

# Frequency and Duration of Self-initiated Task-switching in an Online Investigation of Interrupted Performance

Sandy J. J. Gould, Anna L. Cox, Duncan P. Brumby

UCL Interaction Centre, University College London, Gower Street, London, UK  
s.gould@cs.ucl.ac.uk, anna.cox@ucl.ac.uk, brumby@cs.ucl.ac.uk

## Abstract

Online experimentation has been shown to be a reliable method for collecting data quickly and easily. Nevertheless, it is a method of data collection that can have a negatively affect confidence in results because of the loss of experimental control that comes with moving experiments online. To relieve these concerns, researchers have developed a number of tools and techniques for evaluating the performance of online participants and crowdworkers. In this paper we develop a measure of participants' attentiveness during an online data-entry experiment. We explored participants' propensity to switch away from the experiment to other tasks, and the duration of time that they spent away from the experiment during each switch. Our results show that switching is prevalent, even when participants are asked not to switch to other tasks.

## Introduction

Online experiments are a reliable way of collecting experimental data quickly, cheaply and efficiently (see, e.g., Kittur, Chi, and Suh 2008; Dandurand, Shultz, and Onishi 2008). This is the case across a range of experiments, even those that rely on time-sensitive performance measures, such as reaction times (Komarov, Reinecke, and Gajos 2013). Despite this work, there are still concerns about the validity of data collected online. For instance, without experimenter supervision, participants might satisfice, finding shortcuts to completing a study. Participants might also be distracted by other tasks, like email or social networking, in the absence of an experimenter. Researchers have tried to address these concerns by developing systems to help weed-out poorly performing participants (Rzeszutowski and Kittur 2012), or by making interventions that encourage participants to take more care during studies (Kapelnier and Chandler 2010).

Though prior work might give researchers confidence in results by maintaining standards, it tells us little about why participants' performance deviates in the first place. It could simply be that online participants might not care about the quality of their work. Alternatively it might be that work done outside a controlled lab is more prone to external distractions and interruptions. It is well-known that people tend to spontaneously interrupt activities, sometimes because they are bored or frustrated, or perhaps because they realize they need to be doing something else (Jin and Dabbish 2009). Determining if and when participants are self-interrupting during online experiments is an important step toward understanding the environment in which online experiments are conducted, something that ought to be considered when interpreting results of online studies.

In this paper we present an investigation of self-initiated interruptions during the execution of an online routine data-entry task. We looked at how often participants switched to other activities and how long they were gone for. Our results show that the majority of participants frequently switched to other activities during the experiment, on average for around fifteen seconds at a time. Our results contribute to our understanding of participants' extra-experimental activities during online work and have implications for researchers involved in both crowdsourcing and online experimentation.

## Method

A total of twenty-two participants (14 female) with a mean age of 26 years ( $SD=8$  years) took part in the study. Participants were drawn from an online university subject pool and were paid £7 in Amazon vouchers for approximately one hour of their time. Sign-up and participation were handled end-to-end by the online system.

The experiment used a within-subjects 2x2 factorial design. The independent variables were interruption relevance (levels: relevant, irrelevant) and timing (levels: within-subtask, between subtask). Our measures were the frequency and duration of switches, that is, how often participants switched, and how long they were gone for. We captured this information by hooking into the focus and blur events fired when participants switched away from the experiment. Participants were told before the start of the study that it was essential that they didn't leave the experiment and that there would be reductions in remuneration for those who did. In reality, all participants were paid the same amount at the end of the study.

The task in this experiment was the *Pharmacy Task*, an adaptation of the *Doughnut Machine* (Li et al. 2006), a routine data-entry task. Participants are given a set of 'prescriptions' that contain values that must be copied into one of the five subtasks that make up the task. From time-to-time, participants were interrupted by the program and required to work on a secondary task. The relationship between these system-generated interruptions and self-initiated task-switching is out of scope of this paper.

## Results

We focus on the frequency and duration of task-switching behavior during the experimental trials (i.e., excluding instruction, training trials). First, we looked at how often participants switched away from the experiment by counting the number of times the experiment lost window focus. During the experiment trials, which took an average of 32 minutes in total to complete ( $SD=21m$ ), the median number of extra-experimental task switches was 6 with a mean of 12 (range 0 – 86). Switching was prevalent in the sample; 60% of participants (13 of 22) switched to something else at least once during the experimental trials. Over the whole study, including instruction and training, 80% of participants switched at least once.

The second measure of interest was the duration of these switches; how long did participants spend on other activities before returning to the experiment? For participants who switched, the average duration of a switch was 16 seconds, although there were large individual differences in the duration of switches ( $SD=33s$ ).

## Discussion

Our results show that the majority of participants switch away from experiments and do so frequently. When participants switch, it is often for non-trivial periods of time. This was despite participants being told that their remuneration could be reduced if they left the task. That participants do not obey instructions without incentive has

been previously documented in the literature (Kapelner and Chandler 2010) but we were surprised that the prospect of losing a proportion of the compensation available had little effect on participants' task-switching behavior.

This work demonstrates that self-interruptions are prevalent in online studies, as they are in traditional work environments (see Jin and Dabbish 2009). This should be concerning for online experimentalists, as laboratory work has traditionally sought to minimize external distractions during experiments. While we do not report the effect of task-switching on performance in this paper due to space constraints, previous work demonstrates that interruptions can have a deleterious effect on performance. More work is therefore required to understand the effects of discretionary task-switching on task performance in online experiments.

Unfortunately, it was not possible to see what participants were doing when they switched away from the task, only that they had switched. Nevertheless, this work shows that online participants cannot be relied upon to focus on an experiment to the exclusion of other things in their environment. Researchers should record task-switching and keep it mind when interpreting results.

## Acknowledgements

Work funded by CHI+MED, EPSRC grant EP/G059063/1.

## References

- Dandurand, Frédéric, Thomas R. Shultz, and Kristine H. Onishi. 2008. 'Comparing Online and Lab Methods in a Problem-solving Experiment'. *Behav Res Methods* 40 (2): 428–434.
- Jin, Jing, and Laura A Dabbish. 2009. 'Self-interruption on the Computer: a Typology of Discretionary Task Interleaving'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1799–1808.
- Kapelner, Adam, and Dana Chandler. 2010. 'Preventing Satisficing in Online Surveys'. In *CrowdConf*.
- Kittur, Aniket, Ed H. Chi, and Bongwon Suh. 2008. 'Crowdsourcing User Studies with Mechanical Turk'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 453–456.
- Komarov, Steven, Katharina Reinecke, and Krzysztof Z. Gajos. 2013. 'Crowdsourcing Performance Evaluations of User Interfaces'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
- Li, Simon Y. W., Anna L. Cox, Ann Blandford, Paul Cairns, and A. Abeles. 2006. 'Further Investigations into Post-completion Error: The Effects of Interruption Position and Duration'. In *Proceedings of the 28th Annual Meeting of the Cognitive Science Conference*, 471–476.
- Rzeszotarski, Jeffrey, and Aniket Kittur. 2012. 'CrowdScape: Interactively Visualizing User Behavior and Output'. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, 55–62.