation, the mean accepted task price for that situation increases.

We evaluate task quality using noisy videos as the gold standard. We can check how carefully workers pay attention to the task by analyzing the detection performance of the noisy videos. We used recall, precision, F-measure and accuracy as evaluation criteria. These criteria are defined as follows using TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative), where the noisy videos are positive.

$$Recall = TP / (TN + TP), Precision = TP / (FP + TP)$$

$$F\text{-}measure = 2 / (1 / Recall + 1 / Precision)$$

$$Accuracy = (TP + FN) / (TP + TN + FP + FN)$$

In the task, workers are required to answer the questionnaire after watching a three-minute video. Therefore, tasks completed earlier than three minutes are considered to be invalid, or cheated. We evaluate the average completion time and the ratio of invalid (shorter than three minutes) tasks as intermediate criteria for estimating task quality.

Finally, we asked the participants for subjective opinions about the relationship between their situations and performance. Namely, we asked three questions: "In which situation do you feel the most difficult to accept the task?" "In which situation do you feel the most difficult to concentrate on the task?" and "Have you declined low price tasks?" along with their reasons.

Method of Analysis

To determine whether worker situations affect their performance, we conducted statistical tests. All worker situations are grouped into two categories, as previously described. Namely, busyness and fatigue are grouped into high and low (HB, LB, HF and LF defined in the Condition section). The presence of companions is grouped into "with someone" (WS) or "alone" (AL). Therefore, condition groups consist of two unpaired groups. Since task completion rate and task accuracy are binominal data, i.e., tasks completed or not, accurate or not, we used Chi-square test to compare the significance of the difference. As for mean accepted task price and task completion time, we can assume normality in their distribution. Therefore, we used t-test to compare the differences between the condition groups.

RESULTS

The recruited 50 participants received 27 notifications per day, 567 notifications in total during the three-week experimental period. A total of 47 participants out of 50 participants completed at least three tasks during the period. The total number of completed tasks was 3,596. The response rate of the out-of-task questionnaire was 36%. Table 3 shows the number of worker situations answered in the in-task and out-of-task questionnaire. The quantity of answers in each situation are sufficient for statistical analysis, and are relatively evenly distributed among the situations. In the experiment, we used the number of tasks for the statistical test. The coincidence of worker situations for the same time exceeded 90% between in-task and out-of-task questionnaires, which indicates that workers were able to accurately recall their situations for specific times of day that were selected in the out-of-task questionnaires.

Task Completion

First, we examine the relationship between worker situation and task completion. In order to compare task completion rate among worker situations, we need to collect the situation of workers when they have ignored the task. We used the out-of-task questionnaire for this purpose. Table 4 shows the task completion rate for each situation, absolute and relative difference of task completion rate, and p value of Chi-square test of the differences within each situation.

Situation	in-task	out-of-task
HB	1340	3727
LB	2256	4486
HF	1829	4200
LF	1767	4013
WS	1879	5366
AL	1717	2847

Table 3. The number of worker situations answered in the in-task and out-of-task questionnaire. The numbers of each situation were large enough for statistical analysis, and were free from large deviation among the situations.

Situation	Task Completion (%)	Difference Abs. / Rel. (%)	p value
HB	25.7	7.74 / 30.1	.0000
LB	33.4		***
HF	29.7	2.28 / 7.72	.7554
LF	30.1		
WS	28.2	5.03 / 17.9	.0000
AL	33.2		***

Table 4. Task completion rate across worker situations. Workers with high busyness (HB) decreased task completion rate by 30.1% relative compared to workers with low busyness (LB) (p<.001).

When workers are busy, the task completion rate has decreased by 7.74% (30.1 % relative decrease) compared to when workers are not busy. Our Chi-square test with Yates' continuity correction reveal that the difference is significant ($\chi^2(1, N=8213) = 57.8$, $p < .001$, $\varphi = 0.084$). As for fatigue, no significant difference between the task completion rates for high fatigued workers and low fatigued workers is observed. When workers are accompanied with someone, task completion rate has decreased by 5.03% (17.9 % relative decrease) compared to when workers are alone, The Chi-square test reveals that this difference is also significant ($\chi^2(1, N=8213) = 22.2$, $p < .001$, $\varphi = 0.052$). From these results, it can be concluded that task completion rate greatly decreases when workers are busy or with other people.

Task Price

We measured mean accepted task price to analyze the relationship between worker situations and task price. Since task price is uniformly distributed from ¥10 to ¥50 with a median of ¥30 yen, if workers tend to decline cheaper tasks (like ¥10) more frequently in a specific situation, the mean task price of that situation increases. Unlike task completion rate, calculation of mean accepted task price requires only the information of completed task data. Therefore, we used only the worker situations collected by in-task questionnaire. Table 5 shows mean accepted task price for each situation, difference of mean accepted task price, and p value of t-test for the difference.

Situation	Mean accepted task price (yen)	Difference (%)	p value
HB	32.8	5.95	.0050
LB	30.9		**
HF	32.2	4.02	.0543
LF	31.0		
WS	32.8	7.67	.0003
AL	30.5		***

Table 5. Mean accepted task price across worker situations. Workers with someone else (WS) increased mean accepted task price by 7.67% compared to workers alone (AL) (p<.001).

When workers are busy, the mean accepted task price has increased by 5.95%, compared to when workers are not busy, where a Welch's t-test reveals that the difference is significant ($t(2808) = 2.29$, $p < .001$). The mean accepted task price has increased 4.02% for fatigued workers compared to when workers do not feel fatigue, but the difference is not statistically significant according to the t-test ($t(3592) = -0.09$, $p = .0543 > .05$). When workers are with someone, mean accepted task price has increased by 7.67% compared to when workers are alone. The t-test reveals that the difference is significant ($t(3561) = 3.43$, $p < .001$). These results indicate that when workers are busy or with someone else, they decline lower priced tasks more frequently.

Based on the above results, we conclude that workers' busyness and the presence of companions cause significant differences in both task completion rate and mean accepted task price. This result can be explained by the findings in prior work [40], where task completion and task price have strong relations with each other.

Task Quality

We examine the relationship between worker situation and task quality. In the task, workers are asked to watch three-minute streaming video, and answer the questionnaire about quality of communication service. We assigned videos which contain intentional noise as the gold standard with a 25% probability. The direct measurements for evaluating task quality are recall, precision, F-measure and accuracy for detecting noisy video, as explained in the Experimental Design section. As an indirect measurement, we evaluate the task completion time. We compare average task completion time across the worker situations. The tasks that were completed shorter than three minutes are considered to be invalid. We also measure the rate of invalid tasks. Table 6 shows mean and median of task completion time, the invalid task rate, and p value of t-test for the difference in the mean task completion time.

When workers are either busy, fatigued or with someone, mean task completion time is significantly shorter ($p < .001$) compared to when workers are not busy, not fatigued or alone, respectively. Particularly, when workers feel fatigued, mean task completion time is 20.1% shorter compared to

Situation	Mean (sec)	Median (sec)	Invalid (%)	p value
HB	177	192	49.0	.0000
LB	202	225	33.5	***
HF	171	152	51.3	.0000
LF	215	229	26.8	***
WS	179	208	46.6	.0000
AL	208	224	31.3	***

Table 6. Mean and median of task completion time, the invalid task rate, and p value of t-test for the difference in mean task completion time. Workers feeling high fatigue (HF) decreased mean task completion time by 20.1%, and roughly doubled the invalid task rate, compared to workers feeling low fatigue (LF).

when workers don't feel fatigued. The rate of invalid task of fatigued workers is approximately two times larger than that of workers without fatigue. On crowdsourcing platforms, some lazy workers do not complete tasks in the expected manner. Without introducing techniques to control task quality, 30% of tasks can contain poor results [9]. The results of our experiment indicate that, similar to prior work, about 30% of tasks are invalid when workers are not busy, not fatigued or alone. However, when workers feel high fatigue and busyness, or are with someone else, the invalid task rate significantly increases, which results in the increase of low quality tasks.

Table 7 shows the relationship between worker situations and task quality measured by recall, precision, F-measure and accuracy. We designed this task assuming that telecommunication companies utilize the collected information to detect low quality areas of communication services. In such scenarios, criteria such as recall, precision and F measure are considered to be important. On the other hand, since recall and precision are in a trade-off relation, applying the statistical test is inappropriate. Thus, we applied Chi-square test to the accuracy measure to examine whether the detection accuracy of noisy video differs across worker situations.

We assigned videos containing 15 sec noise during the 3 min video as the gold standard with the probability of 25%. Due to the low probability of noise appearance, and highly demanding task design where workers are required to concentrate on the mobile phone screen for three minutes, the task was so difficult that the F-measure ranges around 0.2 to 0.4. Significant differences were observed in the accuracy across worker situations. On the basis of the F-measure, HB, HF and WS was lower compared to LB, LF and AL, respectively.

When workers feel high fatigue, both the recall and the precision has decreased compared to when workers feel low fatigue. F-measure has decreased from 0.364 to 0.265 (37.4% relative decrease), which was the largest decrease among all the worker situations. Our Chi-square test with Yates' continuity correction revealed that the difference in

Situation	Precision	Recall	F measure	Accuracy	p value
HB	0.567	0.188	0.282	0.742	.7193
LB	0.461	0.261	0.333	0.748	
HF	0.478	0.184	0.265	0.730	.0379
LF	0.500	0.287	0.364	0.761	*
WS	0.449	0.191	0.268	0.733	.0730
AL	0.529	0.277	0.364	0.759	

Table 7. Recall, precision, F measure and accuracy of the tasks across the conditions. Workers feeling high fatigue (HF) significantly decreased the accuracy of the tasks compared to the workers feeling low fatigue (LF).

the accuracy is significant ($\chi^2(1, N=3596) = 4.31$, $p < .05$, $\varphi = 0.035$). This result indicates that workers under fatigue could not concentrate on the task because fatigue decreases their productivity and causes errors, as described in our hypothesis. Similar to worker fatigue, when workers are with someone else, the precision, the recall and the F-measure have also decreased. This result indicates that workers have failed to concentrate on the tasks, either due to the distraction from their companions, or by attempting to complete the task as quickly as possible so that they will not keep the companions waiting. However, the p value of the Chi-square test was 0.073, which does not show significant difference in the accuracy of the task. As for busyness, high busyness increased the precision but decreased the recall. This is probably because workers did not check the detail of the video in busy situations, thus were able to detect only explicitly apparent noise, and ignored minor noise.

Subjective Opinion

After the experiment, we asked the participants for their subjective opinions about the relationship between worker situation and performance. Regarding task completion and task quality, we asked in which situations they feel difficulties to accept tasks, and in which situations they feel difficulties to concentrate on tasks, along with their respective reasons. Of the 50 participants who participated in the experiment, 42 participants (84%) answered the questionnaire. Table 8 shows the number and the rate of answered situations where workers feel difficulties to accept tasks. Table 9 shows the number and the rate of answered situations where workers feel difficulties to concentrate on tasks.

In Table 8, busyness was selected as the most significant factor which applies difficulty to workers to accept tasks. The second most significant factor was companions. These results are consistent with the results of the quantitative evaluation, where the task completion rate was the lowest when workers are busy, and was the second lowest when workers are with someone else. Common reasons why busy workers feel difficult to accept the tasks are, *"I couldn't check the notification of my private phone while I'm at my office"* and *"I couldn't find three minutes on my busy schedule."* Common reasons why workers with someone feel difficult to accept the tasks was, *"I felt it rude to watch the*

Situation	# of answers	% of answers
High busyness	21	50.0
High fatigue	4	9.5
With someone	10	23.8
Others	7	16.7

Table 8. The number and ratio of situations where workers feel difficulties to accept tasks. Many workers answered that busyness and companions make them difficult to accept tasks.

Situation	# of answers	% of answers
High busyness	21	50.0
High fatigue	4	9.5
With someone	15	35.7
Others	2	4.8

Table 9. The number of ratio of situations where workers feel difficulties to concentrate on tasks. Similar to the result in Table 8, many workers answered that busyness and companions make them difficult to concentrate on tasks.

video when I was with someone." These opinions indicate that the experiment of Misra [42] that we referred for the establishment of our hypothesis is also valid for mobile crowdsourcing platforms.

Table 9 indicates the busyness is the largest, and the presence of companions is the second largest factor to apply difficulties to workers to concentrate on the task. However, the results of the previous quantitative experimental results show that workers most lost their concentration when they were feeling high fatigue. This result indicates that workers' feelings and their actual behavior are different. The main reason why busy workers feel difficult to concentrate on the tasks is, *"I couldn't keep my eyes on my mobile phone due to the busyness."* The common reason of workers who selected companions was, *"I cared about my companions and tried to complete the task as quickly as possible."* These opinions are considered to be the cause of decrease in task quality. Although only a few workers selected fatigue as the reason, one of the workers commented that *"I was too sleepy to keep watching the video until the end."* This indicates that it may be difficult for workers to notice the reduction of task quality caused by their own fatigue.

DISCUSSION

Our research question in this paper is, "*does worker situation affect task completion rate, accepted task price and task quality on mobile crowdsourcing platforms?"* By drawing economics and psychology literature, we established three hypotheses about worker situations (busyness, fatigue and the presence of companions). We begin our discussion by highlighting whether these hypotheses are supported by our experiments. Table 9 shows the correspondence matrix of our hypotheses and experimental results.

Table 10 illustrates that most of our hypotheses are supported by the experimental results. The only hypothesis that was not able to be confirmed was the decrease of task quality of

Situation	Completion	Price	Quality
Busyness	Supported	Supported	NA
Fatigue	NA	NA	Supported
Companion	Supported	Supported	Not supported

Table 10. Correspondence matrix of hypotheses and experimental results about the relationships between worker situations and task performance. In most part of the matrix, our hypotheses are supported by experimental results, but the task quality reduction with companions was not supported.

workers with companions, since the Chi-square test did not show a significant difference with p value of 0.067 (> .05). We will discuss the details of the effect of worker situations on task completion rate, accepted task price and task quality, as follows.

For task completion rate, existing studies on notifications to mobile devices mainly focus on reducing the burden of interruption caused by the notification, and increasing the response rate toward the notification. By analyzing the log data of mobile devices, momentary situations like switching applications on the phone are detected as a suitable timing for notification [44]. However, results from such existing work cannot be applied to mobile crowdsourcing platforms because workers consider the amount of time necessary to complete a task before deciding whether or not to accept it. In the experiment, we assumed a task which requires more than three minutes to complete. As some participants have reported, we observe that workers hesitate to accept tasks when they are busy or with someone else, because they are unconfident to complete the task under such situations.

Existing studies report that task completion and task price have strong relations with each other [40]. We were also able to observe such results in our experiments, that busyness and the presence of companions both affect task completion rate and mean accepted task price. When workers are busy, they have many alternatives that increase the opportunity cost of the mobile crowdsourcing task. Workers are supposed to regard the time with companions as one of such valuable alternatives. After the experiment, we asked workers whether they have declined low price tasks. As the result, 50% of workers answered yes, and 50% answered no. One of the common reasons of workers who answered yes for this question was that the task was not cost effective. One of the subjects was clearly conscious about the idea of opportunity cost, mentioning that *"I prioritize the ¥50 task even when I am busy, but I prioritize other things for the ¥10 task."* Contrarily, typical reasons of workers who answered no for this question were, *"I accepted the task even with low price because it increases my income"* and *"I decided whether I accept the task or not only based on my situation."* These opinions indicate that worker situations affect their decisions on the acceptance of low priced tasks.

For the task quality, mental fatigue has been reported to increase errors in continuous tasks [11]. However, we haven't found any studies on the relationship between

worker fatigue and task quality on mobile crowdsourcing platforms, where small tasks are separately pushed to workers. Our experimental results indicate that worker fatigue affects task quality even for one-time tasks.

Let us discuss how the findings from our experiments can be leveraged in practical situations. The experimental results show that worker situations affect mobile crowdsourcing performance of task completion rate, task price and task quality. In order to utilize worker situation on practical mobile crowdsourcing platforms, new technologies which can accurately obtain worker situations are required. Development of such technologies, such as sensor devices and methods to analyze signals from these sensors, and also their implementation to smartphones is expected to make this possible in the near future.

We present our analysis to explore the possibility to leverage our results currently, i.e., without the anticipation of future technology developments. First, we summarize the degree of busyness, degree of fatigue, and the presence of companions obtained from out-of-task questionnaire. Figure 1 shows the ratio of workers in HB (High Busyness), HF (High Fatigue) and WS (With Someone) situations defined in the Experimental Design section, grouped by weekdays, weekends, and time of day. According to Figure 1, the ratios of busy workers and fatigued workers, and the ratio of workers with someone are lower in the morning of both weekdays and weekends. These times of day are considered suitable for requesting tasks. The ratio of workers with someone increases around noon on weekday but relatively decreases on weekends. Figure 1 shows the summary of the situations of all workers, but life styles are different for each worker, such as working days and hours. We can optimize task completion rate, task price and task quality by requesting tasks for workers with suitable situations obtained from their activity logs. The lower part of Figure 1 shows the mean accepted task price for each time of day on weekdays and weekends. From the figure, there seems to be a correlation between worker situations and task price. The correlation coefficient between the rate of busy workers and the mean accepted task price was 0.58 on weekdays, and 0.50 on weekends, which indicates a moderately positive correlation. This finding indicates that requesters can save the cost of mobile crowdsourcing by requesting tasks in time slots when workers are not busy.

The rest of this section discusses the limitations of this research. First, we focus on the representativeness of the mobile crowdsourcing platform and the task used in the study. We conducted the experiment on our own mobile crowdsourcing platform, mainly because there were no commercial mobile crowdsourcing platforms in Japan which support push notifications like Gigwalk. Our platform operates similarly to Gigwalk and others, in that it is capable of providing tasks to workers based on specific location and time. We believe that the findings from this platform can be generalized to other mobile crowdsourcing platforms that support push notification to workers. However, the applicability to each mobile crowdsourcing platform should be examined accordingly.

We used a survey task to evaluate the quality of communication service, where workers are asked to watch a video and answer a set of questionnaires. There are several reasons to think this task can be generalized to other mobile crowdsourcing tasks. First, there are many labeling tasks for video contents on crowdsourcing platforms. Second, "survey" is one of the top twenty keywords on Amazon Mechanical Turk [31]. Third, information sensing tasks such as reporting weather and measuring amount of traffic are all popular ways of use for mobile crowdsourcing platforms [53]. We designed the task so that workers require more than three minutes to complete the task, because workers would accept all tasks without paying attention to the task price if we set a short completion time. In our experiment, workers were required to use up to several tens of seconds to confirm the notification on their mobile phone and check the price of the task, prior to task acceptation. The switching cost for suspending current activity is also large. If we set the task completion time to one minute, the additional time cost to accept the task will exceed time necessary to complete the task, which is calculated as sunk cost. As the result, workers would accept all tasks or decline all tasks regardless of the task price. We designed the three-minute task to avoid this problem. As a result, 50% of workers decided whether they accept the task or not after confirming the task price, which shows that our experimental design worked well.

In this paper, we have asked the degrees of busyness and fatigue, and the presence of companions as typical worker situations. Analyzing worker situations in more detail is our future work. For example, categorizing the reasons of busyness, the cause of fatigue and the conversational content with the companion would be helpful to deeply understand the effect of worker situations on their performance. Existing study about interruptibility shows that conversation context affects response rate toward notifications from mobile devices [48], indicating the importance of such effects.

Another remaining question is, *"Are there any other situations that affect worker performance?"* We do not think that our three hypotheses cover all possible situations which affect mobile worker performance. Many alternative situations can be listed, e.g., extrinsic factors such as weather and social events, and physical conditions such as means of transportation and indoor/outdoor. Contributions of this paper are the definition of the three hypotheses drawn from economics and psychology studies, and the examination of their actual effect to improve performance on mobile crowdsourcing platforms. The evaluation of other worker situations is our future work.

Another limitation is the validness of our hypotheses for tasks that are more excessively associated with worker movement, e.g., field surveys, which are also popular on mobile crowdsourcing platforms. In these tasks, more

Day	Situation	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Weekday	HB	0.378	0.426	0.488	0.516	0.534	0.566	0.581	0.535	0.515	0.518	0.548	0.628	0.612	0.618
	HF	0.394	0.384	0.401	0.447	0.486	0.544	0.548	0.568	0.560	0.599	0.635	0.688	0.678	0.718
	WS	0.550	0.595	0.635	0.656	0.674	0.716	0.721	0.688	0.674	0.637	0.668	0.692	0.623	0.509
Weekend	HB	0.215	0.252	0.289	0.252	0.291	0.333	0.351	0.387	0.367	0.356	0.406	0.471	0.526	0.444
	HF	0.363	0.339	0.341	0.358	0.409	0.458	0.477	0.544	0.541	0.586	0.558	0.563	0.603	0.651
	WS	0.549	0.566	0.610	0.614	0.615	0.678	0.690	0.668	0.700	0.671	0.685	0.697	0.744	0.683

Day	Task Price	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Weekday		28.7	30.2	30.9	31.2	32.2	31.6	30.3	27.4	30.8	30.1	30.5	31.7	32.5	32.9
Weekend		29.3	30.7	28.6	30.6	31.8	29.7	32.6	30.1	29.8	29.8	31.7	32.1	34.5	28.8

Figure 1: The ratio of workers in HB (High Busyness), HF (High Fatigue) and WS (With Someone) situations (upper table) and mean accepted task price (lower table) grouped by weekdays, weekends and time of day. Worker situations and mean accepted task price differ depending on the time and day.

sophisticated worker situations (for example, staying inside or outside) are considered to affect worker performance. In related papers. We limited the scope of tasks into those which do not require physical movement, in order to simplify the research issue. However, since tasks associated with worker movement are one of the major tasks on mobile crowdsourcing platforms, evaluation of the effect of worker situations for these tasks is an interesting area to explore in the future.

We have conducted the experiment without the existence of other crowdsourcing tasks. On commercial mobile crowdsourcing services, many tasks are available at the same time. Evaluation of the relationship between worker situations and their task selection under multiple task conditions is also another interesting direction of future work.

CONCLUSION

In this paper, we have investigated the effect of worker situations such as busyness, fatigue and the presence of companions on task completion rate, task price and task quality on mobile crowdsourcing platforms. In order to develop reasonable hypotheses of worker situation that affect mobile crowdsourcing performance, we draw on economics and psychology literature. We conducted a between-subjects experiment with 50 participants using a mobile crowdsourcing app to examine our hypotheses. The experimental results show that task completion rate decreased by 30.1% in relative value when workers are busy. Mean accepted task price increased by 7.6 % when workers are with companions. Task quality decreased by 37.4% in relative value of F-measure when workers feel fatigue. Subjective opinions obtained after the experiment also supported our hypotheses in that workers feel difficult to accept the tasks when they are busy and with someone else.

On mobile crowdsourcing platforms, there are strong demands from task requesters to complete time/location-specific tasks with quick response and higher quality. Our experiment show a possibility to solve these requesters' demands from the observation of the effect of worker situations on task completion rate and task quality. As a practical method to collect worker situations, we show that worker situations differ depending on the time and day. In the near future, technical advances in sensor devices and analytics of human activity will realize more highly accurate estimation of worker situations. At that time, we believe our study will greatly contribute to increase mobile crowdsourcing performance.

We summarize the contributions of our study as follows:

- We have established hypotheses about the effect of worker situations on mobile crowdsourcing performance by drawing economics and psychology literature.
- We have carried out a mobile crowdsourcing field experiment with 50 participants and investigated valid worker situations in quantitative and qualitative ways.
- We have shown a practical use scenario to apply worker situations on mobile crowdsourcing by showing the difference of worker situations depending on the time and day.

So far, we have investigated our hypotheses with a quality of communication service survey task. In order to expand our findings to the tasks associated with worker movement and multi-task conditions, we plan to establish hypotheses about the factors that affects worker performance and behavior in those settings.

ACKNOWLEDGEMENTS

We would like to thank Michael S. Bernstein of Stanford HCI group for valuable discussion about research design. We would like to thank Yasuyuki Nakajima, Toshiaki Tanaka, Hajime Nakamura, Tsuzuki Tojo, Yasuhiro Takishima and Kazunori Matsumoto of KDDI Research, Inc. for supporting the study. We also thank all of the participants in our experiment.

REFERENCES

1. Field Agent, https://www.fieldagent.net/
2. Gigwalk, http://www.gigwalk.com/
3. Adamczyk, P. D. and Bailey, B. P. If not now, when?: the effects of interruption at different moments within task execution. In Proc. CHI, 2004, 271-278.
4. Alt, F., Shirazi, A., Schmidt, A., Kramer, U. and Nawaz, Z. Location-based crowdsourcing: extending crowdsourcing to the real world. In Proc. NordiCHI 2010, 13-22.
5. Antin, J. and Shaw, A. Social desirability bias in reports of motivation for US and India workers on Mechanical Turk. In Proc. CSCW, 2011.
6. Armstrong, G. B. and Sopory, P. Effects of background television on phonological and visuo-spatial working memory. Communication Research, 24, 5, 1997, 459-480.
7. Bailey, B. P. and Konstan, J. A. On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. Computers in Human Behavior, 22, 4, 2006, 685-708.
8. Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P. and Kraut, R. Using social psychology to motivate contributions to online communities. In Proc. CSCW, 2004, 212-221.
9. Bernstein, M. S., Little, G., Miller, R. C., Hartmann, B., Ackerman, M. S., Karger, D. R., Crowell, D. and Panovich, K. Soylent: a word processor with a crowd inside. Communications of the ACM, 58, 8, 2015, 85-94.
10. Boksem, M. A., Meijman, T. F. and Lorist, M. M. Mental fatigue, motivation and action monitoring. Biological psychology, 72, 2, 2006, 123-132.
11. Boksem, M. A., Meijman, T. F. and Lorist, M. M. Effects of mental fatigue on attention: an ERP study. Cognitive brain research, 25, 1, 2005, 107-116.
12. Cai, C. J., Iqbal, S. T. and Teevan, J. Chain reactions: The impact of order on microtask chains. In Proc. CHI, 2016.
13. Campagne, A., Pebayle, T. and Muzet, A. Correlation between driving errors and vigilance level: influence of the driver's age. Physiology and Behavior, 80, 2004, 515-524.
14. Chang Y. J., Paruthi G. and Newman M. W. A field study comparing approaches to collecting annotated activity data in real-world settings. In Proc. UbiComp, 2015, 671-682.
15. Chatzimilioudis, G., Konstantinidis, A., Laoudias, C. and Zeinalipour-Yazti, D. Crowdsourcing with smartphones. IEEE Internet Computing, 16, 5, 2012, 36-44.
16. Czerwinski, M., Cutrell, E. and Horvitz, E. Instant messaging: Effects of relevance and timing. In Proc. HCI, 2000, 71-76.
17. Dow, S., Kulkarni, A., Klemmer, S. and Hartmann, B. Shepherding the crowd yields better work. In Proc. CSCW, 2012, 1013-1022.
18. Faber, L. G., Maurits, N. M. and Lorist, M. M. Mental fatigue affects visual selective attention. PLOS one, 7, 10, 2012, e48073.
19. Faggiani, A., Gregori, E., Lenzini, L., Luconi, V. and Vecchio, A. Smartphone-based crowdsourcing for network monitoring: Opportunities, challenges, and a case study. IEEE Communications Magazine, 52, 1, 2014, 106-113.
20. Goldman, A. and Johansson, J. K. Determinants of search for lower prices: An empirical assessment of the economics of information theory. Journal of Consumer Research, 5, 3, 1978, 176-186.
21. Gould, S., Brumby, D., Cox, A., Gonz´alez, V., Salvucci, D. and Taatgen, N. Multitasking and interruptions: a sig on bridging the gap between research on the micro and macro worlds. In Proc. CHI, 2012, 1189-1192.
22. Gronau, R. The intrafamily allocation of time: The value of the housewives' time. The American Economic Review, 63, 4, 1973, 634-651.
23. Gupta, A., Thies, W., Cutrell, E. and Balakrishnan, R. mClerk: enabling mobile crowdsourcing in developing regions. In Proc. CHI, 2012, 1843-1852.
24. Hembrooke, H., and Geri G. The laptop and the lecture: The effects of multitasking in learning environments. Journal of computing in higher education, 15, 1, 2003, 46-64.
25. Ho, J. and Intille, S. S. Using context-aware computing to reduce the perceived burden of interruptions from mobile devices. In. Proc. CHI, 2005, 909-918.
26. Horton, J. J., and Chilton, L. B. The labor economics of paid crowdsourcing. In Proc. of the 11th ACM conference on Electronic commerce, 2010, 209-218.
27. Hosio, S., Goncalves, J., Lehdonvirta, V., Ferreira, D. and Kostakos, V. Situated crowdsourcing using a market model. In Proc. UIST, 2014, 55-64.
28. Hoßfeld, T., Seufert, M., Hirth, M., Zinner, T., Tran-Gia, P. and Schatz, R. Quantification of YouTube QoE via crowdsourcing. In Multimedia, In Proc. ISM, 2011, 494-499.
29. Ikeda, K. and Bernstein, M. S. Pay It Backward: Per-Task Payments on Crowdsourcing Platforms Reduce Productivity. In Proc. CHI, 2016, 4111-4121.
30. Iqbal, S. T. and Horvitz, E. Notifications and awareness: A field study of alert usage and preferences. In Proc. CSCW, 2010, 27-30.

31. Ipeirotis, P. G. Analyzing the amazon mechanical turk marketplace. XRDS: Crossroads, the ACM Magazine for Students, 17, 2, 2010, 16-21.
32. Johnson, M. B. Travel time and the price of leisure. Economic Inquiry, 4, 2. 1966, 135-145.
33. Kaup, F., Jomrich, F. and Hausheer, D. Demonstration of NetworkCoverage–A Mobile Network Performance Measurement App. In Proc NetSys, 2015.
34. Kreifeldt, J. G. and McCarthy, M. E. Interruption as a test of the user-computer interface. In Proc. the 17th Annual Conference on Manual Control, 1981, 655-667.
35. Krishna, R., Hata, K., Chen, S., Kravitz, J., Shamma, D. A., Fei-Fei, L. and Bernstein, M. S. Embracing error to enable rapid crowdsourcing. In Proc. CHI, 2016.
36. Law, E., Yin, M., Joslin Goh, K. C., Terry, M. and Gajos, K. Z. Curiosity killed the cat, but makes crowdwork better. In Proc. CHI, 2016.
37. Le, J., Edmonds, A., Hester, V. and Biewald, L. Ensuring quality in crowdsourced search relevance evaluation: The effects of training question distribution. In Proc. CSE, 2010, 21-26.
38. Lippman, S. A. and McCall, J. The economics of job search: A survey. Economic inquiry, 14, 2, 1976, 155-189.
39. Liu, Y., Lehdonvirta, V., Alexandrova, T. and Nakajima, T. Drawing on Mobile Crowds via Social Media. Case UbiAsk: Image Based Mobile Social Search Across Languages. Multimedia Systems, 18, 1, 2012, 53-56.
40. Mason, W. and Watts, D. J. Financial incentives and the "performance of crowds". In Proc. HCOMP, 2009, 77-85.
41. Marmorstein, H., Grewal, D. and Fishe, R. P. The value of time spent in price-comparison shopping: Survey and experimental evidence. Consumer Research, 19, 1, 1992, 52-61.
42. Misra, S., Cheng, L., Genevie, J. and Yuan, M. The iPhone effect: the quality of in-person social interactions in the presence of mobile devices. Environment and Behavior, 48, 2, 2016, 275-298.
43. Musthag, M. and Ganesan, D. Labor dynamics in a mobile micro-task market. In Proc. CHI, 2013, 641-650.
44. Okoshi, T., Ramos, J., Nozaki, H., Nakazawa, J., Dey, A. K. and Tokuda, H. Attelia: Reducing user's cognitive load due to interruptive notifications on smart phones. In Proc. PerCom, 2015, 96-104.
45. Pool, M. M., Koolstra, C. M. and Van der Voort, T. H. Distraction effects of background soap operas on homework performance: An experimental study enriched with observational data. Educational Psychology, 23, 4, 2003, 361-380.
46. Retelny, D., Robaszkiewicz, S., To, A., Lasecki, W. S., Patel, J., Rahmati, N., Doshi, T. Valentine, M. and Bernstein, M. S. Expert crowdsourcing with flash teams. In Proc. UIST, 2014, 75-85.
47. Roberts, R. Getting the most out of life: The concept of opportunity cost. The Library of Economics and Liberty, 5, 2007.
48. Schulze, F. and Groh, G. Conversational context helps improve mobile notification management. In Proc. Mobile HCI, 2016, 518-528.
49. Shaw, A. D., Horton, J. J. and Chen, D. L. Designing incentives for inexpert human raters. In Proc. CSCW 2011, 275-284.
50. Speier, C., Valacich, J. S. and Vessey, I. The influence of task interruption on individual decision making: An information overload perspective. Decision Sciences, 30, 2, 1999, 337-360.
51. Teodoro, R., Ozturk, P., Naaman, M., Mason, W. and Lindqvist, J. The motivations and experiences of the on-demand mobile workforce. In Proc. CSCW, 2014, 236-247.
52. Vaataja, H., Vainio, T., Sirkkunen, E. and Salo, K. Crowdsourced news reporting: supporting news content creation with mobile phones. In Proc. MobileHCI, 2011, 435-444.
53. Vaish, R., Wyngarden, K., Chen, J., Cheung, B. and Bernstein, M. S. Twitch crowdsourcing: crowd contributions in short bursts of time. In Proc. CHI, 2014, 3645-3654.
54. Yu, L., André, P., Kittur, A. and Kraut, R. A. Comparison of social, learning, and financial strategies on crowd engagement and output quality. In Proc. CSCW 2014, 967-978.
55. Zick, C. D. and Bryant, W. K. Alternative strategies for pricing home work time. Home Economics Research Journal, 12, 2, 1983, 133-144.
56. Zijlstra, F. R., Roe, R. A., Leonora, A. B. and Krediet, I. Temporal factors in mental work: Effects of interrupted activities. Occupational and Organizational Psychology, 72, 2, 1999, 163-185.