Restoring the Context of Interrupted Work with Desktop Thumbnails

Adam Rule, Aurélien Tabard, Karen Boyd, Jim Hollan

Restoring the Context of Interrupted Work with Desktop Thumbnails

Adam Rule, Karen Boyd, Jim Hollan (acrule, hollan@ucsd.edu, karenboyd@gmail.com)
Department of Cognitive Science, UC San Diego, 9500 Gilman Drive, San Diego, CA 92093 USA

Aurélien Tabard (aurelien.tabard@univ-lyon1.fr)
Université de Lyon, CNRS
Université Lyon 1, LIRIS, UMR5205, F-69622, France

Abstract

Knowledge work is frequently interrupted. Interruptions enable collaboration and bring timely information, but they disrupt the fragile context of ongoing activities. Computers, now ubiquitous in knowledge work, have improved in their ability to track and restore digital context (documents and files), but they do a poor job of helping users restore mental context: the ideas, intentions, and motivations behind their work. Thumbnail images are an efficient way to help computer users re-find documents; we ask if they can also be used to restore mental context. We tested how three manipulations to thumbnails of personal computer screenshots impact their ability to help viewers recognize past activities and recall accurate and detailed context. In a 2-week study we found that thumbnails of portions of the screen need to be larger than thumbnails of the entire screen for successful activity recognition and that static screenshots prompted more accurate contextual recall than animations.

Keywords: interruptions; memory; thumbnails; activity-based computing

Introduction

Knowledge work is frequently interrupted. One 13-month study found that knowledge workers switch activities once every 12 minutes (e.g. from writing a report to calling a customer) and that the majority of these activities (57%) are interrupted (Mark, Gonzalez & Harris, 2005). This fragmentation is not necessarily bad. Interruptions bring timely information, foster collaboration, and direct attention to interesting or urgent work. But resuming interrupted activities takes time and energy. Early research estimated that programmers spend 15 minutes after each interruption "regaining concentration" (Solingen, Berghout & Latum 1998). Another field study by Iqbal and Horvitz (2007) confirmed that people spend a significant amount of time restoring their mental state after re-finding relevant papers and documents. Likewise, knowledge workers find it harder to restart interrupted tasks than to start new ones (Czerwinski, Horvitz & Wilhite, 2004).

Resuming knowledge work, whose non-routine problem solving activities routinely involves manipulating collections of information in external artifacts and internal memory, takes cognitive effort. While interruptions can make artifacts hard to find, physically or digitally obscuring them under the artifacts of intervening tasks, they are particularly disruptive to the contents of working memory. These memories get harder to reactivate the longer an interruption lasts (Altmann & Trafton, 2002) and their retrieval can be "blocked" if the interrupting activity is sufficiently similar to the suspended one (Gillie & Broadbent, 1989). We refer to these memories as an activity's mental context. This context can be quite complex, especially for dynamic, creative, and non-linear activities like data analysis, programming, and writing. It often includes information that is difficult to represent externally or ephemeral such as intentions ("I'll call Bill next"), motivations ("this letter needs to sound more professional"), and ideas ("this graph may look better on a log scale").

Activity-Based Computing

Knowledge work increasingly involves computers, but computers have historically done a poor job supporting interrupted and interleaved work (Bannon et al., 1983). In the realm of restoring artifacts, switching tasks on a computer can take a flurry of finding, opening, and rearranging windows. Activity-Based Computing has attempted to address this issue by making it easier to manipulate groups of documents (Bardram, Bunde-Pedersen & Soegaard, 2006). The paradigm can trace its roots back to Henderson and Card's *Rooms* work in 1986 and has seen a resurgence of interest in recent years (Dumais et al., 2003; Kaptelinin 2003; Karger et al., 2005; Rattenbury & Canny 2007) but has yet to have a significant presence in mainstream operating systems. Regardless, this research only addresses half the problem. Computers provide even less support for what is often the harder part of restarting an interrupted activity: restoring mental context (Iqbal & Horvitz, 2007).

Visual Memory

One promising approach to restoring mental context was demonstrated by Cangiano et al. (2009) who found that people who watched screen-recordings of past work were able to remember contextual details such as why they were working on a particular project or who they were talking to at the time. Humans have excellent visual memory. Participants shown 2560 images for 10 seconds each were able to correctly select which of two images they had seen before 90% of the time when tested two days later (Standing et al., 1970). A more recent study found that subjects were able to correctly identify the image they had seen before 87% of the time when the only difference between the images was the pose of an object, such as the arrangement of beads on an abacus (Brady et al., 2008).

Cangiano et al. demonstrated that humans can do more than discriminate; they can also attach complex meaning to images, even passively taken ones. Sellen et al. (2007) confirmed this result, finding that images taken automatically by a wearable camera cued as detailed of episodic memories as images taken by actively pressing a button. Passively recording and then visualizing knowledge work may ease the burden of resuming interrupted activities by helping people restore mental context that was never externalized in artifacts.

Thumbnails

Research on visualizing and recognizing computer activity has focused on website thumbnails. Kaasten, Greenberg & Edwards (2002) showed that people could reliably (80% of the time) identify a website domain they had previously visited (e.g. www.cnn.com) using a 132 x 132px thumbnail image of that website and could identify its exact topic (e.g. a CNN article on the 2008 election) with a 208 x 208px thumbnail. Moreover, Teevan et al. (2009) found that people were able to re-find previously visited websites more quickly using thumbnails than page titles.

But, these studies are limited to web-browsers and the recognition of well-defined pages. While useful, they do not address the complex nature of real-world work that typically involves not only multiple applications but also dynamic, creative, and non-linear activities. Our work expands on this research by asking if thumbnails can be used to represent not only documents, but also whole activities and their mental context.

Research Question & Hypotheses

Q: How do the cropping, animation and rehearsal of a thumbnail showing a computer desktop impact its ability to help people recognize past activities and recall accurate and detailed context.

Cropping

Small thumbnails that show a full desktop can be ambiguous; windows are tiny, text is unreadable, and it is hard to see the user's cursor. Keeping the thumbnail the same size but having it show only a small portion of the screen, such as the area around a user's cursor, may support better activity identification as users see one region in detail and may be able to mentally "fill in" the rest from memory (Figure 1).

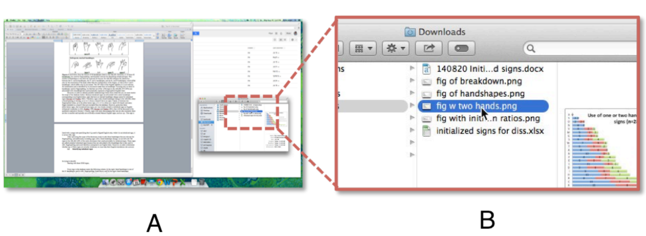

Figure 1: Cropped thumbnails (B) can be the same size as full-screen thumbnails (A) but show less of the screen

H1: People will recognize activities with smaller thumbnails if those thumbnails show a cropped portion of the screen. By forcing people to "fill-in" screen content, cropped thumbnails will cue less accurate but more detailed memories than full-screen thumbnails as people search memory for context rather than read it off the thumbnail.

Animation

By displaying activity over time, animations provide more information than individual screenshots, potentially helping users distinguish similar activities and retrieve more specific memories. However, making sense of animations takes focus and may cause viewers to neglect context held in memory (Tversky, Morrison, & Betrancourt 2002).

H2: Animations will cue more accurate but less detailed memories than screenshots.

Rehearsal

Rehearsing memories makes them easier to recall, but can also change their contents (Hupbach et al., 2008). Asking people to reflect on particular moments may make those moments easier to recall, but can also restrict the number of details they remember about them.

H3: Memories for moments that have been rehearsed will be more accurate but less detailed than those for unrehearsed moments.

Methods

Participants and Materials

Six graduate students (4 female, ages 23-29) recorded their workday (Mon-Fri) computer activity for two weeks using a modified version of Selfspy[1], an open-source key-logger. The computers used for recording included stand-alone laptops, laptops connected to an external monitor, and all-in-one desktops. The tool took a screenshot every time a participant clicked or typed, and every 30 seconds while their computer was awake but idle. All screenshots were stored on an SD card. Participants received $50 in gift cards.

Procedure - Recording

Participants could pause the recording at any time and were instructed to do so whenever they used their computer to communicate with someone outside the study. Every 30 minutes, Selfspy showed the participant a recent screenshot and asked them "What are you doing?" Participants could ignore these experience samples, but were encouraged to fill out as many as possible. At the end of each day, participants were asked to provide additional detail on up to five samples from earlier that day. While viewing each sample's screenshot and textual description participants answered the following questions:

1. What do you remember about this moment?

[1] https://github.com/activityhistory/selfspy

2. How do you know these details? (I remember this moment exactly, I know from experience, or I'm guessing)
3. How well does this image represent what you were doing? (Very well, Somewhat, or Not at all)

These responses gave us a detailed baseline description of what participants were doing at the time of the sample. Work by Brandt, Weiss and Klemmer (2007) has shown that asking users to write short descriptions in the moment and fill in details later reduces the impact of experience sampling interruptions but maintains response quality.

Procedure - Testing thumbnails

Participants attended a one-hour lab session at the end of each week to review up to 40 thumbnails representing moments from the prior week.

The thumbnails varied along three conditions, meaning we tested 8 variations in a 2x2x2 design:

A. Screenshot or animation
B. Full screen or cropped
C. Rehearsed moment (e.g. described in detail at the end of a day earlier that week) or unrehearsed moment (randomly selected)

We tested the four types of screenshot thumbnail during the first week's review (i.e. screenshot condition with BxC). Similar to Kaasten et al. (2002), when the participant pressed a start button, a 20px high thumbnail would appear. (The width of the thumbnail depended on the recording computer's screen ratio). Every two seconds the thumbnail grew 20px taller and proportionately wider until the participant recognized the represented activity and pressed a button to stop the growth. Full screenshots showed the participant's entire screen whereas cropped screenshots showed an expanding area around the user's mouse. Participants then answered the three questions used in the end-of-day debriefs and moved on to the next thumbnail.

We tested the four types of animation thumbnail in the second week. Each animation showed a time-lapse of five minutes of computer activity played at 5x normal speed. Full animations showed the participant's full screen at the native resolution of the recording computer, while cropped animations showed a 520px high area around the participant's cursor location at the time of recording. (520px was the height at which our first four participants could recognize 80% of the activities in their week 1 cropped screenshots). Participants pressed a button to start each animation, pressed a second button once they recognized the represented activity, and then filled out the debrief questions before moving on to the next thumbnail.

Two participants had too few debriefed experience samples in the second week to include their data. In total, we collected 145 responses to screenshots and 70 to animations.

Measures

We measured participant's responses in six ways:

1. Thumbnail size (for static thumbnails) or duration (for animations) at the point of activity recognition
2. Self-rated Memory Strength and Thumbnail Appropriateness
3. Memory Accuracy
4. Episodic Detail
5. Event Specificity
6. Time Discussed

We coded each response for measures 3-6. Memory Accuracy was coded on a scale of how well the activity described in the review matched the activity described in the end-of-day debrief. The levels included 0) no match, 1) partial match, and 2) mostly matches. Accuracy could only be coded for rehearsed thumbnails since unrehearsed thumbnails had no end-of-day debrief responses for comparison. For Episodic Detail, we recorded the number of contextual details shared about an activity including who, what, where, when, why and feeling information. For Event Specificity, we recorded the number of events described in the response at the Action (i.e. "copy and pasting"), Activity ("editing this discussion slide"), or Project ("working on my presentation") level. Lastly, we coded whether the text mentioned events that took place in the Past, Present, or Future relative to the time the thumbnail was taken.

Two authors (AR and KB) coded the responses to the end-of-day debriefs for the first week of the study and iterated the coding rubric until they achieved a Cohen's Kappa of >0.60 for each category. They then separately coded the thumbnail responses.

Results

We fit a mixed linear model to each measure, using animation, cropping, and rehearsal as our predictor variables. Significant effects were detected by removing individual predictors from each model and using a one-way ANOVA to test for differences between the full and reduced models.

Size of Screenshot Thumbnails

Participants reliably (80% of the time) recognized their activity when full screenshots thumbnails were 320px tall ($\bar{x}$ =240px, σ=154px) and when cropped screenshot thumbnails were 460px tall ($\bar{x}$ =370px, σ =210px). For our participants, these heights correspond to thumbnails that were 30% and 45% the height of their screen, respectively. This difference was significant (χ^2=23.50, df=2, $p<0.001$). Rehearsing, however, did not have an effect on the required thumbnail size ($\bar{x}_r$=306px, $\bar{x}_u$=305px, χ^2=1.79, df=2, $p>0.05$).

Duration of Animation Thumbnails

Participants reliably (80% of the time) recognized their activity before the 9.5 second mark of full-screen animations ($\bar{x}$=8.0s, σ=4.1s) and the 14.5 second mark of cropped animations ($\bar{x}$=8.1s, σ=5.3s). This difference was not significant (χ^2=1.79, df=2, p>0.05) nor was the difference between rehearsed and unrehearsed thumbnails ($\bar{x}_r$=8.2s, $\bar{x}_u$=7.9s, χ^2=1.91, df=2, p>0.05).

Memory Strength, Thumbnail Appropriateness, and Accuracy

Participants tended to rate memories cued by animations as stronger than those cued by screenshots ($\bar{x}_a$=1.53, $\bar{x}_s$=1.26, max=2, χ^2=8.87, df=4, p=0.064) but gave similar ratings across the cropping and rehearsal conditions. Thumbnails were given similar ratings of appropriateness across all conditions. However, memories were more accurate when cued by screenshots than animations ($\bar{x}_s$=1.70, $\bar{x}_a$=1.27, max=2, χ^2=11.628, df=2, p<0.01).

Episodic Detail

Participants were remarkably consistent across conditions in the number and type of details they shared in their responses. While responses to animations were significantly shorter ($\bar{x}_a$=34.1 words, $\bar{x}_s$=61.0 words χ^2=16.68, df=4, p<0.01), they were equally likely to include each of the six types of contextual information we tracked. Figure 2 shows the frequency of each type of contextual information across all conditions.

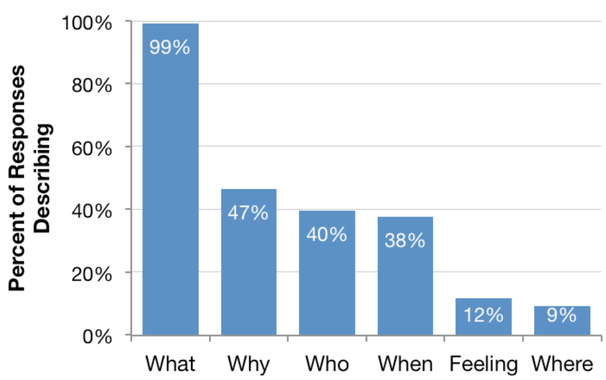

Figure 2: Participants recalled why, who, and when context in a large proportion of responses

Event Specificity

Participants were consistent across conditions in the abstractness of events they described. Of those cases where participants mentioned multiple levels of events, such as talking about an activity and then describing the individual actions that made up that activity, they were much more likely to start with high-level events and then describe lower-level ones (as happened 71% of the time) than to go from describing low-level events to high-level ones (28% of the time) (χ^2=21.6, df=1, p<0.001). Thus statements of the form "*Here I was trying to find car insurance again* [high], *trying to get a quote from a different company* [low]" were more common that those of the form "*I'm copying labels in the dictionary* [low] *so that I can make a count of how many certain types of signs there are in the sign language and handshake dictionary* [high]".

Time Discussed

Participants were also consistent across conditions in how often they discussed past events (29% of responses) or future events (24%). Nearly half (47%) of responses included a reference to either past or future events. Of those cases where multiple time periods were mentioned, participants were equally likely to move forward or backward in time (χ^2=0.096, df=1, p=0.757).

Discussion

Cropping

Cropped thumbnails had to be larger than full-screen thumbnails to cue recognition of past computer activity, but there was no difference in the accuracy or detail of the memories they cued. This evidence does not support H1.

Looking at how participants used thumbnails helps explain why. In some cases a small cue such as a unique photo brought back a set of memories. For example, upon seeing a 160px high cropped screenshot showing part of a gorilla, one participant recognized that they were "*watching a webpage to look up some information about Koko the gorilla and his vocalization abilities.*" But more often, participants pieced together what they were doing from multiple, distributed cues. For example, in the following response, the participant uses the fact that two windows were open simultaneously to remember a past meeting, why they had that meeting, and what they planned to do afterwards. "*Here I have the Github page for DIVY open and I also have the Terminal window open so this is when I was meeting with John… so I could try out some of my data clustering with this code to see if it works for me and then we could show it to the Smith lab.*" The cropped screenshot thumbnail that cued this response had to expand to be 640px high before the participant recognized their activity.

We also found that the mouse is not a good measure of attention. Often the mouse was over a blank or unremarkable portion of the screen so participants had to wait a long time before the cropped thumbnail included relevant information.

Animation

Participants tended to be more confident of memories prompted by animations but these memories were surprisingly less accurate. Animations prompted shorter responses but these included as much contextual information (who/what/where/when/why/feeling) as responses prompted by static screenshots. This evidence does not support H2.

One cause of this mismatch between perceived and actual accuracy was that in several cases participants simply described what was happening in an animation and ignored mental context. For example, compare the following end-of-day debrief and review text cued by an animation.

Debrief: "*So I released a batch of twenty participants through Mechanical Turk this morning at 9am and for some reason the traffic was very slow today. So until 2pm I think I had only gotten about ten people. But then I didn't want to* *wait until I get all twenty so I went ahead and analyzed the data just to see the pattern and saw that it wasn't in the direction that I wanted."*

Review: "*So I was making a pivot table and I was trying to see if I saw any pattern from my data, but I'm not sure if it was before I got all my data or after.*"

The debrief focuses on motivating events and talks about the activity abstractly, "analyzed the data", whereas the review text is much more specific, "making a pivot table", but fails to mention the larger context.

Rehearsal

We found no difference in the memories cued by rehearsed and unrehearsed thumbnails. Several factors could have led to this result. First, there may have been too much time between debriefing and reviewing. Whatever memory benefit the rehearsal provided may have dissipated by the time of review. Alternatively, our participants, all graduate students, may have been working on too few projects for rehearsing to have a targeted effect. Since graduate students usually work on a few large projects, letting them rehearse five memories from the day may have effectively let them reflect on most of their projects. This evidence does not support H3.

Content of Memories

A large proportion of responses described why an activity was taking place (47%), who was involved (40%) and when it occurred (38%). Some explanations of why an activity occurred were quite complex, describing multiple, conflicting motivations or a sequence of actions that depended on the current action. Descriptions of when an activity took place were often relative such as "before starting my day" or "during lab meeting." Absolute time (e.g., 3pm) was rarely mentioned, though participants did use the computer clock in thumbnails to distinguish similar activities or confirm guesses, such as "*Oh, yep. Tuesday morning. That would make sense.*"

To check if these contextual details were recalled from memory or could have been read directly from the thumbnail we conducted a post-hoc analysis. On average, responses included 4.47 (σ =1.89) episodic details and of these, 1.95 (σ =1.60) seemed to be reconstructed from memory. Specifically, the majority of feeling information (98%), why (81%), and where (78%) information seemed to be reconstructed from memory whereas when (47%), who (41%) and what (34%) information were more likely to have been represented directly in the thumbnail.

Participants mentioned activities in most of their responses (86%) such as "*going through and finding references from other references*" or "*trying to find pictures of my two advisors*". They also mentioned actions in most responses (56%), but these were often just descriptions of what was on the screen, "*I see that I was sending an email to someone*".

Participants also frequently mentioned events leading up to or after the thumbnail (47% of responses). Many of the past events were motivation, and while many of the future events included motivations, a number were outcomes "*this isn't the way I ended up solving* [this problem]."

Limitations

Given privacy concerns with recording computer activity, we only recruited 6 participants, who provided 215 responses. These participants were graduate students whose work often involves a small number of projects. Other working styles could have produced different results. We also did not balance presentation of screenshots and animations across the two weeks of our study, so the differences between screenshots and animations may be an artifact of having an extra week to practice giving responses.

Implications for Design

Thumbnails of full-screen images

Without a better predictor of users' attention than the mouse (e.g. gaze), full-screen thumbnails can be smaller than cropped thumbnails and covey as detailed of memories.

Our results also suggest that thumbnails still work for recognizing cross-application activities, but need to be larger (320px high for 80% recognition) than those used for recognizing website (208px high according to Kaasten et al. (2002)). This result is encouraging since a website thumbnail is a summary of one window, but a desktop thumbnail contain more content (overlapping windows, toolbars, etc.). This size increase may also be an artifact of computer monitors increasing in pixel density over the last decade.

Animations are less effective than expected

We found that animations, at least 5x time-lapses, cue less accurate memories than screenshots and found no evidence that they produce memories that are any more detailed. Showing activity over time may be better accomplished with small-multiples or thumbnails that allow scrubbing.

No need for rehearsing

We did not find evidence that random experience sampling and rehearsing improved memory in our tasks. This is in line with Sellen et al.'s (2007) finding that passively taken images were as good of triggers for remembering past events as actively taken images. This does not diminish the value of bookmarking past events to create landmarks for future reviewing, but one should not expect these events to be more memorable than others.

Conclusion

Existing computer systems do a poor job of helping users restore the mental context of interrupted activities. This restoration is particularly challenging for dynamic, creative, and non-linear activities that span multiple applications. Given humans' excellent visual memory, we investigated how visual thumbnails of computer desktops could help restore the memories that make up this context. Our work extends the scope of previous thumbnail research from representing previously visited websites to representing cross-application activities and their context.

Our two-week field study of six graduate students confirmed that thumbnails cue a significant amount of contextual information. We compared eight types of thumbnails: full screenshots, cropped screenshots, full animations, and cropped animations (each with or without rehearsal). Across conditions, a non-trivial proportion of thumbnail responses included information about what activity was taking place (99%), why it was happening (47%), who was involved (40%), and when it happened (38%). Between conditions, we found that 80% of activities could be recognized with full screenshot thumbnails that were 320px high, whereas thumbnails cropped around a region of interest (cropped screenshots) needed to be 460px high. Animations cued more targeted but less accurate memories than screenshots, which is surprising as we expected that seeing the activity unfold would lead to more accurate memories. Rehearsal had no discernable impact on the accuracy or detail of recalled context.

There is a rich design space to be explored in future research. For example, we plan to explore composing activity-specific thumbnails from snippets of multiple application events (e.g., a series of movements between applications related to accomplishing a specific activity) rather than focusing exclusively on thumbnails of a specific screen location at a specific moment in time. Animations may be more useful in showing these tightly coupled sequences. Another important next step will be to document how the contextual information cued by thumbnails is actually used to restart interrupted activities, as the use of this information will inform how it should be represented.

Our findings support the notion that visual cues can be used to recall the detailed context of complex activities. This finding is important because recovering context is a key first step towards restarting interrupted activities.

Acknowledgments

This research was funded by NSF grant #1319829

References

Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science*, 26(1), 39–83.

Bannon, L., Cypher, A., Greenspan, S., & Monty, M. L. (1983). Evaluation and analysis of users' activity organization. *Proceedings from CHI '83* (pp. 54–57). ACM.

Bardram, J., Bunde-Pedersen, J., & Soegaard, M. (2006). Support for activity-based computing in a personal computing operating system. *Proceedings from CHI '06* (pp. 211–220). ACM.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, 105(38), 14325-14329.

Brandt, J., Weiss, N., & Klemmer, S. R. (2007). txt 4 l8r: lowering the burden for diary studies under mobile conditions. *Proceedings from CHI '07* (pp. 2303–2308). ACM.

Cangiano, G. R., & Hollan, J. D. (2009). Capturing and restoring the context of everyday work: A case study at a law office. *In Human Centered Design* (pp. 945–954). Springer.

Czerwinski, M., Horvitz, E., & Wilhite, S. (2004). A diary study of task switching and interruptions. *Proceedings from CHI '04* (pp. 175–182). ACM.

Dumais, S., Cutrell, E., Cadiz, J. J., Jancke, G., Sarin, R., & Robbins, D. C. (2003). Stuff I've seen: a system for personal information retrieval and re-use. *Proceedings from SIGIR '03* (pp. 72–79). ACM.

Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research*, 50(4), 243–250.

Hupbach, A., Hardt, O., Gomez, R., & Nadel, L. (2008). The dynamics of memory: Context-dependent updating. *Learning & Memory*, 15(8), 574–579.

Iqbal, S. T., & Horvitz, E. (2007). Disruption and recovery of computing tasks: field study, analysis, and directions. *Proceedings from CHI '07* (pp. 677–686). ACM.

Kaasten, S., Greenberg, S., & Edwards, C. (2002). How people recognise previously seen Web pages from titles, URLs and thumbnails. *In People and Computers XVI-Memorable Yet Invisible* (pp. 247–265). Springer.

Kaptelinin, V. (2003). UMEA: translating interaction histories into project contexts. *Proceedings from CHI '03* (pp. 353–360). ACM.

Karger, D. R., Bakshi, K., Huynh, D., Quan, D., & Sinha, V. (2005). Haystack: A customizable general-purpose information management tool for end users of semistructured data. *In Proc. of the CIDR Conf.*

Mark, G., Gonzalez, V. M., & Harris, J. (2005). No task left behind?: examining the nature of fragmented work. *Proceedings from CHI '05* (pp. 321–330). ACM.

Rattenbury, T., & Canny, J. (2007). CAAD: an automatic task support system. *Proceedings from CHI '07* (pp. 687–696). ACM.

Sellen, A. J., Fogg, A., Aitken, M., Hodges, S., Rother, C., & Wood, K. (2007). Do life-logging technologies support memory for the past?: an experimental study using sensecam. *Proceedings from CHI '07* (pp. 81–90). ACM.

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science*, 19(2), 73-74.

Teevan, J., Cutrell, E., Fisher, D., Drucker, S. M., Ramos, G., André, P., & Hu, C. (2009). Visual snippets: summarizing web pages for search and revisitation. *Proceedings from CHI '09* (pp. 2023–2032). ACM.

Tversky, B., Morrison, J., & Betrancourt, M. (2002). Animation: can it facilitate? *International Journal of Human-Computer Studies*, 57, 247–262.

Van Solingen, R., Berghout, E., & van Latum, F. (1998). Interrupts: just a minute never is. *IEEE Software*, 15(5), 97–103.