Not Now! Supporting Interruption Management by Indicating the Modality and Urgency of Pending Tasks

Chih-Yuan Ho, Mark I. Nikolic, Molly J. Waters, and **Nadine B. Sarter,** Ohio State University, Columbus, Ohio

Operators in complex event-driven domains must coordinate competing attentional demands in the form of multiple tasks and interactions. This study examined the extent to which this requirement can be supported more effectively through informative interruption cueing (in this case, partial information about the nature of pending tasks). The 48 participants performed a visually demanding air traffic control (ATC) task. They were randomly assigned to 1 of 3 experimental groups that differed in the availability of information (not available, available upon request, available automatically) about the urgency and modality of pending interruption tasks. Within-subject variables included ATC-related workload and the modality, frequency, and priority of interruption tasks. The results show that advance knowledge about the nature of pending tasks led participants to delay visual interruption tasks the longest, which allowed them to avoid intramodal interference and scanning costs associated with performing these tasks concurrently with ATC tasks. The 3 experimental groups did not differ significantly in terms of their interruption task performance; however, the group that automatically received task-related information showed better ATC performance, thus experiencing a net performance gain. Actual or potential applications of this research include the design of interfaces in support of attention and interruption management in a wide range of event-driven environments.

INTRODUCTION

Operators in complex event-driven domains face considerable and often competing attentional demands. They need to perform a wide range of tasks in coordination with other human and machine agents. Some tasks need to be performed immediately and concurrently with other activities; others can be delayed and executed in sequence. Similarly, interactions with other agents may require an immediate response or allow for delayed attending. The challenge for operators is to schedule their various activities and manage attentional resources as effectively as possible. They need to complete tasks and interactions in a timely manner and, at the same time, avoid unnecessary interruptions of ongoing activities and lines of reasoning. Such interruptions have the potential to lead to performance

breakdowns in the form of errors and reduced efficiency on both the interrupting and the interrupted tasks (e.g., Dismukes, Young, & Sumwalt, 1998; Gillie & Broadbent, 1989; Latorella, 1998).

Most currently used interruption cues do not support operators in managing attentional resources. They often lead to a data-driven reorientation of the operator's attention away from its prior focus to the interruption signal before its importance can be determined. Also, interruption cues tend to be uninformative and thus fail to support operators in deciding whether and when to attend to an incoming signal or new task. This can lead to the unintentional dismissal or the preemptive integration of an interruption (McFarlane & Latorella, 2002). One promising way to improve the situation is to support preattentive reference (i.e., the preattentive

evaluation of interruption signals; Woods, 1995). There are two prerequisites for preattentive reference: (a) Operators must be provided with (at least) partial information about the nature and cognitive requirements of the pending task or tasks to enable them to decide quickly whether an immediate shift in attention to a new task or interaction is warranted or should be delayed – the focus of the present research. (b) The information must be presented in such a way that it allows for peripheral access (i.e., such that it can be picked up in parallel with ongoing lines of reasoning; Sarter, 2002; Woods, 1995). Few efforts have been made to support preattentive reference through informative interruption cuing.

One notable example is the work on likelihood alarm displays by Sorkin, Kantowitz, and Kantowitz (1988). In an attempt to address problems with binary alarms, which are rather uninformative but prevail in most work environments, those authors developed an alarm system in which information about event likelihood is encoded into the alerting signal. They have shown that by providing this additional information, likelihood alarm displays enable operators to determine the relative urgency and the likely benefit of attending to an alarm, thus improving their allocation of attention among multiple tasks and events.

Another line of research on attention and task management was conducted more recently by Funk and Braune (1999). They developed computational aids, such as the Cockpit Task Management System (CTMS) and the Agenda Manager, to help pilots initiate, monitor, prioritize, and terminate tasks and goals. The CTMS, for example, provides the operator with information about task state (upcoming, active, terminated), status (satisfactory vs. unsatisfactory performance), and priority. This system has been tested and shown to reduce task misprioritization and the number of incomplete tasks in the context of simulated flight deck operations.

One last example is Latorella's (1998, 1999) work on interruption management on commercial flight decks – a special case of task and attention management. Latorella's studies investigated the influence of the modality (visual vs. auditory) of both the task that is being interrupted and the one that is interrupting an ongoing task on pilot performance and strategies. Her

results confirmed the expected performance benefits of cross-modality conditions (e.g., a visual task should be easier to perform concurrently with an auditory task than with another visual task), the more compelling nature of auditory interruptions, and the high level of resistance to interruptions of ongoing auditory tasks. However, they also suggested some more complex interaction effects between combinations of tasks in the visual and auditory modality.

As part of this work, Latorella (1999) developed the interruption management stage model, which specifies five interruption management behaviors: (a) oblivious dismissal (interruption not detected); (b) unintentional dismissal (significance of interruption signal not interpreted correctly); (c) preemptive integration (interruption task immediately started and completed, thus intruding on ongoing task); (d) intentional dismissal (interruption correctly interpreted, deliberate decision not to perform interruption task); and (e) intentional integration (interruption and ongoing task considered as set and integrated by operator). The first three behaviors are symptoms of a breakdown in task and interruption management, whereas intentional dismissals and intentional integration are desirable responses that should be supported through information and interface design.

The present experiment built on the latter two studies. It examined the effects of providing partial information about pending tasks on operators' scheduling strategies and overall task performance. It went beyond the earlier research by (a) including information on the time remaining to perform a task, (b) varying the availability of the information (not available, available upon request, and automatically available), and (c) including a tactile task to examine cross-modal interference between the sense of touch and vision and audition, respectively.

Information on the modality of a pending task was provided based on the prediction of multiple resource theory (Wickens, 1984) that the performance of concurrent tasks suffers to the extent that these tasks share the same resources (including processing code, stage, and, most important for the present research, task modality). Thus knowledge of the modality in which a task will be presented should be considered valuable, as it supports intentional dismissal and

intentional integration by helping an operator to anticipate and thus avoid resource competition and intramodal interference. Information about the time remaining to perform a pending task was made available to aid in fine-tuning task scheduling, as suggested by Latorella (1998). In cases where concurrent performance of two tasks in the same modality could not be avoided due to the urgency of the pending task, performance costs were anticipated for the interruption task (in the form of incorrect counts) and/or for the air traffic control (ATC) task (in the form of traffic conflicts, missed handoffs, or incorrect altitude assignments).

METHOD

Participants

The participants were 48 graduate and undergraduate students from the Ohio State University (42 men and 6 women). The students were recruited from several colleges across campus, with the largest group of participants being enrolled in the College of Engineering (32 participants). None of the participants had any experience with performing ATC tasks. The average age for all students was 23.5 years (SD = 4.1). Participation was voluntary, and all participants were paid $10/hr for their collaboration. They were randomly assigned to one of three experimental groups, which differed with respect to the availability of information concerning the nature of interruption tasks.

Experimental Design

This study employed a 3 × (3 × 2 × 2 × 2) mixed-model design in which the availability of visually presented information – none, request (available upon request), and auto (available automatically) – concerning interruption tasks was the between-subjects variable. Participants in the first group (none) did not receive any task-related information. Those in the request group could ask for information about the modality of low-priority tasks and the time remaining to perform the task. The auto group automatically received this information. The within-subjects variables in this study were the modality (visual, auditory, tactile), the priority (urgent vs. low priority) and frequency (high vs. low) of the

interruption task, and the ATC workload level (high vs. low).

Tasks

Primary task. Participants were asked to perform a simulated ATC task involving digital air-ground communication (so-called DataLink communication). The simulation was run on a Pentium III desktop computer. Participants' primary task was to monitor the progress of airplanes and avoid potential conflicts in their sector. They also had to climb and descend aircraft to predefined altitudes at certain points along their flight path and accept and issue handoffs from and to other sectors in a timely manner. Finally, they were asked to report any unusual events (such as an airplane deviating from its assigned route) as quickly as possible.

Interruption task. While performing the ATC task, participants were asked to handle interruption tasks that were presented via the visual, auditory, or tactile modality. A pending task was indicated by an initial cue (a red flashing box at the top left of the interruption message list; see Figure 1). Participants were asked to push the space bar on the keyboard as soon as they noticed the initial cue. Each interruption task consisted of the presentation of 12 to 15 slow and fast pulsing patterns. Participants were told to count and report the number of fast patterns only. Visual interruption tasks consisted of two filled circles flashing at the bottom of the screen (see Figure 1); auditory tasks involved beeping sounds that were presented via a headset; and the tactile signals were vibrations delivered by tactors to participants' left and right inner wrists.

Interruption tasks varied in terms of their priority. Urgent tasks started immediately after participants responded to the initial cue. If participants missed the cue, the task started automatically after 5 s. All urgent tasks were announced by an orienting cue: two white circles flashed once before the start of a visual task, a distinct tone was presented before an auditory task, and both tactors buzzed briefly before the onset of a tactile task. In the case of low-priority tasks, participants could delay task initiation for up to 2 min.

A time stamp on the interruption message list served as an external reminder of the pending interruption task. When presented with a
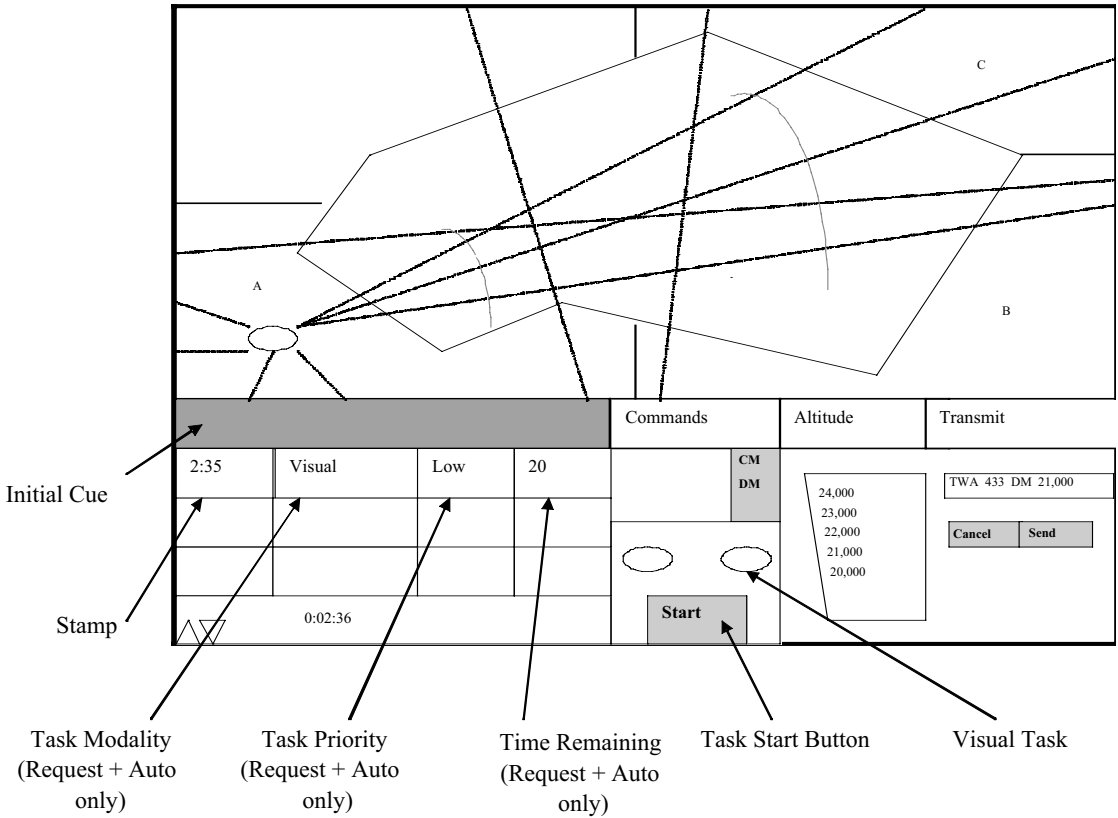
Commands | Altitude | Transmit

Initial Cue

2:35 | Visual | Low | 20

CM DM

24,000
23,000
22,000
21,000
20,000

TWA 433 DM 21,000

Cancel | Send

Stamp

0:02:36

Start

Task Modality
(Request + Auto
only)

Task Priority
(Request + Auto
only)

Time Remaining
(Request + Auto
only)

Task Start Button

Visual Task

Figure 1. The ATC simulation and task-related information.

low-priority task, participants in the request group could click anywhere on the message list to request additional information about task modality, priority, and the time remaining for performing the task. In contrast, participants in the auto group received this information automatically as soon as the initial cue appeared.

Procedure

Participants completed two sessions, each lasting approximately 1.5 hr. In the first session, participants received 30 min of training on the ATC task. They then performed the ATC task on their own for another 20 min. Finally, they received 20 min of training on handling the different types of interruption tasks in parallel with the ongoing ATC task. Only participants who made fewer than three mistakes in any ATC task or interruption task during the last 20 min of combined training were asked to participate in the actual experiment. Only 1 participant was rejected as a result of failing to meet this cri-

terion. The actual experimental session started with a 10-min review of the ATC and interruption tasks, followed by a 1-hr experimental scenario.

Scenario

The experimental scenario consisted of four different phases that varied in terms of workload and interruption frequency (Table 1). During the high-workload period, participants were working approximately 16 airplanes in their sector. This resulted in higher monitoring demands

TABLE 1: Overview of the Four Scenario Phases

Phase	Duration (min)	Workload	Interruption Frequency
LH	10	L(ow)	H(igh)
HH	10	H(igh)	H(igh)
HL	20	H(igh)	L(ow)
LL	20	L(ow)	L(ow)

and the need for more control inputs as compared with the low-workload period, in which participants were controlling approximately 8 airplanes. The scenario also varied in terms of the frequency of interruptions. Participants received one urgent and one low-priority interruption task in each of the three modalities (visual, auditory, and tactile) during each of the four phases. Two of those phases (one each of the high- and low-workload conditions) lasted for 20 min (resulting in low interruption frequency), and the duration of the other two phases was only 10 min (high interruption frequency). Workload and interruption frequency were counterbalanced.

Participants also had to detect and report the deviation of three different airplanes from their original route. All deviations occurred during high-workload conditions. The first deviation occurred while participants performed an urgent visual interruption task. The second airplane deviated when participants performed an urgent tactile interruption task. A third deviation took place when a low-priority tactile interruption task was presented.

Dependent Measures

The indicators of participants' primary (ATC) and interruption task performance listed in Table 2 served as dependent measures in this study.

RESULTS

Latency of Information Request

On average, participants in the request group accessed task-related information for low-priority interruption tasks within 5.9 s ($SD = 9.4$) of detecting the initial cue. In 61% of the cases, the information was requested within 2 s (see Figure 2).

Low-Priority Task Initiation Delay

For low-priority interruption tasks, participants could wait as long as 2 min before initiating the task. A repeated-measures analysis of variance (ANOVA) yielded a significant main effect for task modality. On average, visual interruption tasks were initiated after 22.4 s, whereas auditory and tactile tasks were started after 15.7 and 14.7 s, respectively, $F(2, 46) = 6.738$, $p < .01$ (see Figure 3). Bonferroni tests indicated that the visual tasks were delayed significantly longer than the auditory and tactile tasks. There was no significant difference between the latter two.

Incidence of Concurrent Performance of ATC and Interruption Tasks

A repeated-measures ANOVA yielded significant main effects for both workload, $F(1, 45) = 89.95$, $p < .0005$, and task modality, $F(2, 90) = 26.87$, $p < .001$. During periods of high ATC workload, concurrent performance of ATC and interruption tasks was observed significantly more often (see Figure 4). Also, auditory and tactile interruption tasks were performed concurrently with ATC more often than visual interruption tasks (see Figure 5).

Interruption Task Performance

The three experimental groups did not differ significantly in terms of their overall interruption task performance. The request, auto, and none groups performed accurately in 90.3%, 90.6%, and 91.9% of the interruption tasks, respectively. Next, the effects of simultaneity and modality were examined. There was no

TABLE 2: Dependent Measures

ATC performance
    No. of traffic conflicts (<5 miles of separation when aircraft within 1000
        feet of altitude)
    No. of late or missed handoffs
    Detection of course deviation
Interruption task performance
    Latency of information request
    Low-priority task initiation time
    Incidence of concurrent task performance (ATC and interruption task)
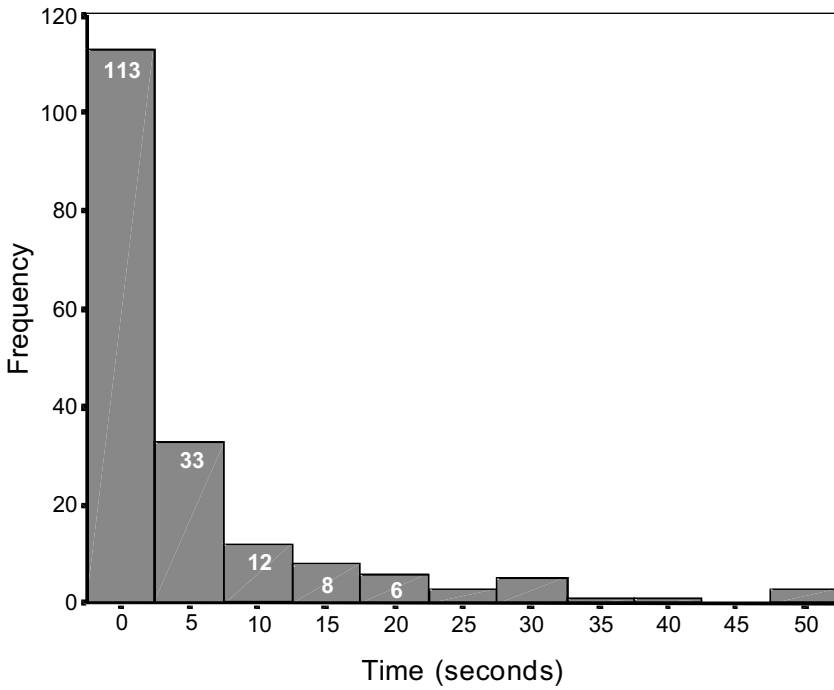   Correct counts of fast patterns

Figure 2. Time between detection of initial cue and request for additional information on low-priority tasks (for the request group only).
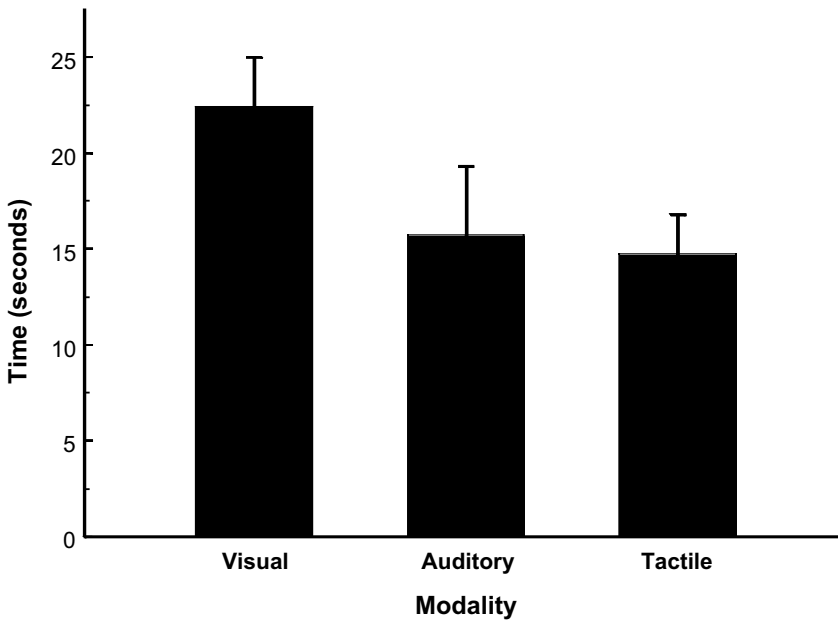

Figure 3. Low-priority task initiation time as a function of interruption task modality (for the request and auto groups only).
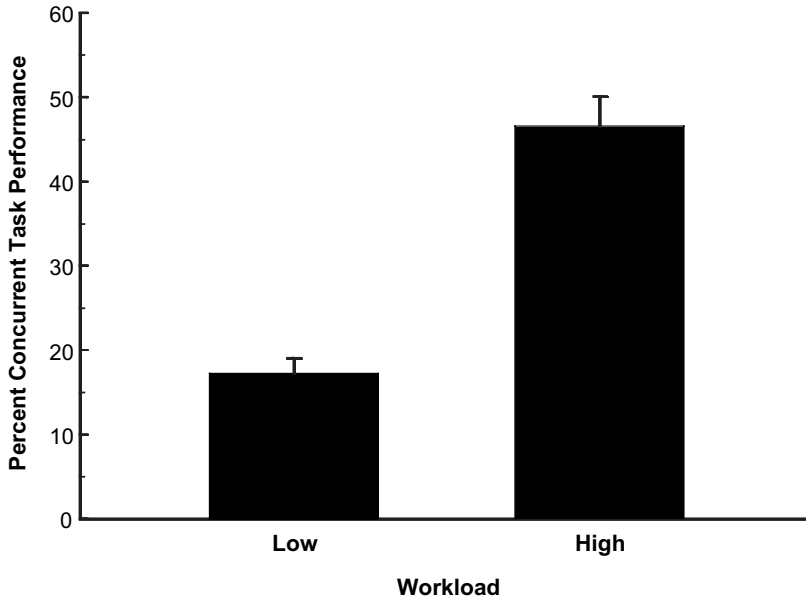
Figure 4. Incidence of concurrent performance of ATC and interruption tasks as a function of ATC workload.

significant difference among the three interruption tasks – visual, auditory, and tactile – when they were performed in isolation. However, a significant interaction between modality and simultaneity was found, $F(2, 54) = 3.844$, $p = .027$, such that performance on the visual and tactile (but not on the auditory) interruption

tasks suffered significantly when these tasks were performed concurrently with ATC, $t(33) = –3.874$, $p < .001$, and $t(44) = –2.362$, $p = .023$, respectively, based on paired t tests (see Figure 6).

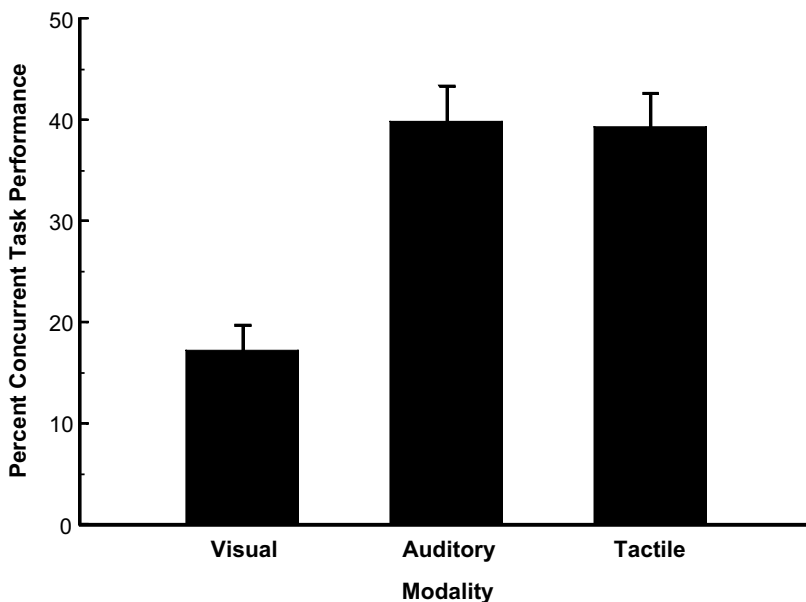Interruption task performance was also compared for low-priority interruption tasks, for

Figure 5. Incidence of concurrent performance of ATC and interruption tasks as a function of task modality.
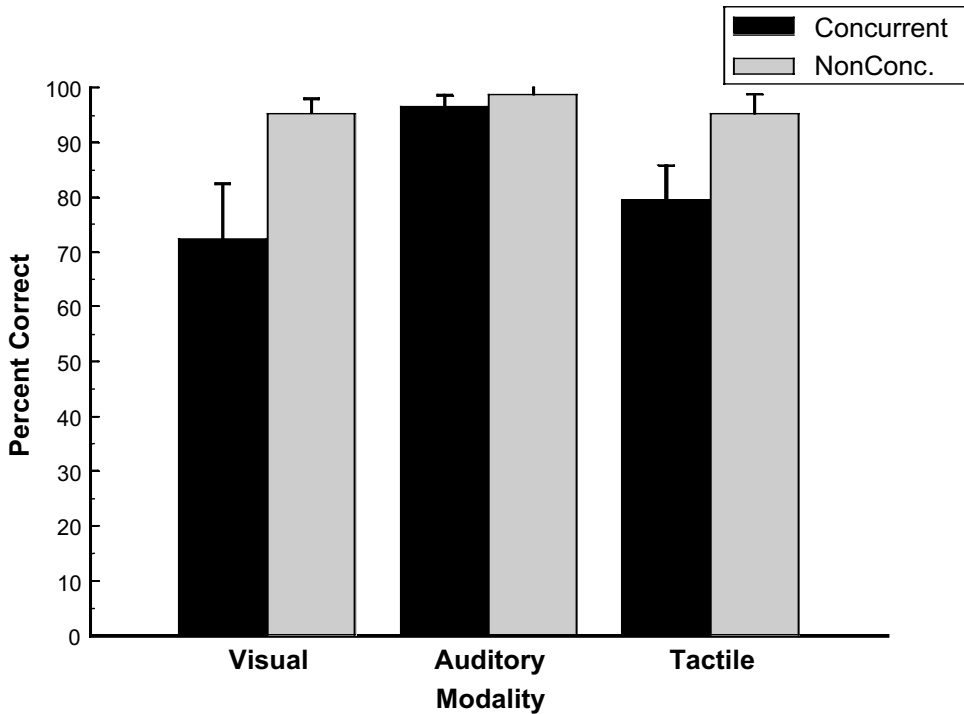
Figure 6. Interruption task performance as a function of task modality and simultaneity.

which participants had control over the initiation delay, and urgent tasks, which started automatically. Overall, participants performed significantly better on low-priority interruption tasks, $F(1, 45) = 6.29$, $p = .016$. They correctly counted the number of fast patterns in 93.3% of the low-priority tasks as opposed to 88.5% of the urgent tasks (see Figure 7). A closer examination of this finding shows that it is attributable primarily to the fact that participants in the request group performed significantly better (in terms of correct counts of fast patterns) on low-priority tasks ($M = 85\%$, $SD = 11\%$) than on urgent ones ($M = 96\%$, $SD = 6.5\%$), $t(15) = -3.13$, $p = .007$ (see Figure 8).

ATC Performance

The number of delayed or missed handoffs in the auto group (15 cases) was significantly smaller than in the none group (38 cases), $F(1, 31) = 4.738$, $p = .038$. However, there was no significant difference between the request group (22 cases) and the other two groups. There was also a trend toward fewer traffic conflicts in the auto group (one conflict) as compared with

the none and request groups (four conflicts each). Finally, participants in the auto group tended to detect the deviation of an airplane from its course more often (94%) than did participants in the none group (89%) and the request group (81%).

Table 3 provides an overview of the performance data for the three experimental groups (none, request, and auto) for both the ATC and the interruption tasks.

DISCUSSION

The findings from this study illustrate the effectiveness of providing partial information about the nature of pending tasks for supporting task and interruption management. The perceived high value of task-related information is illustrated by the fact that in 61% of the low-priority tasks (which allowed for a 2-min delay), participants in the request group accessed this information within 2 s of detecting the initial cue. They chose to do so even though this step required interaction with the visual interface and, thus, a short orientation away from their
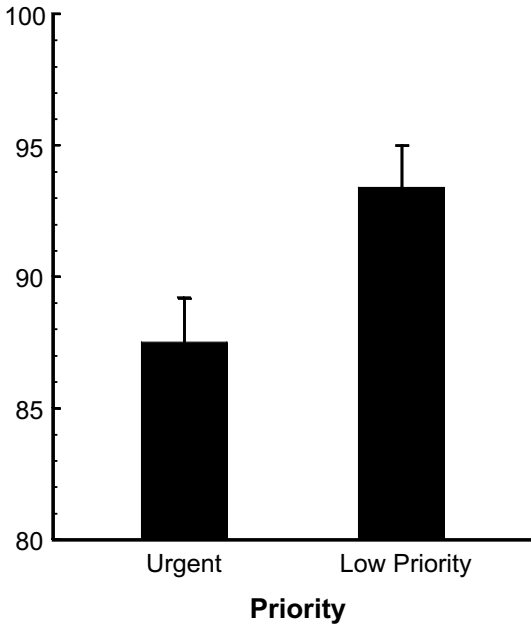
Figure 7. Interruption task performance as a function of task priority.

tasks, knowledge of task modality led participants to delay visual tasks significantly longer than both the auditory and tactile tasks. Also, concurrent performance of a visual interruption task with the ATC task was observed less often than with tasks in either of the other two modalities. Instead, participants tried to complete ATC-related tasks first.

The appropriateness of this strategy, and the predictions made by multiple resource theory, are supported by the fact that interruption task performance suffered the most when participants had to perform two visual tasks simultaneously. It is interesting to note that the performance of tactile interruption tasks (but not of auditory interruption tasks) also suffered significantly when they were performed concurrently with the visual ATC task. This appears to contradict findings from some earlier studies that examined cross-modal interference between visual and tactile tasks (e.g., Sklar & Sarter, 1999; Spence & Driver, 1997). For example, the tactile signals that were presented in the Sklar and Sarter study to alert pilots to unexpected mode changes did not interfere with the continuous visual task of flying the airplane. Note, however, that this can be explained by the discrete nature of those cues; in contrast, the present study employed a continuous tactile task.
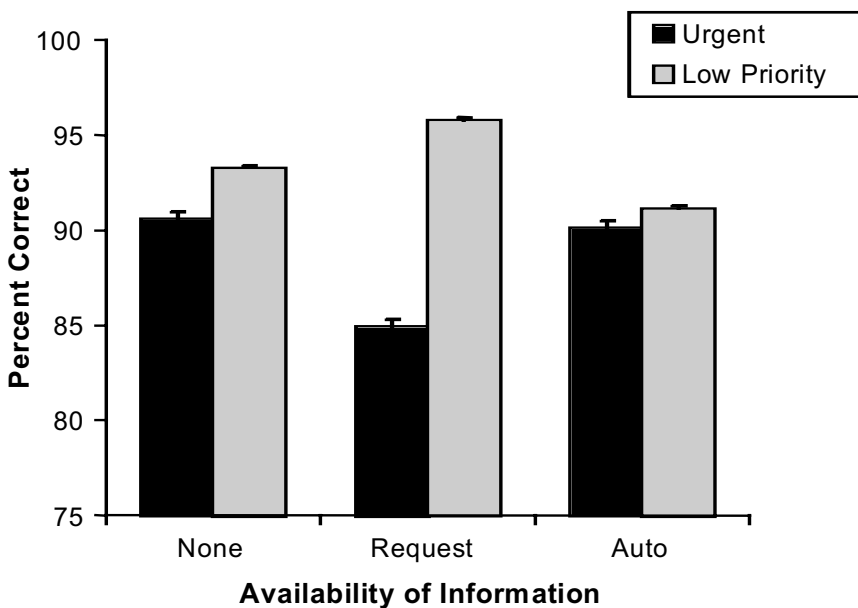
primary ATC task. The information not only was perceived to be useful but also actually helped operators schedule and manage tasks and associated attentional resources, as had been predicted by Latorella (1999) and McFarlane and Latorella (2002). In the case of low-priority

Figure 8. Interruption task performance as a function of task priority and information availability.

TABLE 3: Overview of Performance Data for the Three Experimental Groups

	None	Request	Auto
Interruption Task Performance (% correct counts)			
Visual	93.7	90.3	90.6
Auditory	93.0	92.1	93.8
Tactile	89.1	88.3	87.5
Air Traffic Control Performance			
Delayed/missed handoffs	38	22	15
Traffic conflict(s)	4	4	1
Detection of course deviation (%)	89	81	94

Interruption task performance was significantly better in the case of nonurgent tasks, for which participants had some control over the scheduling of the task. They could engage in what Latorella (1999) referred to as a "negotiated interruption" (as opposed to an immediate interruption), in which the operator has control over when to deal with an interruption. A more detailed analysis showed that this finding is attributable to the performance difference between urgent and nonurgent tasks in the request group only. The performance of the none and auto groups did not differ for these two types of tasks. This result confirms that control over task initiation is not sufficient for improving task scheduling; instead, it needs to be combined with at least partial information about the nature of a pending task (Latorella, 1999). As explained earlier, the request group did not receive task-related information in the case of urgent tasks but could, and did, request the information for low-priority tasks. It appears that this helped them to avoid unintentional dismissals and preemptive integration of interruption tasks.

Overall, the availability of task-related information did not result in a difference in interruption task performance for the three experimental groups. To some extent, this finding may be the result of a ceiling effect: All three groups performed at or above 90% accuracy. A contributing factor may have been the amount of effort that participants invested in their competing tasks. Participants who did not receive information about the interruption tasks may have tried to compensate for this deficiency by investing more resources in those tasks. In contrast, those participants who received task-related information may have decided that they could focus on the ATC tasks instead. This interpretation is supported by the fact that participants in the auto group performed better on some aspects of the ATC task. Thus, although the interruption task performance of this group was not superior, they still achieved a net gain in terms of overall performance.

CONCLUSIONS

In conclusion, the findings from this study confirm the usefulness of providing operators with partial information about the nature of competing attentional demands for improving task management. As systems continue to become more complex and involve larger numbers of collaborating agents, the need for this form of support will probably increase.

To meet this need, a number of issues still need to be addressed. In the current experiment, information on pending tasks was presented in visual form, thus requiring participants to orient their attention away from the ongoing visual ATC task for at least a brief moment. A more effective way to provide this information and thus fulfill the second criterion for preattentive reference – namely, peripheral access to information – would be to present it via other modalities, such as hearing or touch. Also, most research has focused on situations involving only two or few competing demands when, in fact, an operator could find himself or herself in a much more complex situation. This will require even more effective support for task management and may call for the involvement of an external mediator who can assess, predict, and adapt to operator interruptibility based on cognitive models, physiological measurements,

or shared access to the operator's workspace (see, e.g., Czerwinski, Cutrell, & Horvitz, 2000; Kuperman & Perez, 1988; Latorella, 1998). Progress in these areas is needed in order to avoid researchers and designers that are being outpaced, once again, by rapid developments in the modern workplace.

ACKNOWLEDGMENTS

REFERENCES

Czerwinski, M., Cutrell, E., & Horvitz, E. (2000). Instant messaging and interruption: Influence of task type on performance. In C. Paris, N. Ozkan, S. Howard, & S. Lu (Eds.), *OZCHI 2000 Conference proceedings* (pp. 356–361). Narrabundah: Ergonomics Society of Australia.

Dismukes, K., Young, G., & Sumwalt, R. (1998). Cockpit interruptions and distractions: Effective management requires a careful balancing act. *Aviation Safety Reporting System Directline, 10.* Retrieved July 28, 2004, from http://asrs.arc.nasa.gov/directline_issues/dl10_distract.htm

Funk, K., & Braune, R. (1999). The agenda manager: A knowledge-based system to facilitate the management of flight deck activities. In *Proceedings of the 1999 World Aviation Congress* (pp. 1–15). San Francisco: Society of Automotive Engineers.

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

Kuperman, G. G., & Perez, W. A. (1988). A frame-based mission decomposition model. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 135–139). Santa Monica, CA: Human Factors and Ergonomics Society.

Latorella, K. A. (1998). Effects of modality on interrupted flight deck performance: Implications for DataLink. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 87–91). Santa Monica, CA: Human Factors and Ergonomics Society.

Latorella, K. A. (1999). *Investigating interruptions: Implications for flight deck performance* (Tech. Memorandum NASA/TM-1999-209707). Hampton, VA: NASA-Langley Research Center.

McFarlane, D. C., & Latorella, K. A. (2002). The scope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction, 17,* 1–61.

Sarter, N. B. (2002). Multimodal information presentation in support of human-automation communication and coordination. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research* (pp. 13–36). New York: JAI.

Sklar, A. E., & Sarter, N. B. (1999). "Good vibrations": The use of tactile feedback in support of mode awareness on advanced technology aircraft. *Human Factors, 41,* 543–552.

Sorkin, R. D., Kantowitz, B. H., & Kantowitz, S. C. (1988). Likelihood alarm displays. *Human Factors, 30,* 445–459.

Spence, C., & Driver, J. (1997). Cross-modal links in attention between audition, vision, and touch: Implications for interface design. *International Journal of Cognitive Ergonomics, 1,* 351–373.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention* (pp. 63–101). Orlando, FL: Academic.

Woods, D. D. (1995). The alarm problem and directed attention in dynamic fault management. *Ergonomics, 38,* 2371–2393.

Chih-Yuan Ho received her M.S. in industrial and systems engineering from Ohio State University in 2004.

Mark I. Nikolic is a postdoctoral student at Ohio State University, where he received his Ph.D. in industrial and systems engineering in 2004.

Molly J. Waters is a graduate student in the Department of Industrial and Systems Engineering/Institute for Ergonomics at Ohio State University, where she received a B.S. in industrial and systems engineering in 2001.

Nadine B. Sarter is an associate professor in the Department of Industrial and Operations Engineering and the Center for Ergonomics at the University of Michigan. She received her Ph.D. in industrial and systems engineering from Ohio State University in 1994.