Supporting Task Resumption Using Visual Feedback

Yikun Liu, Yuan Jia, Wei Pan, Mark Pfaff
School of Informatics and Computing
Indiana University, Indianapolis
Indianapolis, IN 46202
{yikliu, jiayuan, panwei, mpfaff}@iupui.edu

ABSTRACT
For information workers, maintaining high productivity relies on timely task resumption after interruptions, which are frequent. However, people's task resumption ability is compromised by disruptive environments and human cognitive limitations. We propose that a helpful intervention is to provide visual feedback about the duration of a suspended task. Results from a controlled study show significantly shorter average off-task time by adding visual feedback. Further, using emotionally attachable visual objects in the visualization has the additional benefit of decreasing off-task time without increasing stress. Punishment-oriented persuasion strategies produced faster resumption, but also caused higher stress levels. Combined with other related results, we discuss the implications for design as well directions for future study.

Author Keywords
Interruptions; task resumption; visual feedback; persuasive computing;

ACM Classification Keywords
H.4.1. [Information Systems Applications]: Office Automation - Time Management; H.5.2. [Information Interfaces and Presentation]: User Interfaces — Theory and Methods

General Terms
Human Factor; Design

INTRODUCTION

Information exchange and task coordination are often achieved by group members interrupting each other through face-to-face or computer-mediated communication. At any moment of work, among all the tasks a person is involved in, there typically is a set of primary tasks that are of higher priority and demand more cognitive control and attentiveness to produce meaningful results. To maintain workflow, workers need to resume primary tasks in a timely fashion after an interruption, whether from a team member or self-initiated. However, task resumption is becoming increasingly difficult given the fragmentation of work [19]. For example, O'Connaill and Frohlich [39] found that 41% of the time interrupted workers failed to resume their original tasks. For group work, failure to resume tasks is not only directly attributed to unmanaged group interruptions, but also in turn impedes the completion of cooperative work. After all, it is these individuals who need to manage his/her own share of work and keep up with schedules in order to collectively advance the group's work. Thus, a successful and efficient cooperative environment should also incorporate assistive techniques that maintain personal working efficiency within the inevitable interruptions associated with group work.

To study potential interventions, the purpose of this study is to test whether providing visual feedback regarding the duration of task suspension can improve primary task resumption. Providing dynamic visual cues indicating the length of the task suspension helps a worker become more aware of it and its implications. The visual feedback serves two primary purposes: 1) reminding the worker of the existence of the suspended task and its status and 2) motivating an immediate resumption of the suspended task.

CSCW'14, February 15 - 19 2014, Baltimore, MD, USA
Copyright 2014 ACM 978-1-4503-2540-0/14/02…$15.00.
http://dx.doi.org/10.1145/2531602.2531710

RELATED WORK

Interruption at workplaces

Researchers of communication and workplace interruptions have been studying their negative effects for many years [12, 36]. Not only do interruptions create difficulties for maintaining task flow, they also cause stress and frustration [36]. Constant interruptions often lead to discontinuity of primary tasks [39] and procrastination [27]. Existing intervention methods in group work contexts include minimizing occurrences of interruptions by either controlling notifications [24] or broadcasting the individual's availability in order to regulate mutual interruptions between group members [13]. However, filtering or delaying the arrival of interruptions could result in missing important messages and losing the optimal timing for dealing with the interruptions.

In addition, researchers also identify an equally if not more significant type of interruption: self-initiated interruptions

[14]. Self-interruptions are initiated by internal causes rather than explicit external sources. A previous study indicates that self-interruptions are more disruptive than external ones [14]. Since self-interruptions do not have explicit sources, it becomes very difficult to suppress them by managing the sources, which are inside the human mind.

Thus it is proposed that an alternative approach is to help task resumption regardless of its source. Previous designs in supporting task resumption focus on preserving and resuming context of the original task to enhance goal retrieval from memory [15, 18]. These attempts try to ease the transition by providing some kind of memory aid. Though it is important to support goal retrieval, it neglects to address other relevant problems behind failed resumption: missing the resumption cues and the lack of motivation for resumption. These issues apply equally to external and self-interruptions.

Failure to resume at the right time

Common to both external and self-interruptions, a critical moment in an interruption cycle is when the interrupting task is completed, or its urgency has been contained, and thus the person is supposed to return to primary tasks. However, numerous factors influence a worker's decision making at this moment, including environmental characteristics and individuals' cognitive limitations.

Subtle Cues

Firstly, the opportune timing for resumption is usually not well defined. For example, a person browsing Facebook feeds often does not have a clear cue signaling the proper end of the activity. On the other hand, ending a conversation and hanging-up the phone clearly signals the end of a phone call interruption. In cases like Facebook browsing, it often depends on the person to be aware of the time to end that task. Different people have different ways of defining how much time is too much for deferring task resumption. In other cases, interruptions are cascaded in such a way that resumption cues for previous interruptions are shadowed by new interruptions and old interruptions are never returned to [35].

Ego Depletion

Secondly, people need to be sufficiently alert to cues and fully motivated to control their behaviors to complete the resumption. Research in self-control indicates that this critical cognitive resource for self-regulation diminishes over time, which is coined as "Ego Depletion" by Baumeister and colleagues [1]. Inzlicht and Schmeichel [23] further reveals that lack of motivation and dampening of attention are the culprits for the failure of self-control. When people are less motivated to exert self-control, they pay more attention to cues that meet their predominant response tendencies. With low motivation and divided attention, especially at later stages of task execution, information workers become more susceptible to impulsive tendencies to go off-task.

Disruptive Environment

Thirdly, working environments nowadays increasingly enable self-indulgence, which does a disservice to self-regulation. Part of the reason why resumption cues are too subtle to catch can be attributed to the principles of modern information system design, which emphasize capturing user attention and indulging continuous browsing [38]. Techniques like push notifications, shortcuts, and related links all try to encourage frequent visits and extended stays on these information services, which in many cases are entirely off-task. Group working environments such as open cubicles also result in a higher interruption rate [14].

Visual Feedback

Based on these previous findings, a feedback mechanism is needed to boost self-regulation when self-control strength is low. Visual feedback works as a "discrepancy-reducer" in self-control by decreasing the gap between one's self-control goals and actual behaviors [5]. In order to detect such discrepancies, a monitoring process is key to successful control [7]. Similar to the Cybernetic model in engineering [42], this monitoring process operates on a negative feedback loop in which feedback is given in proportion to the discrepancy. A monitoring lapse of neglecting a goal often leads to failure in self-control [29]. Thus it is predicted that using visual feedback that highlights the goal and discrepancy could enhance this monitoring process and improve self-control behaviors.

Emotionally-Attachable Objects as Motivators

Many persuasive devices or applications implement an emotional motivator, referred to as an *emotionally-attachable object*. For example, UbiGarden [11] presents an aesthetically pleasing display of a growing garden to represent the amount and variety of the user's exercise activities. Fish'n'Steps [34] encourages physical activity using an animated visualization of a fish in a tank, which also includes other players' fish. Designs in persuasive computing use similar objects as a way of encouraging target behaviors [8, 32]. To mitigate motivation loss, amiable visual feedback encourages users to respond to an unfavorable state of the object with the motivation to return it to a favorable state, hence serving as a motivation for immediate task resumption.

Persuasion Strategies

Another factor concerns the persuasion strategy for motivation. Operant Conditioning theory in behavioral change uses the general category of punishment" and "reward" as two basic persuasion strategies [41]. Punishment emphasizes negative feedback for unfavorable behaviors while reward emphasizes giving positive feedback for favorable behaviors. This study tests which of the two general strategies could have a larger effect on task resumption.

METHODOLOGY

Experiment Setting

To test the effect of visual feedback on task resumption, we implemented two forms of computer animations as visual stimuli: a dynamically changing flower (rose) and a traditional horizontal progress bar (See Figure 1). With a total of 1024 frames, the flower animation changed from being fully bloomed to withered and dying. To establish direct comparison, the progress bar also updated in 1/1024 increments. Both flower and progress bar start at 2/3 of full scale. Both visualizations occupied the full screen on an extended display attached to the primary working display. For the control group, the extended display showed an empty black screen. Participants were told to only operate on the primary display, leaving the extended display unoccupied. We enforced the display to be full screen on the extended monitor in order be consistent across groups.

Figure 1. Sample Flower Frames and Progress Bar

Our choice of using a blooming/withering flower is inspired by and partly in accordance with design strategies put forward by Consolvo et al [10]. A blooming/withering flower satisfies most of their design strategies such as Aesthetic, Unobtrusive, Abstract/Reflective and Controllable. The principle of Positive is tested as an independent variable (Persuasion Strategy) while Trending/Historical and Comprehensive are not fully implemented. The deviation from some of these principles is due to the lab-based nature of this study. Moreover, Consolvo et al's design strategies are drawn from projects targeted at changing long-time lifestyles, e.g. physical activity, while the goal of our tool is more interested in correcting short-time task management behaviors.

Using the Windows API, we developed a logging tool that constantly monitored the topmost window in the participant's primary display. This was used as an indication of what the participant was looking at any moment and determined whether the participant was on or off task at that moment.

With two visualizations (Flower and Progress Bar) and two persuasion strategies (Punishment and Reward), this experiment has 2x2 factorial design plus the Empty condition as a control group, receiving no visual feedback and consequently no persuasion strategies. In the "Punishment" condition, the longer a person was deviating from the windows that were related to primary tasks (such as a Microsoft Word document), the more withered the flower appeared, or the less full the progress bar became, depending on the visualization condition. In the reward condition, the longer a user stayed on windows that are related to primary tasks, the more the flower blossomed or the more the progress bar became full. The transition of direction in both visualizations was instantaneous: the moment a user changed the topmost window from primary tasks to work-unrelated task, or vice versa, the direction changed to reflect this.

The punishment and reward strategies were manipulated by adjusting the change rate of the visualizations, as described in Table 1. Fast animation changed at the rate of 4 frames per second while slow animation changed at the rate of 2 frames per second.

Visualization	Punishment	Reward
Flower	Slowly blossom/ Fast decay	Fast blossom/ Slowly decay
Progress Bar	Slowly grow/ Fast reduce	Fast grow/ Slowly reduce

Table 1. Implementation of Persuasion Strategies

Calibration on the Flower Frames

The actual difference between two neighboring frames in the flower visualization is almost undetectable from a human eye. However, given the change rate described earlier, over 2~3 seconds the difference should be obvious to participants.

To have at least a roughly linear representation of flower florescence, we calibrated the flower image frames based on florescence ratings from a pilot study. We selected nine equally-spaced frames from the flower visualization with each representing 10% increment for the whole range of florescence from 10% blossomed to 90% blossomed. We put these images on Amazon Mechanical Turk (AMT) along with images representing the two extreme states of flowering: the fully withered and the fully blossomed, labeled with a score of 0 and 10 respectively. We asked the participants on AMT to rate the 9 randomly presented frames in terms of their level of florescence in comparison with the two extremes. In return, 52 participants from AMT responded with their ratings. The mean values are shown in Table 2. Though the perceived florescence did not follow the spacing of the frames perfectly, only the midpoint (5/10) was substantially higher than anticipated.

Actual Frame	**1/10**	**2/10**	**3/10**	**4/10**	**5/10**
Mean Ratings	1.87	2.85	3.04	4.62	6.23
Actual Frame	**6/10**	**7/10**	**8/10**	**9/10**	
Mean Ratings	6.65	7.58	7.88	8.56	

Table 2. Florescence Ratings from AMT Study

Hypotheses

Our first set of hypotheses was focused on studying differences caused by visualizations and persuasion strategies on time spent on off-task activities:

- H1: Participants with visual feedback (Flower or Progress Bar) will spend shorter off-task time than participants without visualization.
- H2: Participants with the Flower will spend shorter off-task time than participants with the Progress Bar visualization.

We were also interested in whether participants would perform more voluntary switching behaviors without visual feedback. Voluntary switching activities are task switching that are driven by their own volition, including self-interruptions and subsequent voluntary switching after the initial interruption.

- H3: Participants with visual feedback (Flower or Progress Bar) will have less voluntary switching than participants without visual feedback.

We also wanted to observe depletion effects in this study as predicted by Ego Depletion theory, thus:

- H4: Participants will have slower task resumption in later stages of the experimental session.

Finally, we were also interested in whether different visualizations or persuasion strategies would cause more psychological effects such as stress, thus:

- H5: Participants with visual feedback (Flower or Progress Bar) will have a higher subjective stress level than participants without visual feedback.
- H6: Participants receiving Punishment feedback will have higher subjective stress levels than participants receiving Reward feedback.

Participants

30 participants were recruited from a Midwestern university campus via online and in-person solicitation. Participants were randomly assigned into three visualization groups with ten in each group. Each participant in the intervention groups (Flower or Progress Bar) was randomly assigned to either the Punishment or Reward persuasion strategy, resulting in five Punishment recipients and five Reward recipients in each of the intervention groups. This allowed us to compare between five independent groups: Empty, Flower/Punishment, Flower/Reward, Progress Bar/Punishment and Progress Bar/Reward. Each participant was compensated with a $10 Amazon gift card at the end of the study.

Procedure

Framing

It is known that people will behave differently when they know they are being tested [6]. To increase our chances of capturing the behaviors of interest, despite the inherent limitations of a laboratory setting, and to cover our true intent of studying interruption behaviors, we framed our study as an evaluation of new office working environments, in which the participant would perform a typical office task: editing Microsoft Word documents.

After reading and signing the consent form, participants were asked to fill out a questionnaire collecting their demographic information and social media/computer usage. The facilitator then introduced several "new office setting elements" to the participant, including green plants, a new type of office chair, some newly designed lamps, and a seat beside a large window in order to provide an outdoor view and natural light (Figure 2). Participants in the intervention groups were also introduced to the visualization intervention (flower or progress bar). The facilitator introduced each of these office setting elements with equal emphasis and in random order. This design intended to minimize participants' suspicion toward the visualization as the tool under evaluation.

The monitoring application detects inputs from the mouse and keyboard and considers any input as legitimate as long as the Microsoft Word windows stay on the topmost position. This knowledge, if shared with participants beforehand, could result in participants abusing the system by just randomly using the keyboard/mouse to fake productivity. Therefore we did not tell them these operational details beforehand, but only told them that the visualization changes based on their progress on the primary task.

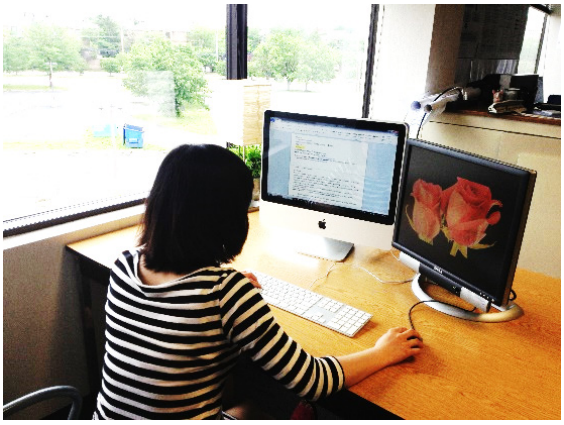

Figure 2. Experiment Setting

Task

The primary task was editing a Word document continuously for 40 minutes. The document was ill-formatted with inconsistent fonts, unaligned paragraphs, messy tables and other features requiring editing. They were given four such documents and asked to fill the time by performing the editing task, but not to be concerned about making a specific amount of progress. Participants were told that this task is of low importance and has no deadline. Each document was at least 20 pages long and they were not expected to finish any of four documents.

They were reminded that neither editing quality nor speed was important to the study. They were encouraged to relax and perform the task at the rate and in the way they feel comfortable. They were told that no audio or video recording devices were implemented in the room and that they were allowed to perform any habit they may have when doing computer-based tasks, e.g. listening to music, checking emails, etc. All these instructions aimed to lower participants' awareness of being in an experiment and encourage spontaneous behaviors. Participants were still reminded that the editing task is the primary goal during the experiment and relinquishing the editing task is not appreciated.

Interruptions

Mimicking a group work environment, distracting interruptions were sent over Google Chat in the form of task requests. These external interruptions forced a baseline level of interruptions in the task in addition to participants' voluntary interruption behaviors such as self-interruptions and delayed resumptions. Typical requests included "Please go check your Facebook feeds and follow any links or stories for as long as you like" and "Please go check new uploads in your YouTube subscription and watch any video you like." All these interruption destinations (e.g. Facebook and YouTube) were drawn randomly from the participants' answers to the question regarding their most frequently visited websites in the pre-study survey. Participants were told that they did not have to reply to the facilitator upon completing the requests. All subsequent actions were of their own volition. This was to eliminate the possibility that replying in the chat box would constitute an additional resumption cue for the participants.

The logging program continuously broadcasts the participant's window changing activities to a web service by sending out the window titles (not the actual content of the windows). The facilitator in another room then was able to monitor participants' window-switching activities via a browser-based client. Based on moment-by-moment updates, the facilitator decided when to send the next request according to the following rules: 1) if the participant is on the editing task, wait for five minutes and send a request; 2) if the participant is off task, do not send requests; 3) when the participant resumes the editing task, reset the timer and wait for another five minutes to send the next interruption; 4) if the participant self-interrupts and goes off task before the five minute timer is up, stop the timer and wait for the participant to resume on his/her own.

Both the logging and monitoring program were not revealed to the participant until the end of the study. At the end of the study we had signed consent from participants to allow us to use the log file after the file was reviewed by the participants. No participants requested any deletion of entries from the log files.

Pre- and post-tests

Before the editing task, each participant went through a five-minute digit-span test. A digit-span test asks subjects to repeat a sequence of digits immediately after seeing the sequence. The digit-span tests were used to deliberately deplete participants' cognitive resources before entering the actual experiment. As adopted in Ego Depletion studies [31], this technique is used to make the depletion effect more detectable.

After the editing task, participants were asked to fill out a modified NASA TLX [20], with one additional question about self-reported stress level, as used by Mark and colleagues [36]. Participants also rated how much they liked the visualization after the study. The study concluded with a semi-structured interview about the participant's experience using the tool, their motivations behind interruption behaviors, and their general impression toward the visualizations.

DATA COLLECTION

The logging tool captures all the window-switching activities and writes all activities into a file. Each entry contains the timestamp of the switching as well as the title of the current topmost window. This log file provides the ability to determine if a participant is on or off of the editing task at any moment of the study. Two measures were extracted from the log files as dependent variables:

Average Off-task Time (AOT): For each participant, we computed the average duration of off-task time as a measure of primary task resumption speed. This quantity was calculated by dividing their total off-task time by the number of window-switches away from Microsoft Word to other windows. Incidents where the participant peeked at other windows and switched right back to the editing task within 10 seconds were excluded. This is based on the observation from the data that a task deviation that lasts shorter than 10 seconds usually ends at the Google Chat window. This suggested that such deviations were intended at checking for new requests, rather than leaving the editing task. AOT is preferable to using total off-task time as a measure of resumption speed because the total duration of the experiment varied from participant to participant due to the off-task time being open-ended. The five minute limit for on-task time before an interruption created a floor effect where even the most efficient workers suffered a baseline penalty without being rewarded for remaining on-task for the full duration.

Voluntary Switching (VS): For each participant, we counted the incidents of window switching that were driven by their own volition. Specifically, we counted the number of different webpages a participant visited other than the websites requested by the facilitator. This includes all the subsequent visited websites after an initial interruption as well as all the target websites visited by self-initiated interruptions.

RESULTS

Task Switching:

Overall, participants spent on average 45.0% ($M = 0.45$, $SD = 0.15$) of time off tasks which amounts to about 18 minutes of off task time if we normalize total experiment time to 40 minutes. The mean AOT across all groups was 169 ($M = 169.46$, $SD = 111.50$) seconds. On average, participants made 114 ($M = 114.0$, $SD = 82.37$) window switches. Participants made 8.53 ($M = 8.53$, $SD = 3.55$) switches away from the editing task. Of those switches, a mean of 4.86 ($M = 4.86$, $SD = 2.23$) were self-initiated.

Average Off-task Time (AOT)

One-way ANOVA on five between-subject groups revealed significant differences of AOT[1] across the groups ($F(4, 20) = 3.21$, $p < 0.05$). Planned contrasts showed the Empty group ($M = 249.60$, $SE = 43.81$) had significant higher AOT than the rest of the groups ($t(20) = -2.72$, $p < 0.05$), confirming H1 that providing the visualization significant shortens average off-task time compared to the control. Further, Flower group's AOT ($M = 119.03$, $SE = 15.35$) was also significantly shorter than that of Empty group ($t(20) = -2.64$, $p < 0.05$). Progress Bar ($M = 139.76$, $SE = 27.51$) also produced a significantly shorter AOT than the Empty group ($t(20) = -2.32$, $p < 0.05$), but the Flower and Progress Bar did not have a significantly different AOT ($t(20) = 0.40$, *ns*). This rejects H2 and indicates that the Progress Bar and Flower are not significantly different in term of resumption effectiveness. Figure 3 illustrates AOT among the three groups. The fact that having a visualization produced lower AOT is partially supported in interview data. For example, P24 mentioned: "As soon as I see the petals get a little brown edge, I jump back to the Word. I want to keep it normal!"

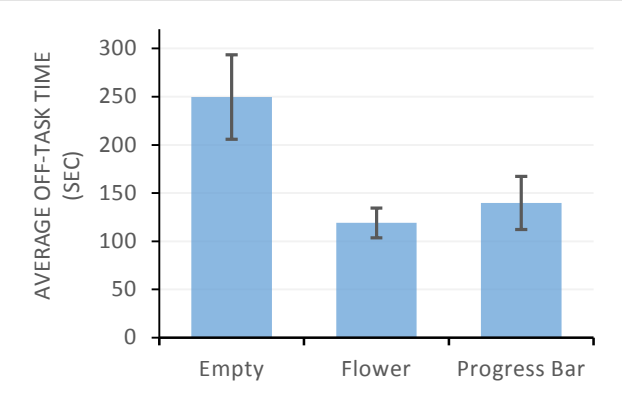

Figure 3. AOT across Visualization Groups

For persuasion strategies in the intervention groups, a two-way Factorial (2x2) ANOVA of both visualizations (excluding Empty) and persuasion strategy revealed a significant main effect of persuasion strategy ($F(2, 16) = 11.04$, $p < 0.05$), but no significant main effect of visualization and no significant interaction effect between the two. Regardless of visualization, Reward ($M = 170.23$, $SE = 21.07$) produced significantly longer AOT than Punishment ($M = 88.56$, $SE = 14.20$). Figure 4 shows the AOT between two strategies.

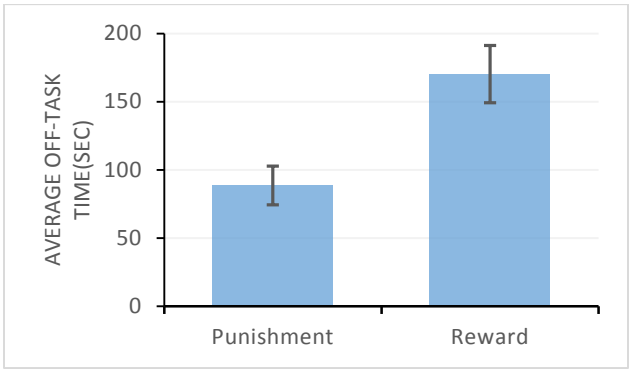

Figure 4. AOT between Persuasion Strategies

Even though there was no significant interaction effect between visualization and persuasion strategy, Figure 5 shows that the Progress Bar is marginally more sensitive to persuasion strategies.

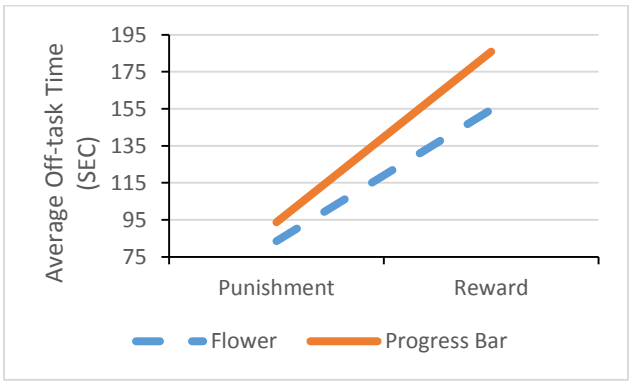

Figure 5. Interaction between Visualization and Persuasion Strategy on AOT

Factorial ANOVA on gender and visualization failed to find a significant main effect of gender or interaction effect of gender and visualizations in terms of average off-task time.

Depletion Effect

To investigate the depletion effect over time, we extracted all off-task time (OT) from log files. We also calculated the time into the study for each entry based on the timestamps. A multiple regression was run to predict OT from time and visualization type. Both variables significantly predicted OT, $F(2, 253) = 5.89$, $p < .005$, and both added significantly to the prediction, $p < .05$. Time's $\beta = 0.26$, meaning that regardless of the visualization type, OT

[1] AOT was log10 transformed for a normal distribution.

increases over time. There was no significant interaction between visualization and time into the study.

Voluntary Switching (VS)

The mean number of voluntary switches was 34.7 (SE = 4.06) for the Flower group, 30.5 (SE = 3.52) for the Progress Bar group, and 42.0 (SE = 4.60) for the Empty Group. In terms of persuasion strategy, the mean VS was 31.3 (SE = 3.61) for the Punishment group and 33.9 (SE = 4.05) for the Reward group. The number of voluntary switches was not significant across the five groups based on one-way ANOVA ($F(4, 20) = 1.86$, *ns*). Excluding the Empty group, a two-way factorial ANOVA of visualization and persuasion strategy also failed to find significant results ($F(3,16) = 0.516$, *ns*). This result rejects H3 and concludes that there is no significant difference in terms of voluntary switching across visualization groups.

Subjective Workload

Subjective Stress

In answering the stress related question in the NASA TLX questionnaire ("How stressed do you feel throughout the task?"), the Flower group produced the lowest subjective stress level (M = 2.40, SE = 0.50), followed by the Empty group (M = 2.70, SE = 0.63), and the Progress Bar group with highest subjective stress level (M = 9.20, SE = 1.78). Non-normal distribution of subjective stress level required nonparametric analysis (Kruskal-Wallis), which indicated there was a significant subjective stress level difference across the three visualization groups ($H(2) = 11.23$, $p < 0.05$). Mann-Whitney tests were used to perform pair-wise comparisons using a Bonferroni corrected significance level of .0167 (3 comparisons). The Flower group's subjective stress level was significantly lower than the Progress Bar group ($U = 12.00$, $p < .0167$). The Empty group's subjective stress was also significantly lower than Progress Bar group ($U = 13.50$, $p < .0167$). The difference of subjective stress levels between the Flower group and Empty group was not significant. This rejected H5 but also shows that the Progress Bar elicited higher stress levels than other two visualizations. Figure 7 shows subjective stress across the three visualization groups.

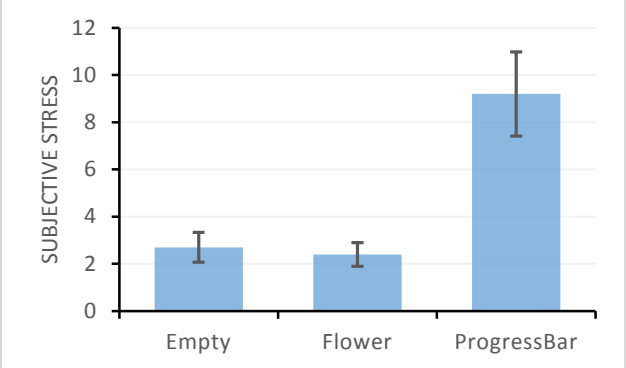

Figure 7. Subjective Stress across Visualization groups

A Mann-Whitney test also revealed that within the intervention groups, the Punishment strategy produced significant higher subjective stress levels (M = 8.20, SE = 1.96) than the Reward strategy (M = 3.40, SE = 0.92; Mann-Whitney's $U = 23.50$, $p < 0.05$). This result confirms H6 that Punishment recipients have higher stress levels than Reward recipients. Figure 8 shows the subjective stress level between persuasion strategies. In our interview, P18 confirmed the result that Punishment produced higher stress level by saying: "I feel like the flower's speed of decaying is faster than recovering and it made me more nervous".

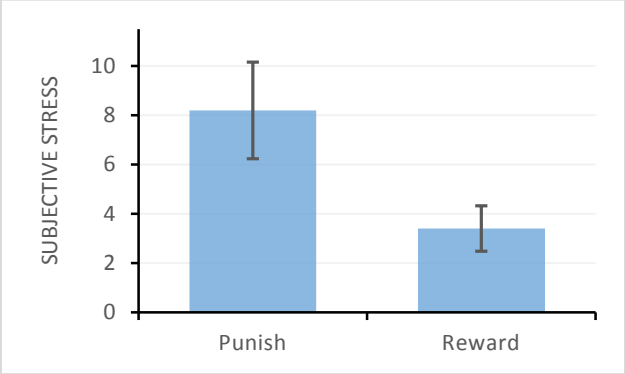

Figure 8. Subjective Stress between Persuasion Strategies

Subjective Time Pressure

For the survey question regarding time pressure ("How hurried were you while doing the task?"), visualization types did not have a significant effect on subjective time pressure. However, persuasion strategy did: Punishment produced a significantly higher self-reported time pressure (M = 9.30, SE = 1.71) than the Reward strategy (M = 2.80, SE = 0.65; Mann-Whitney's $U = 13.50$, $p < 0.05$).

Subjective Likability and Gender

Even though gender and visualization did not have a significant interaction effect on AOT performance, they did have significant interaction effect on self-reported liking of the visualization in post-test survey. Factorial ANOVA revealed a significant interaction effect between visualization and gender ($F(3,16) = 17.74$, $p < 0.05$). Specifically, male participants rated both the Progress Bar (M = 3.14, SE = 0.55) and Flower (M = 3.00, SE = 0.649) similarly, while female participants liked the Flower visualization significantly higher (M = 5.80, SE = 0.65) than the Progress Bar (M = 2.00, SE = 0.84). Participant 19 (Male) confirmed his preference toward the progress bar by saying that "it gives direct feedback and I can see where I am clearly."

DISCUSSION

On Visualization

The significant difference of AOT across visualization groups indicates that providing visual feedback helps

participants return to tasks faster. Even though the Flower group produced a somewhat lower mean AOT than the Progress Bar group, the effect of adding an emotionally attachable element like the flower does not significantly help. On the other hand, the Progress Bar elicited a significantly higher level of subjective stress than the Flower intervention, indicating that in the long run, immediate and direct feedback such as a progress bar could wear people out. It is thus recommended that amiable and emotionally attachable visualizations could achieve a similar level of effectiveness in boosting task resumption while creating less stress on users. However, due to the limitations of a lab-based study, how well these visualizations could support long-term use needs to be further studied in a real workplace environment.

On Persuasion Strategy

The punishment strategy produced significantly faster resumption while also inducing significantly higher subjective stress levels and subjective time pressure. This result shows the double-bladed effect of using the punishment method: it has higher effectiveness but also creates higher pressure on users. The Reward group's lower stress level and time pressure as well as longer off-task time may be explained by some information gathered in the post-test interview. Many Reward recipients mentioned the opportunities of "taking a breath" granted by the Reward intervention. Once they attained a high level with the visualization (almost fully blossomed flower or almost full progress bar), they gave themselves a license to take a break by visiting non-work related websites. Since the reduction of the progress bar or flower is much faster in the Punishment condition, those participants did not perceive themselves as having this license. This "self-licensing" behavior is often seen as a motivation deficit or justification for immediate gratification and over-indulgence [16]. Designing for a relaxing condition like Reward requires vigilance for excessive self-licensing behaviors and must balance the manipulation between encouraging high concentration and allowing mental and physical relaxation. The result from this study also gives evidence as to why the reward strategy could help persuasive technology's continuous use, as predicted by Consolvo et al [10].

On Depletion Effect

As predicted by Ego Depletion theory [1], a depletion effect on AOT was observed over time for all groups. Thus H4 is confirmed. However, no significant interaction was found between the time into the study and the visualization type. Further study is still needed to investigate how visual feedback could influence the depletion effect.

On Voluntary Switching

The analysis on voluntary switching did not yield significant results. This is probably due to the relatively small sample size and limited study time (40 minutes on average). It might also have to do with the fact that some websites are better at grasping user attention and making users stay (e.g. YouTube) than others and that is why the significant results in AOT is not reflected on the number of webpages visited.

On Taking Breaks

It is worth noting that tools such as these that continuously encourage people to get back to work should not prohibit taking breaks during working hours. Numerous studies indicate the benefits of taking breaks at workplaces [21, 26]. Research in attention restoration advocates taking breaks as a way of replenishing attention resources [2]. However, for more and more information workers, taking breaks often means visiting non-work-related websites, which tend not to replenish cognitive resources, compared to restorative activities such as taking a stroll in the park [4].

As predicted by previous research, visiting these websites as a way of taking breaks is counterproductive in replenishing cognitive resource. Firstly, these websites are what Kaplan and colleagues refer to as *hard fascinations* that fixate one's attention and disallow any reflection about anything else. Thus they are less effective in attention restoration [30]. On the contrary, *soft fascinations* are natural objects such as flowers, plants, or woods that promote reflection on other things and are validated ways of attention restoration. Secondly, Zeigarnik shows that unfinished issues could persist in memory [44]. These *attention residues* [33] create interferences with normal task execution by disrupting cognitive functioning [3, 28]. It is recommended that visual feedback be integrated with scheduled breaks to produce a sustainable working flow.

On Multitasking Tendencies

Ophir et al [40] aims to debunk the myth of multitasking efficiency in the digital age. Their work recognizes that people are not as good at multitasking as they think they are, especially given the saturation of information in today's world. Their results indicate that heavy media multitaskers are habituated to "explorative" rather than "exploitative" ways of information processing and they are more attracted to off-task stimuli. This means they are more likely to be disrupted by the external environment and more likely to interrupt themselves to explore novel stimuli.

Therefore, based on Ophir et al's results and those learned from this study, an intervention for heavy media multitaskers could aim to limit the extent of this "explorative" way of information processing by providing visual feedback reminding them about the extent and duration of off-task behaviors. This will narrow their attention allocation to more important tasks. In other words, for multitaskers, tools like those in this study serve as a container that prevents irrational time and attention allocation to secondary or irrelevant tasks.

Design Implications

The effectiveness of providing visual feedback is confirmed in this study. The insignificant AOT differences

between the Flower and Progress Bar shows that either dynamic cue can effectively remind and motivate task resumption. In addition, rewarding strategies are potentially more suitable for long time use since they produced lower stress level, as recommended by [10].

Confirming the visualization likability results, our interview data identified a theme of male participants reacting rather lukewarm toward the Flower visualizations. This indicates potential individual differences in terms of acceptance of different visualizations. Previous studies confirmed that in persuasive computing, designers should consider gender as a factor in persuasive design [43]. For long-term use in real-life work environments, different people will respond differently to visualizations as motivators [10]. It is important to match individual differences to maximize the motivation effect, or offer customizations that improve effectiveness.

A higher-level design suggestion is to recognize the differences between two modes of operating a computer: recreational and serious working. System designers should take into account the human cognition characteristics under the two modes and provide support accordingly.

LIMITATION AND FUTURE WORK

Task Urgency

One of the side-effects of this framing was misrepresenting real-life behaviors. 13 out of the 30 participants mentioned in their interview that if they knew the task was urgent, they would have spent much less time on irrelevant tasks and focus more on the editing task. Indeed, task urgency can increase rate of task completion [9]. Thus the results of this study do not directly apply to contexts with high task urgency, whose effects on motivation for task completion could be large enough such that visual feedback is unnecessary. Further studies are needed to explore how external factors such as task urgency would enhance or diminish the effect of visual feedback on the rate of task resumption.

On the other side, because of the large effect of task urgency on task resumption and in order to prevent task urgency from masking the effect caused by visual feedback, we had to bring every participant down to a similar low-level task urgency in order to observe how visual feedback alone could influence the resumption efficiency.

Moreover, the results of this study still have validity in certain circumstances. First, workers whose tasks are not deadline-driven depend more on self-regulation capacity than task characteristics to manage interruptions and resumptions. Second, claiming that situational differences would alter their behaviors does not stand as an excuse for self-control failure in general non-urgent scenarios, which characterizes much of routine office work. Lastly, delaying a task to the last minute to increase its urgency is a common type of procrastination behavior identified by researchers as *Arousal Procrastination* [17], which is an extreme example of over-dependence on deadlines.

Complex Real-life Interruptions

We recognize the complexity of real-life interruptions and their formation [35]. In particular, interruptions are often interleaved and the setting of primary tasks often changes over time. This study demonstrated the efficacy of visual feedback in a simplified context and may not be fully replicable when faced with the complexity in real-scenarios.

However, we can still see the potential even with this complexity: as long as there is some relative difference of importance among tasks, more sophisticated tools can be built which intelligently learn about the environment and adjust the settings to accommodate this complexity. Examples include dynamically detecting the primary tasks or pausing the feedback upon important immediate interruptions. Other projects demonstrate the potential of systems that are able to learn about working environments such as [22] and [25].

Sources of Interruptions

Real-life interruptions are not always coming from chat-based communication like this study. As described earlier, some interruptions have explicit cues for resumption while others do not. Thus some interruptions are inherently easier to conclude, e.g. phone call vs. visiting Facebook. Thus the effect of visual feedback should be more pronounced for wrapping up interruptions that are more open-ended.

Limited Sample Size

The limited number of participants prevented us from arriving at other interesting conclusions such as gender differences, voluntary switching behaviors, and the effect on mental workload. These factors remain to be investigated in future studies with larger sample sizes.

Future Work

Follow-up studies are applying field research methods and investigating the usage of such tools in real-life settings. A diary-based study recording daily use of the tool will enable a deeper understanding of the problems of interruptions in collaborative work settings and produce robust designs for addressing them. Natural settings give us opportunities to investigate contextual influences on task resumptions. Such context variables could be: task urgency, task completion level, time of day, nature of the task, and others.

In regard to supporting group motivation, this tool also could be extended to involve multiple group members competing with each other by sharing each other's visualizations. Such effort in gamifying working spaces has seen positive results [37].

CONCLUSION

Disruptive environments and cognition limitations delay timely task resumption after interruptions. The results of this study suggest that the use of visual feedback on task

suspension can function as an effective intervention to help people resume tasks more gracefully and efficiently. Results also show the efficacy of using emotionally attachable objects of both reducing off-task time and controlling stress level. However, further studies are needed to address individual differences on responsiveness to visual persuasion to maximize its effectiveness. Lastly, future work must investigate the effectiveness of this intervention on long-term use and whether it is best used on an ongoing basis, or if it may function most effectively as a tool to train more productive behaviors, and then be discontinued.

ACKNOWLEDGEMENT

The authors thank Michael Durham (www.durmphoto.com) for allowing the use of his time-lapse rose photography in this study.

REFERENCES

1. Baumeister, R. F., et al. Ego depletion: is the active self a limited resource? *Journal of personality and social psychology*, *74*, 5 (1998), 1252.
2. Berman, M. G., et al. The cognitive benefits of interacting with nature. *Psychological science*, *19*, 12 (2008), 1207-1212.
3. Berman, M. G., et al. In search of decay in verbal short-term memory. *Journal of experimental psychology. Learning, memory, and cognition*, *35*, 2 (2009), 317.
4. Bock, G. W. and Ho, S. L. Non-work related computing (NWRC). *Communications of the ACM*, *52*, 4 (2009), 124-128.
5. Botvinick, M. M., et al. Conflict monitoring and cognitive control. *Psychological review*, *108*, 3 (2001), 624.
6. Carlopio, J. *The Hawthorne effect awareness of experimental participation or experimental demands?* , National Library of Canada, Ottawa, 1983.
7. Carver, C. S. and Scheier, M. *Attention and self-regulation: A control-theory approach to human behavior*. Springer-Verlag New York, 1981.
8. Chiu, M.-C., et al. Playful bottle: a mobile social persuasion system to motivate healthy water intake. In *Proc. Proceedings of the 11th international conference on Ubiquitous computing*, ACM (2009), 185-194.
9. Claessens, B. J. C., et al. Things to Do Today . . . : A Daily Diary Study on Task Completion at Work. *Applied Psychology*, *59*, 2 (2010), 273-295.
10. Consolvo, S., et al. Theory-driven design strategies for technologies that support behavior change in everyday life. In *Proc. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2009), 405-414.
11. Consolvo, S., et al. Activity sensing in the wild: a field trial of ubifit garden. In *Proc. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2008), 1797-1806.
12. Czerwinski, M., et al. A diary study of task switching and interruptions. In *Proc. Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2004), 175-182.
13. Dabbish, L. and Kraut, R. Coordinating communication: Awareness displays and interruption. In *Proc. CHI'03 extended abstracts on Human factors in computing systems*, ACM (2003), 786-787.
14. Dabbish, L., et al. Why do I keep interrupting myself?: Environment, habit and self-interruption. In *Proc. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2011), 3127-3130.
15. Daniels, J. J., et al. Support for intelligent interruption and augmented context recovery. In *Proc. Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants.*, IEEE (2002), 7-15-17-21.
16. De Witt Huberts, J. C., et al. License to sin: Self-licensing as a mechanism underlying hedonic consumption. *European Journal of Social Psychology*, *42*, 4 (2012), 490-496.
17. Ferrari, J. R. Psychometric validation of two procrastination inventories for adults: Arousal and avoidance measures. *J Psychopathol Behav*, *14*, 2 (1992), 97-110.
18. Franke, J. L., et al. Recovering context after interruption. In *Proc. Proceedings 24th Annual Meeting of the Cognitive Science Society* (2002), 310-315.
19. González, V. M. and Mark, G. *"Constant, constant, multi-tasking craziness": managing multiple working spheres.* ACM, City, 2004.
20. Hart, S. G. and Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human mental workload*, *1* (1988), 139-183.
21. Henning, R. A., et al. Frequent short rest breaks from computer work: effects on productivity and well-being at two field sites. *Ergonomics*, *40*, 1 (1997/01/01 1997), 78-91.
22. Horvitz, E., et al. Attention-sensitive alerting. In *Proc. Proceedings of the 15th conference on Uncertainty in artificial intelligence.* (1999), 305-313.
23. Inzlicht, M. and Schmeichel, B. J. What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, *7*, 5 (2012), 450-463.

24. Iqbal, S. T. and Bailey, B. P. Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption. In *Proc. CHI '05 extended abstracts on Human factors in computing systems*, ACM (2005), 1489-1492.
25. Iqbal, S. T. and Bailey, B. P. Oasis: A framework for linking notification delivery to the perceptual structure of goal-directed tasks. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *17*, 4 (2010), 15.
26. Jett, Q. R. and George, J. M. Work interrupted: A closer look at the role of interruptions in organizational life. *Academy of Management Review*, *28*, 3 (2003), 494-509.
27. Jin, J. and Dabbish, L. A. Self-interruption on the computer: a typology of discretionary task interleaving. In *Proc. Proceedings of the 27th international conference on Human factors in computing systems*, ACM (2009), 1799-1808.
28. Jonides, J., et al. The Mind and Brain of Short-Term Memory. *Annual Review of Psychology*, *59*, 1 (2008), 193-224.
29. Kane, M. J. and Engle, R. W. Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology-General*, *132*, 1 (2003), 47-70.
30. Kaplan, S. The restorative benefits of nature: Toward an integrative framework. *Journal of environmental psychology*, *15*, 3 (1995), 169-182.
31. Kaplan, S. and Berman, M. G. Directed attention as a common resource for executive functioning and self-regulation. *Perspectives on Psychological Science*, *5*, 1 (2010), 43-57.
32. Kim, T., et al. Coralog: use-aware visualization connecting human micro-activities to environmental change. In *Proc. CHI'09 Extended Abstracts on Human Factors in Computing Systems*, ACM (2009), 4303-4308.
33. Leroy, S. Why is it so hard to do my work? The challenge of attention residue when switching between work tasks. *Organizational Behavior and Human Decision Processes*, *109*, 2 (2009), 168-181.
34. Lin, J. J., et al. Fish'n'Steps: Encouraging physical activity with an interactive computer game. In *Proc. UbiComp 2006: Ubiquitous Computing*, Springer (2006), 261-278.
35. Mark, G., et al. No task left behind?: Examining the nature of fragmented work. In *Proc. Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2005), 321-330.
36. Mark, G., et al. The cost of interrupted work: more speed and stress. In *Proc. Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, ACM (2008), 107–110.
37. Neil, D., et al. Limber: exploring motivation in a workplace exergame. In *Proc. Proceedings of the 2013 conference on Computer supported cooperative work companion*, ACM (2013), 239-242.
38. Nieson, J. Scrolling and Attention.(2010), http://www.nngroup.com/articles/scrolling-and-attention/
39. O'Conaill, B. and Frohlich, D. Timespace in the workplace: Dealing with interruptions. In *Proc. Conference companion on Human factors in computing systems*, ACM (1995), 262-263.
40. Ophir, E., et al. Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 37 (Sep 15 2009), 15583-15587.
41. Skinner, B. F. Two types of conditioned reflex: a reply to Konorski and Miller. *J. Gen. Psychology*, *17*, 272 (1937), 79.
42. Wiener, N. Cybernetics: or Control and Communication in the Animal and the Machine (1948).
43. Zanbaka, C., et al. Can a virtual cat persuade you?: the role of gender and realism in speaker persuasiveness. In *Proc. Proceedings of the SIGCHI conference on Human Factors in computing systems*, ACM (2006), 1153-1162.
44. Zeigarnik, B. Über das Behalten von erledigten und unerledigten Handlungen. [Remembering completed and uncompleted tasks]. *Psychologische Forschung*, *9*, 1 (1927), 1-85.