If Not Now, When? The Effects of Interruption at Different Moments Within Task Execution

Piotr D. Adamczyk
School of Library and Information Science
University of Illinois
Urbana, IL 61801

Brian P. Bailey
Department of Computer Science
University of Illinois
Urbana, IL 61801

ABSTRACT

User attention is a scarce resource and users are susceptible to interruption overload. Systems do not reason about the costs of interrupting a user during a task sequence. In this study, we measure effects of interrupting a user at different moments within task execution in terms of task performance, emotional state, and social attribution. Task models were developed using event perception techniques, and the resulting models were used to identify interruption timings based on a user's predicted cognitive load. Our results show that different interruption moments have different impacts on user emotional state and positive social attribution, and suggest that a system could enable a user to maintain a high level of awareness while mitigating the disruptive effects of interruption. We discuss implications of these results for the design of an attention manager.

Author Keywords

Interruption, task models, attention, affective state.

ACM Classification Keywords

User Studies (primary keyword), Emotion and Affective User Interface, User and Cognitive models, Usability Testing and Evaluation

INTRODUCTION

Proactive applications such as email agents [16], and browsing assistants [15], are increasingly competing for user attention. Although these applications are well intentioned they do not reason about the impact an interruption has on a user, and even the most well-meaning application has the potential to cause interruption overload. Poorly timed interruptions can adversely affect task performance [3, 9, 10] and emotional state [4, 32]. Whether by changing their Instant Messaging status or blocking web site pop-ups, users try to cope with this problem by explicitly balancing their interruptibility with their need for information.

Strategies for notification have tried to achieve this balance by using novel visual strategies [11], multimodal communication [2, 6], or appropriate temporal moments [21]. The problem with common user solutions is that there is no guarantee that needed information will reach the user. The above methods try to find an alternate mode or particular pattern of interruption that can offload some of the user's cognitive stress. Previous studies have examined the effect of interruption timing on task performance [3, 9, 10]. However, our work differs in that we identify moments for interruption utilizing task models based on event perception research [29, 30, 31], and measure not only the effects of interruption on task performance, but also measure the effects of interruption on a user's emotional state and social attribution.

The interaction between a user and their system can be modeled as a social one [22]. Interpretations of the behavior of an interrupting application can turn users away from future application use [14, 24], and influence attitude towards the information that application provides [28]. In this work, we conducted an evaluation to measure the effects of interruptions at particular moments during task execution. Starting from work in event perception [30], we developed user task models that reflect the user's own cognitive representation of their tasks. We used the models to predict better and worse moments for interruption. In our evaluation, participants performed editing, searching, and media tasks while periodically being interrupted by a news alert service.

Our results show that our predicted best points for interruption consistently produced less annoyance, frustration, and time pressure, required less mental effort, and were deemed by the user, more respectful of the their primary task. In terms of Annoyance, the best moment for interruption showed a 56% reduction compared to the worst condition, and was 43% lower than the random condition. The best moment for interruption conveyed 43% more respect than in the worst condition, and 27% more than the random condition.

RELATED WORK

Visual and Multimodal Strategies for Interruption

McFarlane [18] presents 4 strategies for coordinating interruption in HCI. These are summarized as: immediate, negotiated, mediated, and scheduled. Others [17, 25] have dealt with visual strategies for presenting peripheral information and balancing interruption and awareness. Multimodal communication strategies have also been explored [2, 6]. Each of these approaches depends on some model of sensory attention when deciding how and when to interrupt. The problem remains difficult, as this information may not always be readily available to an application designer.

Our work can help inform the design of these systems. The task models we develop aim to represent the internal hierarchical representation of ongoing behavior directly, providing more accurate timings for interruption in the process. These can help systems decide when to interrupt and add precision to their temporal components.

Temporal Strategies for Interruption

Some studies [8, 9] place moments for interruption towards the beginning, middle, or end of a task. This kind of strategy relates most to Miyata and Norman [20]. The authors explain that task execution occurs in three phases: planning, execution and evaluation. If this applies to a task, a logical extension is that it would also apply to each of the subtasks of a task. As tasks in themselves, every subtask would then contain moments of planning, execution, and evaluation, making task execution a repeated loop of these phases. Another possibility is that the three phases occur in parallel, governed by a single central executive control. There are clearly effects to interrupting during the various phases [32] but associating rough temporal placement (beginning, middle, end) might be an oversimplification of task execution.

Other studies place interruptions between instances of repetitive sequences or, more generally, at breakpoints in a task sequence [5, 19, 21]. The choice of these points is more intuitive but the reasoning behind these locations remains ill defined. Studies of this type sometimes produce internally inconsistent results [19]. Our study differs from other temporal strategies by relying on event perception models to determine moments for interruption.

Event Perception

Studies in psychology have provided insights into the ways in which tasks are decomposed hierarchically in the mind [29, 30, 31]. Observers of events segment ongoing activity into temporal parts and sub-parts that are reliable, meaningful and correlated with ecologically relevant features of the action [29]. In [29] this process of recognizing time-based boundaries was linked to distinct patterns of brain activity. In an experimental follow-up [30], subjects were shown video recordings of tasks being performed, and then asked to communicate or recall the task structure. It was shown that subjects remembered events as hierarchies with two levels, coarse and fine. Coarse breakpoints largely represented the introduction of objects and broad actions on those objects, while Fine breakpoints were the more precise actions in the scene. The study showed how event structure influences recall of tasks and goals, and that moments that are best recalled are those that are more firmly related to schematic action – recognizable and well understood activities.

TASK MODELING

Based on event perception research into such a deep structure involved in the composition of cognitive task hierarchies, a prediction was made regarding moments for interruption.

The best moments for interruption should be between two coarse breakpoints that are, on the whole, better understood and better recalled [7]. Having just completed a schematic event, the subject is utilizing fewer cognitive resources, leaving the rest immediately available for a peripheral task [7, 33]. After interruption, the next schematic event could be quickly recalled from memory and execution could resume with potentially little disruption [1, 34].

The placements of some of the moments for interruption (see Table 2) are similar to those that appear in other studies [3, 9, 21], but they are distinct in a few fundamental ways. Interruption triggers are based on behavior that there is good reason to believe this is significant in the mind of the user, and the interruptions are not associated with a temporal phase, making it easier for them to be applied anywhere during execution.

Our methods for eliciting task models paralleled [30]. Full-color video captures of instances of three tasks were recorded using HyperCam, performed at a screen resolution of 1024x768 pixels. A sample task instance, checking email on a UNIX server, was also recorded. The task video captures were shown on a large projected display to a group of participants. Twenty-five subjects, 16 male and 9 female, participated in the task model elicitation phase. Subjects were instructed as to the difference between fine and coarse breakpoints. The email task was shown first as an example, followed by a breakpoint listing for the task. (see Table 1) Experimental task videos were shown twice. On the first viewing, subjects were instructed to note the coarse breakpoints, and fine breakpoints during the second.

We recorded how often and where a particular breakpoint appeared in subject task models. At its highest levels, agreement among users reached 80%, a value in line with similar research [reviewed in 31]. The tasks were defined within a linguistic tolerance. For example, responses like “Opens Word” and “Launches Editor” were deemed to refer to the same task in an execution sequence.

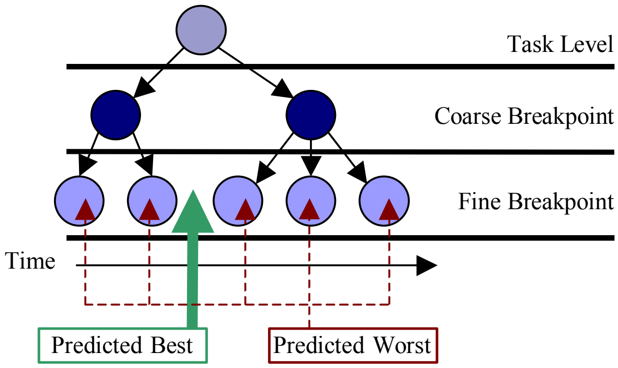

Figure 1. Task model hierarchy

The resulting weights were then used to determine points for interruption in accordance with our hypothesis. Our hypothesis is that points in a task sequence where there is a large degree of agreement across respondent's task models represent the most concretely and commonly understood moments in a task's execution. Our hypothesis is an extension from the fact that it is these moments that people tend to recall most frequently and in the appropriate order [30]. When our hypothesis is applied to the Interruption domain, it allows us to forecast best and worst moments for interruption. Predicted best points for interruption correspond to moments between breakpoints with a large degree of agreement. Predicted worst points correspond to moments between breakpoints with a small degree of agreement. In our task models, presumed best mapped to moments between coarse breakpoints, and presumed worst to moments in the execution of fine breakpoints.

The task models we developed were a majority view of the three task classes, where fine breakpoints can be understood as subtasks of a larger coarse breakpoint. The predicted best points for interruption are those when a user is moving from one well-defined and commonly understood task to another. The presumed worst points are those where a user is involved in a highly subjective sequence of ill-defined and user specific behavior. (see Figure 1)

USER STUDY

We set out to measure the effects of interruption on environmental computing tasks. In addition to traditional performance measures, we were particularly interested in the collection and analysis of information about user emotional state under various interruption timings. We also hoped to gain insight into the role of social attribution in interruption.

Experimental Design

The study was a 3 (task type: Editing, Media, Searching) x 4 (interruption trigger: presumed best, presumed worst, random, no) repeated measures within subjects design. Tasks were grouped by type into 3 sets of 4. Task types were counterbalanced and interruption timings were arranged in a balanced Latin Square to remove any ordering effects.

Coarse-unit Descriptions	Fine-unit Descriptions
Moves to Start Menu	Moves to Start Menu Moves to Apps folder
Selects email application	Selects email application Types in username Types in password
Opens email	Selects email function Goes to message index Moves through messages
Opens message	Opens particular email Scrolls through message Selects text
Copies text from message	Copies text Exits email function
Closes application	Logs off Closes application

Table 1. Coarse and fine unit descriptions for sample task.

Subjects

Sixteen subjects, 13 male and 3 female, were enlisted in the user study. The subject pool consisted of undergraduate and graduate students from various departments at our institution. Subjects were not compensated for their time for this study. Though we did not balance for gender, previous HCI interruption research has shown no gender effect. [35]

Tasks

We devised three experimental task classes for this study. The first was a document-editing task performed in Microsoft Word XP (see Figure 2). Four short (100-115 words) film reviews were combined into a single Word document. Spelling errors were introduced into the text. The document was further modified (words replaced, punctuation removed) and comments inserted with instructions to restore the document to its original state. Four instances of the editing task were created, each with different film reviews, but the same number of spelling mistakes and comments. Subjects were instructed to make the appropriate changes to the document as quickly and accurately as possible, and save the modified document to the Desktop.

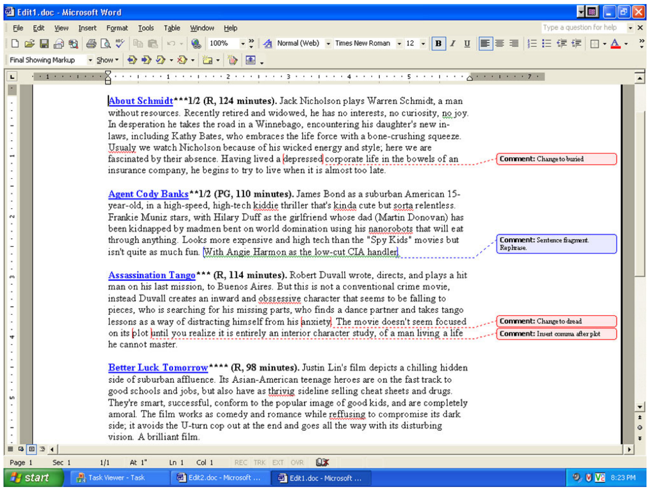

Figure 2. Editing Task

Figure 3. Media Task

The second task class consisted of four short documentary or news media clips (see Figure 3). The narrated film clips were about 1 minute 45 seconds in length and were stored on the computer used by the subject. The subjects were instructed to watch the video clip, write a short summary (one paragraph) in MS Word, and save the summary to the Desktop.

The final task class was a web-searching task. Subjects were provided with clues and instructed to identify a target publication by a member of the faculty at our University. Using Internet Explorer 6, starting from the page listing department faculty, subjects were to find the professor's homepage and their publications listing. Subjects were to locate the citation for the publication in question, copy the text to a MS Word file, and save the file to the Desktop. To counteract any strategic learning effects, clues varied between task instances. Though the task remained the same for all instances, individual professor's homepages were constructed differently enough that learning effects based on page structure were also minimized.

These tasks were developed for a number of reasons. We were interested in seeing whether our approach would be applicable and consistent across a number of different task types. We included tasks that would produce varying cognitive loads on different mental resources. Each task differed in complexity, was varied enough from the others so as to be distinct, and all tasks were still environmental enough to meaningfully reflect the real activity of users.

Interruptions

Interrupting tasks were modeled after those in a similar study [3, 4]. Subjects were confronted with a full screen pop-up containing the first paragraph from a newswire service and, from three radio button options, instructed to select an appropriate title for the entire news story. Subjects were instructed to answer the interrupting question immediately and accurately, and without moving the interrupting window. A full screen pop-up was used so as to assure that interruption would not go unnoticed. This mode allowed for the most immediate and complete notification, and made certain that the time of interruption reflected visual onset.

Hardware and Software

In the user study, we used a client-server application. The client, TaskViewer, was installed on the subject computer and received commands from the TaskServer application installed on the experimenter's computer. All of the required applications were made available through the TaskViewer interface. This removed the need for subjects to navigate the Start Menu, which may have produced a compound task. Files required for the execution of the experiment were stored locally on the subject computer and loaded automatically by the various buttons in the TaskViewer interface. Experiments were performed using two high-end Windows XP PCs. The monitors were 19" CRT monitors with a screen resolution of 1024x768 pixels.

Procedure

Upon entering the study environment we went through an informed consent process with the subject. If consent was given, subjects were provided with instruction, both written and oral, on the tasks they were expected to perform. They were shown a modified NASA-TLX survey (discussed in the next section) and told to complete one after each trial. Subjects were informed that there would be an application running on their computer that would periodically appear with a new task. The interrupting task was described as a paragraph of text with an accompanying multiple-choice question, and was to be answered as quickly and accurately as possible. Finally, subjects were told to take the whole task execution, including any interruptions they may have encountered, into account when filling out the TLX form. Each subject was run individually for a one hour session.

At the beginning of every task class, the experimenter first sent a sample instance of the task type, guided the subject through the use of the TaskViewer interface, and was available to answer any questions that might arise. The experimenter remained out of view during the remainder of

Task	Interruption	Interruption Trigger
Edit	Presumed Best	Upon completing an edit – spelling mistake or comment
	Presumed Worst	During an edit – typing, selecting, or confirming
Media	Presumed Best	After completing summary, but before the save process
	Presumed Worst	During the viewing of the Video clip
Search	Presumed Best	After copying the citation, but before the save process
	Presumed Worst	During the search process, on the publications page

Table 2. Description of Interruption Triggers by Task

the trial task executions. To be able to monitor a subject's on screen activity, a RealVNC server was installed on the subject's computer. The experimenter, through a corresponding client application, watched the subject's task execution in real time.

Before sending over a trial task, an interruption was selected from a pool of 12, and loaded into the TaskServer interface. Our task models provided clear descriptions of the moments in a task's execution that would provoke an interruption. During the presumed best and presumed worst moment, the experimenter noted such behavior, and transmitted the interruption to the client computer. To handle the random condition we first averaged task execution times gathered from a small pilot study. After a trial task was sent to the subject, the experimenter also triggered an interrupt, timed to occur at a random point within the interval determined in the pilot study. This method assured that the interruption would occur before completion of the task.

Measurements

We collected a total of 10 measures. We logged Time on Interruption (TOI) and Time on Task (TOT) from a file generated by TaskViewer. TOT was the total trial time minus the TOI. An approximate value for Resumption Lag, the time a subject takes to switch focus back to primary task after interruption, was also collected.

The remaining 7 measures were the subjective scales in a modified NASA-TLX survey presented to subjects after each trial. The NASA-TLX [13] subjective workload assessment tool has been used in a number of HCI studies [6, 23]. It was chosen over other workload and affect scales, like PANAS, because of its continuous scale and short length. As the survey was administered 12 times, the scale's short length and simple marking strategy made it less likely to confound any of the primary tasks. The continuous scale also allowed for more fine-grained

TLX Value	Effect	F	p
Annoyance	Task	F(2,13) = 0.220	0.806
	Interruption	F(3,12)=10.532	0.001*
Frustration	Task	F(2,13) = 0.449	0.648
	Interruption	F(3,12) = 9.795	0.002*
Time Pressure	Task	F(2,13) =1.695	0.222
	Interruption	F(3,12) = 5.564	0.013*
Own Performance	Task	F(2,13) = 3.550	0.059
	Interruption	F(3,12) = 2.054	0.160
Mental Effort	Task	F(2,13)= 14.614	<0.001*
	Interruption	F(3,12) = 4.436	0.026*
Mental Demand	Task	F(2,13) = 6.411	0.012*
	Interruption	F(3,12) = 2.190	0.142
Respect	Task	F(2,13) = 1.845	0.197
	Interruption	F(3,12) = 4.964	0.018*

Table 3. Main effects for TLX dimensions. Starred results are significant for $\alpha = 0.05$.

responses from users. The survey was administered on paper rather than on the computer to provide a clear distinction between task conditions and to remove any bias from additional computer tasks.

The TLX was modified in two places. The Anxiety scale was replaced by Annoyance, and Physical Demand was replaced by the question; "How respectful was the application of your task?" We chose respect as the dimension of social attribution as it conveyed the sought after component of deference to the primary task. The remaining scales were: Mental Demand, Mental Effort, Frustration, Time Pressure and Own Performance. Descriptions of the scales were made available to the subjects.

RESULTS

A two-way within-subjects analysis of variance was conducted to evaluate the effect of interruption strategies and task type on the various TLX measures. The dependent variables were continuous TLX ratings of 0 to 5. Results are summarized in Table 3. There were no significant interaction effects. The main effects were tested using the multivariate criterion of Wilks' lambda. Post hoc analyses were conducted on the significant main effects. These consisted of paired-samples t tests (6 for interruption main effects, and 3 for task main effects) with familywise error rate controlled across the test using Holm's sequential Bonferroni approach.

TOT, TOI, Resumption Lag

Task type had a main effect on TOT (F (2,11)=4.685, p=0.034). Post hoc analysis showed a significant difference between the means of the Media and Search tasks (t (15)=3.231, p=0.006). There were no significant effects on TOI. We calculated approximate resumption lag values as the time between closing the interrupting task and bringing the primary task window into focus. There were no significant effects on Resumption Lag.

Annoyance, Frustration, Time Pressure

Interruption timing had a significant main effect on a subject's reported frustration and annoyance. Annoyance rose 56% from the best to worst conditions, and 43% from best to random. Frustration rose 49% from the best to worst condition and 20% from best to random. And although absolute values in Time Pressure remained low, the relative increases exhibited the same pattern of growth, 55% higher at worst and 37% higher at random.

For Annoyance there were significant differences between the means for the following pairs of conditions: Random and No (t (15)=4.857, p<0.001), Worst and No (t (15)=-5.482, p<0.001), and Best and Worst (t (15)=-3.732, p = 0.002). For Frustration there were significant differences between the means for the following pairs of conditions: Random and No (t (15)=4.739, p<0.001), Worst and No (t (15)=-5.230, p<0.001), and Best and Worst (t (15)=-3.850, p = 0.002). For Time Pressure there were significant differences between the means for the following pairs of conditions: Best and Worst (t (15)=-4.707, p<0.001), Worst and No (t (15)=-3.489, p<0.001), and Best and Random (t (15)=2.990, p = 0.009).

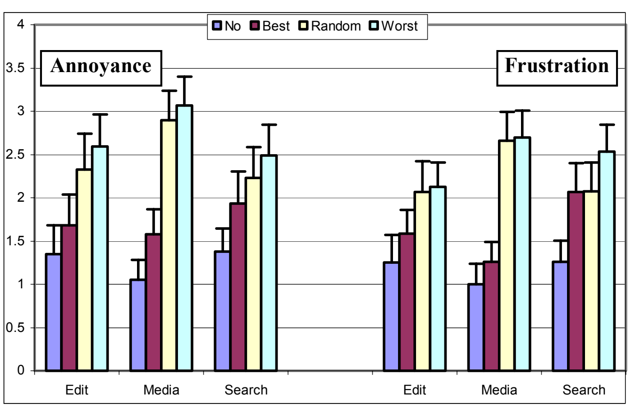

Figure 4. Annoyance and Frustration (0 = low, 5 = high)

Mental Demand, Mental Effort

Task type had a significant main effect on reported Mental Demand and Mental Effort. There was little difference in Mental Effort between Edit and Search tasks, but the Media condition showed a 37% increase over the other tasks. In Mental Demand, Search was reported at rates 14% higher than Edit, and Media at 23% higher.

For Mental Demand, t tests showed no significant differences between the means of task pairs. For Mental Effort there were significant differences between the means for the following pairs of tasks: Edit and Media (t (15)=-4.114, p=0.001), and Media and Search (t (15)=2.999, p=0.009). Also for Mental Effort, there was also significant a significant difference between the means for Random and No (t (15)=3.417, p=0.004).

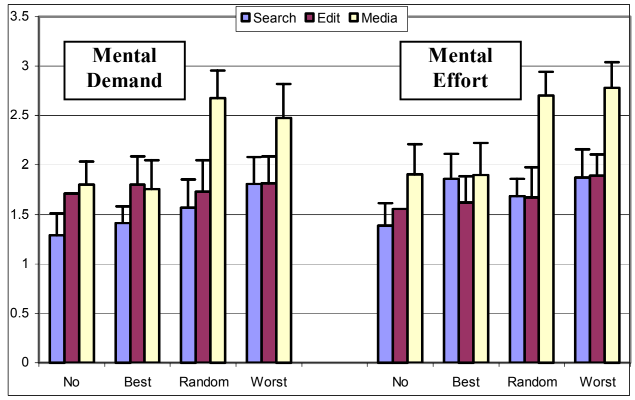

Figure 5. Mental Demand and Mental Effort (0 = low,5 = high)

Respect

The values for respect showed a significant interaction with interruption timing. The best condition had levels of respect 43% higher than the worst condition and 27% higher than at random.

For Respect there were significant differences between the means for the following pairs of conditions: Worst and No (t (15)=3.728, p=0.002), Worst and Best (t (15)=3.474, p=0.003), and Random and No (t (15)=-3.329, p = 0.005).

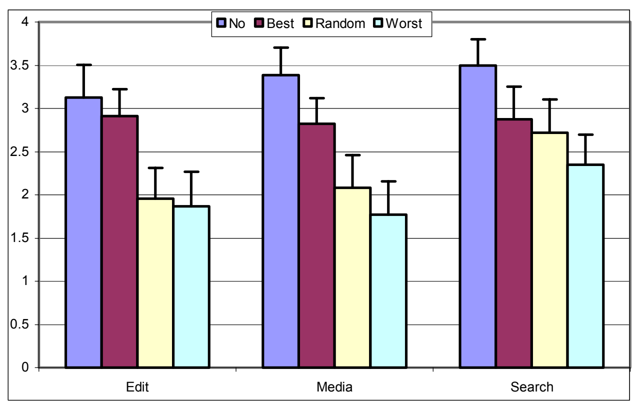

Figure 6. Respect (0 = low, 5 = high)

DISCUSSION

Our results show significant impacts along a variety of scales. Our predicted best points for interruption consistently produced less annoyance, frustration, and time pressure, required less mental effort, and were deemed by the user, more respectful of the their primary task. The predicted Worst condition also underperformed the Random condition on the same measures.

Our results concerning TOT, TOI, and Resumption Lag are slightly inconsistent with previous research. TOT did differ across task types, suggesting that the tasks differed from one another. The lack of main effect of interruption strategy may be due to any impacts being dominated by the comparatively long trial times. The lack of significant interactions for TOI is consistent with past work [4]. The values for Resumption Lag were approximations, as our log recorded events to the second. Differences, if they exist, may have been smaller than the granularity at which the measure was recorded.

Our measures of Annoyance and Frustration indicate the change in effects across the interruption strategies. Our predicted Best moment was effective at minimizing the disruptive effects of an interruption, with values significantly closer to the No interruption condition than either Random or Worst.

The presence of task type main effects and absence of interruption main effects, along the Mental Demand and Mental Effort scales suggests that the relationship between interruption and task is crucial in certain conditions. During the media task for example, the effects of the interruption in the Random and Worst conditions were particularly disruptive. In those conditions the interruption was most likely to occur during, and thereby, completely obscure the video clip. The reading comprehension nature of the interruption made listening to the video narration difficult. In fact many subjects were observed turning down the computer speaker volume until after completing the interruption, only then resuming the primary task.

It may be appropriate for designers to consider alternate modalities for interruption, taking into account the particular relationships between different cognitive resources, and choosing those that would provoke the least additional load [27].

The main effect of interruption strategy on Respect helps to quantify the role of social attribution in interruption. Worst interruption timings were interpreted as signs of disrespect. If a user does not attribute respect to an application, it may cause them to discontinue its use [24]. Equally important is the negative impact disrespectful interruptions have on information use [28].

IMPLICATIONS FOR DESIGN

Our results have implications for the design of an attention manager system. An attention manager attempts to identify opportune moments in a user's task sequence for an interruption to occur [4]. Our results show that such a system could significantly decrease the disruptive effects of interruption on users' emotional state and social attribution.

This study links human interruptibility with a deeper cognitive representation of a user's task. Our task models direct designers to points in a task where a user is more amenable to an interruption, and provide guidelines to help identify better and worse points in a task sequence. To develop an effective attention manager, one must either supply the attention manager with the task models or the attention manager must learn the task models over time. We believe that both directions must be investigated in the future.

Multimodal interruption schemes [2,6] and novel visual presentation styles [17, 25, 35] help reduce the negative effects of interruption. When combined with the kinds of task models outlined in this paper, an attention manager could be equipped with an adaptive rule set for an even more effective interruption policy. Under normal conditions interruptions would be held until the user was available, and the choice of notification style would be only as disruptive as necessary. However, in the case of a particularly important or time sensitive notification, the attention manager could identify an appropriate moment within a time boundary defined by the notification, and select the minimal cue necessary to gain user attention.

Interruption lag [26] can cue a rehearsal process in the user before an interruption (e.g. a phone ringing) to help in resuming the primary task. In an experimental study with a complex task [19], training users in a rehearsal strategy actually decreased task performance. With an informed attention manager system, the user wouldn't be required to rehearse, as the points at which the interruptions would occur would be easy to recall [1, 7, 12, 34].

REFERENCES

1. Altmann E. M. and Trafton J. G. (2002) Memory for goals: An activation-based model, *Cognitive Science*, 26, 39–83.
2. Arroyo E., Selker T. and Stouffs A. (2002) Interruptions as multimodal outputs: Which are the less disruptive? In: *(ICMI'02)*, p. 479-483.
3. Bailey B. P., Konstan J. A. and Carlis J. V. (2000) Measuring the effects of interruptions on task performance in the user interface. In: *(SMC 2000)*, IEEE, pp. 757-762.
4. Bailey B. P., Konstan J. A. and Carlis J. V. (2001) The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In: *Human-Computer Interaction – INTERACT 2001* pp. 593-601.
5. Bannon L., Cypher A., Greenspan S. and Monty M. L. (1983) Evaluation and analysis of users' activity organization. *CHI'83* pp. 54-57.
6. Brewster, S.A., Wright, P.C., & Edwards, A. D. N. The Design and Evaluation of an Auditory Enhanced Scrollbar. CHI 94. pp.173-179.
7. Cohen, S. (1980) Aftereffects of Stress on Human Performance and Social Behavior: A Review of Research and Theory. *Psychological Bulletin* 88 (1), pp. 82-108
8. Cutrell E., Czerwinski M. and Horvitz E. (2001) Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. *INTERACT 2001* pp.263-269.
9. Czerwinski M., Cutrell E. and Horvitz E. (2000) Instant messaging: Effects of relevance and time. In: S. Turner and P. Turner (Eds.), *People and Computers XIV:*

Proceedings of HCI 2000, Vol. 2, British Computer Society, pp. 71-76.

10. Czerwinski M., Cutrell E. and Horvitz E. (2000) Instant messaging and interruption: Influence of task type on performance. In: C. Paris, N. Ozkan, S. Howard and S. Lu (Eds.), *OZCHI 2000 Conference Proceedings*, pp. 356-361.

11. Dantzich, M., D. Robbins, E. Horvitz, and M. Czerwinski. Scope: Providing Awareness of Multiple Notifications at a Glance. *Proceedings of Advanced Visual Interfaces*, 2002.

12. Dix A., Ramduny-Ellis D. and Wilkinson J. (2003) Trigger analysis - understanding broken tasks. In: D. Diaper and N. Stanton (Eds.), *The Handbook of Task Analysis for Human-Computer Interaction*. Lawrence Erlbaum Associates, 2003.

13. Hart, S.G., & Staveland, L.E. (1988). Development of a NASA-TLX (Task load index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), Human Mental Workload pp. 139-183. Amsterdam: North-Holland.

14. Hudson J. M., Christensen J., Kellogg W. A. and Erickson T. (2002) "I'd be overwhelmed, but it's just one more thing to do": Availability and interruption in research management. *CHI'02*, pp. 97-104.

15. Lieberman, H. (1995), Letizia: An Agent that Assists Web Browsing, *IJCAI 95*, AAAI Press.

16. Maes, P., (1994) Agents that reduce work and information overload *Comms. of the ACM* 37 (7) 31-40.

17. Maglio, P. and C.S. Campbell. Tradeoffs in Displaying Peripheral Information. *CHI*, 2000, pp. 241-248.

18. McFarlane D. C. (2002) Comparison of four primary methods for coordinating the interruption of people in human-computer interaction, *Human-Computer Interaction*, 17 (1), 63-139.

19. Miller S. L. (2002) Window of opportunity: Using the interruption lag to manage disruption in complex tasks. To appear in: *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*. Santa Monica, CA: Human Factors and Ergonomics Society.

20. Miyata Y. and Norman D. (1986) Psychological issues in support of multiple activities. In: D. A. Norman and S. W. Draper (Eds.), *User Centered Systems Design: New Perspectives on Human-Computer Interaction*. Hillsdale: Lawrence Erlbaum Associates, pp. 265-284.

21. Monk C. A., Boehm-Davis D. A. and Trafton J. G. (2002) The attentional costs of interrupting task performance at various stages. *(HFES 2002)*.

22. Nass, C., J. Steuer and E. Tauber. Computers are Social Actors. *CHI 94*, pp. 72-78.

23. Rowe, D.R., Sibert, J., & Irwin, D. Heart Rate Variability: Indicator of User State as an Aid to Human-Computer Interaction. *CHI 98*. pp. 480-487

24. Shneiderman, B. Designing the User Interface. Reading, MA. Addison-Wesley-Longman (1998)

25. Somervell, J., C.M. Chewar, and S. McCrickard. Evaluating Graphical vs Textual Secondary Displays for Information Notification. *Proceedings of the ACM Southeast Conference*, Raleigh NC, April 2002.

26. Trafton J. G., Altmann E. M., Brock D. P. and Mintz F. E. (2003) Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal, *International Journal of Human-Computer Studies*, 58, 583-603.

27. Wickens, C. D. & Hollands, J. G. (2000). Engineering psychology and human performance. Upper Saddle River, NJ: Prentice Hall.

28. Xia, L., and Sudharshan, D., Effects of Interruptions on Consumer Online Decision Processes. *Journal of Consumer Psychology*, 12(3), 265-280, Lawrence Erlbaum Associates (2002).

29. Zacks, J., Braver, T.S., Sheridan, M.A., Donaldson, D.I., Snyder, A.Z., Ollinger, J.M., Buckner, R.L., & Raichle, M.E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4(6),* 651-655.

30. Zacks, J., & Tversky, B. (2001). Event structure in perception and cognition. *Psychological Bulletin, 127(1),* 3-21.

31. Zacks, J., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events, *Journal of Experimental Psychology: General, 130(1),* 29-58.

32. Zijlstra F. R. H., Roe R. A., Leonova A. B. and Krediet I. (1999) Temporal factors in mental work: Effects of interrupted activities, *Journal of Occupational and Organizational Psychology*, 72, 163-185.

33. Cellier J.-M. and Eyrolle H. (1992) Interference between switched tasks, *Ergonomics*, 35 (1), 25-36.

34. Sanders G. S. and Baron R. S. (1975) The motivating effects of distraction on task performance, *Journal of Personality and Social Psychology*, 32, 956-963.

35. Bailey, B.P., J.A. Konstan and J.V. Carlis. Adjusting Windows - Balancing Information Awareness with Intrusion. P*roceedings of the 6th Conference on Human Factors and the Web*, 2000.