The benefits of practice with interruptions is step-specific

Kevin Zish (kzish@gmu.edu**)**
George Mason University
4400 University Dr, Fairfax, VA 22030

J. Malcolm McCurry (malcolm.mccurry.ctr@nrl.navy.mil**)**
Harris
4555 Overlook Ave SW, Washington, DC 20375

J. Gregory Trafton (greg.trafton@nrl.navy.mil**)**
U.S. Naval Research Laboratory
4555 Overlook Ave SW, Washington, DC 20375

Abstract

In two studies we investigated the effect of resumption practice following an interruption at the same step in a Computerized Physician Order Entry system (CPOE). The results of both studies showed that error rate decreased with increasing amounts of resumption practice. One reason people may have resumed more accurately following an interruption is improvement in a general resumption process. If true, we would expect that participants could be interrupted at any step in a task and show improved resumption with increased practice. Instead, our results suggest that repeatedly resuming from the same step likely produces associative priming between a specific task, interruption, and step. The associative priming allowed participants to resume more successfully with additional interruption practice, but only for that task-interruption-step triplet.

Keywords: interruption, skill acquisition, practice, errors, memory for goals

Introduction

In complex environments such as emergency rooms and critical care units, medical professionals are interrupted nearly ten times an hour (Chisholm et al., 2001). These interruptions have particularly unsettling effects to patient outcomes and are difficult to remove from the environment.

Although interruptions have been studied across multiple measures, especially error rates and resumption time, few methods exist to make interruptions less disruptive (Edwards & Gronlund, 1998; Cutrell et al., 2001; Eyrolle & Cellier, 2000; González & Mark, 2004; Hodgetts & Jones, 2006; Monk, 2004; Ratwani et al., 2008; Trafton et al., 2003). One consideration for reducing the negative effects of interruptions has been practice of the task. For example, Trafton et al. (2003) and Cades et al., (2006) showed an improvement in the ability to resume a task following repeated exposure to interruptions.

To determine if improvement on task resumption is due to practice on the primary or secondary task, Cades et al. (2011) used a clever paradigm which combined two interruption types with one primary task. Some participants had one or both interruption types presented during the primary task across multiple sessions. The results of their study suggest that participants showed improvement on resuming the primary task only when they received practice with the same interruption type across each session. Interestingly, there was almost no benefit in mitigating the disruptiveness of interruptions with practicing the primary task alone or switching to another interruption type.

The findings from Cades et al. (2011) suggests that people can improve their ability to resume after an interruption. However, the benefit of practice only extends to the resumption of a specific task and interruption pair. When the interruption was switched to an interruption with less or no practice, no improvement was present. Cades et al. (2011) explained the improvement in task resumption following an interruption as the result of the interruption task priming the primary task. Presumably, when a new interruption type is introduced, the new interruption provides little to no priming to the primary task.

Memory for Goals (*MFG*: Altmann & Trafton, 2002), has had great success in explaining why interruptions are so disruptive (Bailey & Konstan, 2006; Ratwani et al., 2008) and why practice improves performance after an interruption for specific task-interruption pairs. *MFG* is an activation-based model built in the *ACT-R* cognitive architecture. Activation is instantiated in *ACT-R* (Anderson, 2007) as the probability that a step in memory will be retrieved.

According to *MFG*, activation in procedural tasks is determined by the strengthening and priming constraints. The strengthening constraint considers how often a goal has been retrieved in the past and the recency of retrieval. As a result, goals that have been retrieved more often will have greater activation than goals that have been retrieved less often. Additionally, goals that have been retrieved in the recent past will have more activation than goals retrieved in the distant past.

The priming constraint suggests that activation for a goal can also increase with environmental cues. For example, an interrupted goal, such as putting cream in coffee, can be cued to resume by seeing that the coffee is still dark in color. One major prediction of the priming constraint is the creation of associative links between goals in a task. The role of an associative link is to prime all future goals of the task (Altmann & Trafton, 2007, 2015).

MFG predicts that strengthening, priming, and the natural decay of activation over time is responsible for cognitive control over the current goal of a task. Using the task of making coffee as an example, the act of pouring coffee into a cup would prime the goal to put in sugar. Sugar would receive the maximum amount of priming if it were always preceded by pouring coffee. In a procedural task, interruptions have the effect of reducing activation by preventing people from rehearsing steps of the task they have already completed successfully. Furthermore, interruptions also allow cumulative priming to future steps of the task from associative links to dissipate. In both cases, a reduction in activation for the correct step of the task increases the chance that people will retrieve the wrong step and make an error.

MFG predicts that a specificity in priming following a task/interruption pair is related to the associative links established in procedural tasks. Following practice with interruptions, improvement in resumption of the primary task could occur in multiple ways. One possibility is that general resumption practice with the same interruption type at any point in the task is important (Cades et al., 2011, 2006; Trafton et al., 2003). A second (and much less convenient) possibility is that improvements in resumption are much more specific and that the result of improvements in resumption only extend to resumption of a primary task, following a specific interruption, for a particular step in the primary task.

Current theories of skill transfer in *ACT-R* suggest that a task-interruption-step triplet is the likely source of improvement in interruption practice paradigms (Singley & Anderson, 1989). According to *ACT-R* if goals are repeatedly performed together, they are grouped into one procedure. Transfer between tasks would be strongest when the goals of one practiced part of the task are identical to the goals following an interruption.

One popular method to measure the effects of practicing interruptions has been to examine resumption lag. Resumption lag is the amount of time it takes to resume a task after an interruption (Altmann & Trafton, 2004; Cades et al., 2011, 2006; Monk, 2004; Trafton et al., 2003, 2005). While time effects (Gray & Boehm-Davis, 2000) are important, errors typically have a more meaningful impact in real-world settings. However, the degree of environmental cues used in applied settings make errors relatively rare. Because errors are rare, investigating if accuracy also responds to practice with a task/interruption pair in the same way as Cades et al. (2011) can be challenging.

In this study we investigated if errors in procedural tasks also respond to practice with interruptions. To help generate more errors for study, we focused on the post-completion step (PCS) of a procedural task, which can be especially error prone (Ratwani et al., 2008). A post-completion step occurs when there is an extra step after the goal of a task has been completed. An error at the post-completion step is called a post-completion error (PCE). A common example of a PCE is failing to take back a bank card after receiving money from an ATM (Byrne & Bovair, 1997).

While forgetting an ATM card can be a nuisance, the medical field has demonstrated many cases in which errors at the post-completion step can have dangerous outcomes. Electronic health record systems (EHRs) and computerized physician order entry systems (CPOEs) have been implemented to reduce many errors related to poor-handwriting and dosage miscalculations (Koppel et al., 2005). However, these systems have created new situations in which errors occur. Sometimes the user can forget to close the record of a patient or switch the record to a new patient after ordering medications, tests, or writing in charts. The post-completion step of closing one order before opening a new one has led to serious errors where medical information or medications are switched between patients (Ash et al., 2004; Hettinger & Fairbanks, 2012; Kim et al., 2006). These errors have resulted in severe harm.

Can resumption practice reduce the rate of PCEs following an interruption? *MFG* suggests that increased practice in resuming the primary task from a task/interruption pair will lead to the formation of associative links between an interruption and the primary task. *MFG* would also predict that associative links should form for a specific task-interruption-step triplet. Specifically, there is a greater benefit to practicing with interruptions when resumption occurs on the same step versus practice with interruptions throughout the task. Priming should be passed from the interruption to the correct step of the primary task when the same step is resumed on multiple occasions. This should result in greater performance only after practicing resumption of a specific step—despite the number of interruptions practiced at other places in the primary task.

Experiment 1

Method

Participants

Undergraduate students who were enrolled in at least one psychology course at George Mason University gave informed consent and participated for partial course credit. Fifty-seven participants completed Experiment 1.

Task

Primary Task Participants were presented with a simulated computerized physician order entry (CPOE) system based on CPOEs used in medical offices to help practitioners fill out patient orders and history (Figure 1). Participants were asked to complete CPOE orders. To complete an order, participants would fill in the sections of a CPOE form using information provided on the monitor. The information needed to complete the order was always on the monitor when the CPOE window was active. Orders were randomly given a status of Urgent, Priority, and Normal. Sets of orders were presented three at a time on the top right of the monitor.

Starting the task required the participant to choose the order with the most urgent status located on the right of the screen. If two items had the same urgency, participants were told that they could choose either order first. A single mouse click activated the order and revealed information that provided all the necessary information to complete the task. Relevant patient information including the name and order type was located at the top of the monitor. Each patient had four order types that participants could select. On the CPOE system this was represented as tabs named Patient Factors, Labs, Medication, and History. Only the first three tabs were used in this experiment. Participants used this information to first select the patient name on the top of the screen and then the order type. Accuracy was enforced by not allowing the participant to continue until both the name and type matched the active order. The system would notify the user with a beep when an incorrect entry or section of the CPOE was selected.

Participants were instructed to use the information on the monitor to fill out the sections of the CPOE form left to right and top down. The system enforced completeness but not correctness for this portion of the task. Once a CPOE order was complete, participants were instructed to click the “Send Email to Doctor” button to notify the attending doctor that the order was submitted. Instructions were to click this button only once to reduce the likelihood of confusing the doctor with identical or near-identical orders. After this button was pressed, a dialog box appeared with the words “Email Sent.” Participants accepted this message by clicking “OK” on the screen.

The final step of the task was to click “Close Record” which would allow the participant to select the next order. Clicking on this button is the post-completion step. Because our focus is on the PCE, our analysis is on the actions following the post-completion step.

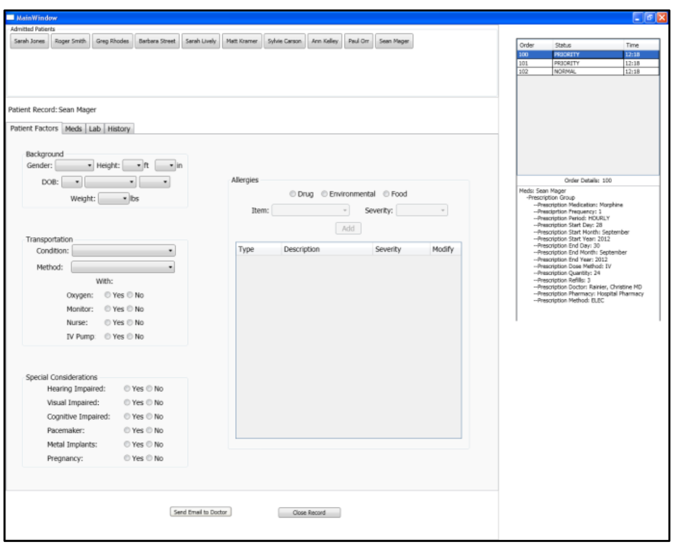

Figure 1. CPOE system main task

Secondary Task Participants were also asked to complete a secondary interruption task. We chose subtraction problems because the process of solving a math problem reduces the likelihood participants can simultaneously rehearse the last step of the primary task. Certain buttons or sections of the CPOE task would trigger an interruption. The buttons or sections of the CPOE form that would trigger an interruption were chosen randomly for each participant. Furthermore, interruptions occluded the entirety of the main task, reducing environmental cues that could be used to easily resume the primary task. Two-digit subtraction problems appeared with four answers in the center of the screen for 15s (Figure 2). Participants were instructed to complete the problems as quickly and as accurately as possible. If the participant chose the incorrect answer, the system would beep, and the selected choice would turn red. Once 15 seconds has passed, the secondary task would immediately be replaced with the primary task.

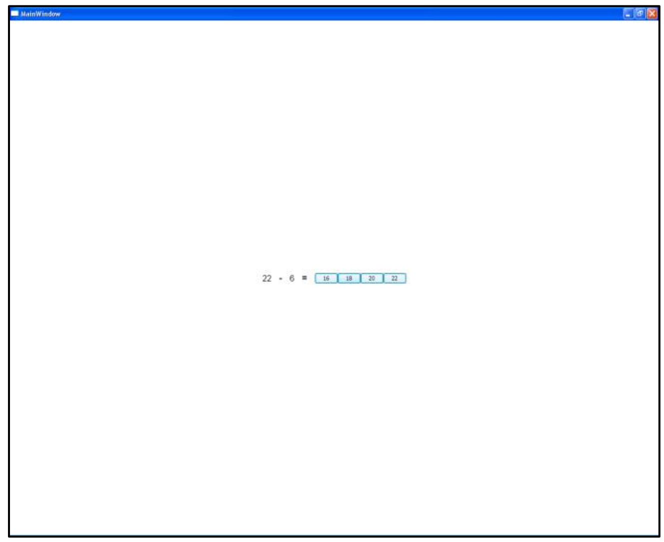

Figure 2. CPOE math interruption.

Design

Participants were assigned to one of three conditions in a 3 (between: interruption practice) x 2 (within: interruption/no interruption control) design. Each condition consisted of 6, 12, or 18 interruptions throughout the task. Only one interruption appeared on any given CPOE form. The CPOE sections and buttons that triggered the interrupting task were equally split between post-completion and non-post-completion widgets. Therefore, three, six, or nine of the CPOE forms were interrupted at the post-completion step. The remaining interruptions were randomly located at non-PCS widgets to decrease the likelihood that participants would prepare for an interruption at the PCS.

Out of 24 CPOE orders that were completed throughout the task, participants had both interruption and non-interruption orders for a within-participants design. Filling in all the information and clicking “Close Record” was recorded as completing an order.

Procedure

Each participant filled out an approved IRB consent form. Demographic information was taken before each participant was trained in the task. Participants were seated approximately 47 cm from the computer monitor. Directions were initially given using a series of screenshots of the CPOE system and the math interruption. Afterwards, three practice orders were presented to the participants to provide the opportunity to ask questions about how the system operates. Participants were asked to complete the orders as quickly and accurately as possible. The experiments were completed without the experimenter being present in the room. All participants were debriefed and dismissed upon completion.

Measures

Accuracy at each section of the CPOE form was collected for all participants. An error at the post-completion step was defined as selecting any other widget other than the "Close Record" button on the CPOE interface. Errors were calculated as a percentage of incorrect selections over the opportunity to correctly complete the post-completion step between interrupted and uninterrupted control trials.

Results

Error Rates

Fifty-seven participants in Experiment 1 made a total of 105 errors at the post-completion step. The proportion of errors were analyzed using a mixed-model ANOVA using the number of interruptions at the PCS as a covariate.

Figure 3 demonstrates the effect of the number of interruptions at the PCS. There was a main effect of interruption such that the error rate was significantly higher at the PCS after an interruption (Interrupted: M = 31.23%; Control: M = 1.20%), $F(1, 55) = 82.76$, $MSE = 25693.00$, $p < .05$, $\eta 2 = .39$. The low error-rate for control trials suggests that the task was well-learned. There was also a main effect of the number of interruptions at the PCS, $F(1,55) = 10.09$, $MSE = 3123.40$, $p < .05$, $\eta 2 = .05$. A significant interaction between interrupted trials and the number of interruptions at the PCS suggests that error rates at the PCS without interruptions did not respond in the same way to an increase in interruptions, $F(1, 55) = 10.82$, $MSE = 3359.00$, $p < .05$, $\eta 2 = .05$.

A polynomial contrast was performed to determine if the error rates changed reliably across increasing numbers of interruptions for the interruption condition. A significant linear contrast suggests that error rates decreased with an increase in practice resuming at the PCS, $t(2,54) = -3.21$, $p < .05$.

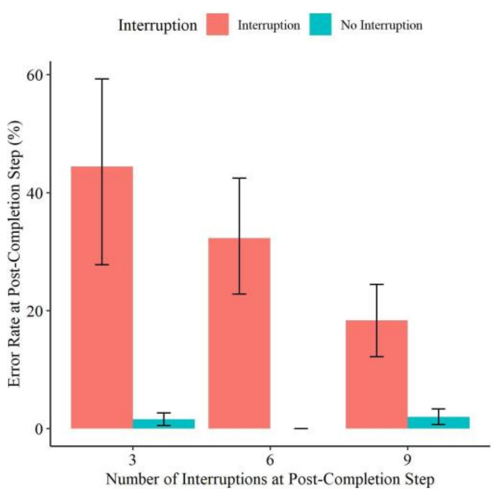

Figure 3. Error results for Experiment 1 including 95% CI.

Experiment 2

In Experiment 1, errors decreased linearly with additional resumption practice at the PCS. One possibility is that errors decreased with additional practice because participants had more practice with interruptions on the task overall. To investigate whether a general increase in resumption practice is the cause of a decrease in errors in Experiment 1, we ran a second experiment which kept the total number of interruptions equal across conditions and varied the proportion of interruptions at the PCS.

Method

Participants

Undergraduate students who were enrolled in at least one psychology course at George Mason University gave informed consent and participated for partial course credit. Forty-six participants completed Experiment 2. Data from Experiment 2 was collected after Experiment 1 during the same semester. Participants from Experiment 2 were not a part of the sample for Experiment 1.

Tasks

The primary task and secondary task for Experiment 2 followed the same protocol as Experiment 1.

Design

Each participant received interrupted and uninterrupted opportunities to complete the task. The two main differences in Experiment 2 were the total number of interruptions and proportion of interruptions at the PCS. For Experiment 2, participants had a total of 18 interruptions meaning that the total practice resuming from an interruption across the task was equal for each condition. Participants were given either 3, 6, or 9 interruptions at the PCS with all other interruptions spread randomly throughout the rest of the task.

Procedure The procedure was the same as Experiment 1.

Measures All measures were the same as in Experiment 1.

Results

Error Rates

Forty-six participants in Experiment 2 made a total of 74 errors at the post-completion step. The proportion of errors were analyzed using a mixed-model ANOVA using the number of interruptions at the PCS as a covariate.

Figure 4 demonstrates the effect of the number of interruptions at the PCS. There was a main effect of interruption such that the error rate was significantly higher at the PCS after an interruption (Interrupted: M = 27.62%; Control: M = 1.11%), $F(1, 44) = 57.75$, $MSE = 16170.00$, $p < .05$, $\eta 2 = .37$. The low error-rate for control trials again suggests that the task was well-learned. Like Experiment 1 there was a main effect of the number of interruptions at the PCS, $F(1,44) = 4.61$, $MSE = 1333.60$, $p < .05$, $\eta 2 = .03$. A significant interaction between interrupted trials and number of interruptions at the PCS suggests that error rates at the PCS without interruptions did not respond in the same way to an increase in interruptions, $F(1, 44) = 5.15$, $MSE = 1441.00$, $p < .05$, $\eta 2 = .03$.

A polynomial contrast was performed to determine if the error rates changed reliably across increasing number of interruptions for the interruption condition. A significant linear contrast suggested that error rates decreased with an increase in practice resuming at the PCS, $t(2,43) = -2.21$, $p < .05$.

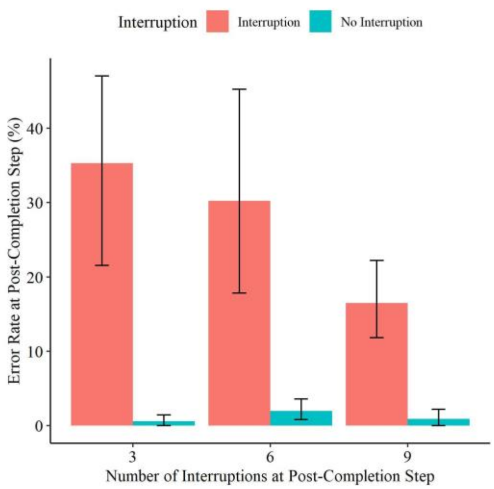

Figure 4 Error results for Experiment 2 including 95% CI.

Discussion

Experiments 1 and 2 investigated the role of resumption practice in mitigating interruptions. Unsurprisingly, participants were more error-prone in both experiments when resuming the post-completion step following an interruption.

The first study considered if increasing resumption practice at the post-completion step reduced error rates. When not interrupted, participant's error rates were low suggesting a well-learned procedure. Participant's had 3, 6, or 9 interruptions at the PCS and an equal number of interruptions at other places in the task.

A significant linear contrast for error rates suggests that as participants increased the amount of practice they had resuming from interruptions, their error rate decreased at the post-completion step. However, it was not clear if the decrease in error rates at the PCS were due to specific practice at the PCS or the increase in practice with interruptions across the task overall. To determine if the decrease in error rates was due to some benefit in resuming at the PCS or general resumption practice, a follow-up study kept the number of interruptions equal across all three conditions. Participants had 18 interruptions across the task with 3, 6, or 9 of the 18 interruptions just following a post-completion step. Again, a low error rate suggested a well-learned procedure.

A significant linear contrast showed that as participants increased the amount of practice resuming at the PCS, error rates decreased.

Theoretical Implications

Cades et al. (2011) demonstrated a decrease in resumption lag when participants practiced resuming from a task-interruption pair. Experiment 1 and Experiment 2 extend their original findings to errors.

More importantly, the increase in performance after an interruption seems not to come from learning to resume at any place in the task. Practice benefits the participant the most when resuming at a specific step in the task. Taken together with Cades et al. (2011), performance increases are found when resumption is practiced with the same task-interruption pair at the same step.

One explanation for why we found improvement in performance for a task-interruption-step triplet comes from *ACT-R*: resumption of the primary task at the same place after a math problem may have formed a new procedure. Resuming the PCS after a math problem became a routine part of the CPOE task. Another explanation comes from *MFG*: persistent practice with the task-interruption-step triplet forms associative links for completing the task correctly. According to the priming constraint of *MFG*, associative links formed by error-free practice would prime the next correct step in the task, regardless of an interruption. Associative links would result in accurate continuity of the procedure.

Because performance was best after practicing a task-specific triplet, this study suggests that the largest benefit from practicing with interruptions is when the conditions of practice overlap substantially with the environmental conditions of resumption. One open question is why Cades et al. (2011) found a benefit for practicing with interruptions even though interruptions occurred at different steps in the task. A likely cause of their results was that participants were able to form associative links for task-specific triplets, but the associative link occurred for different steps. The results of their study would likely be even stronger if resuming from a specific interruption type and step were practiced repeatedly.

Practical Implications

Experiment 1 and Experiment 2 demonstrates that reducing error rate following an interruption is possible. In addition, this study suggests that the greatest benefit from resumption practice can be found from practicing in a manner that is most likely to create an associative link or form a new procedure. Because associative links form at the level of specific goals, it is important to determine the goals of a task.

Training to mitigate the disruptive effects of interruptions in complex environments such as medicine can be profoundly important for improving performance. However, this work suggests that general interruption training is unlikely to improve performance. Instead, a training protocol should consider which primary task is most likely to be interrupted by a particular secondary task. In the case of errors, consideration should be given to identifying the step that is most critical if skipped or repeated.

References

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

Altmann, E. M., & Trafton, J. G. (2004). *Task interruption: Resumption lag and the role of cues*. DTIC Document.

Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin & Review*, *14*(6), 1079–1084.

Altmann, E. M., & Trafton, J. G. (2015). Brief lags in interrupted sequential performance: Evaluating a model and model evaluation method. *International Journal of Human-Computer Studies*, *79*, 51–65.

Anderson, J. R. (2007). How can the human mind occur in the physical universe? Oxford University Press.

Ash, J. S., Berg, M., & Coiera, E. (2004). Some unintended consequences of information technology in health care: The nature of patient care information system-related errors. *Journal of the American Medical Informatics Association*, *11*(2), 104–112.

B. Edwards, M., & Gronlund, S. D. (1998). Task interruption and its effects on memory. *Memory*, *6*(6), 665–687.

Bailey, B. P., & Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior*, *22*(4), 685–708.

Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science*, *21*(1), 31–61.

Cades, D. M., Boehm-Davis, D. A., Trafton, J. G., & Monk, C. A. (2011). Mitigating disruptive effects of interruptions through training: What needs to be practiced? *Journal of Experimental Psychology: Applied*, *17*(2), 97.

Cades, D. M., Trafton, J. G., & Boehm-Davis, D. A. (2006). Mitigating disruptions: Can resuming an interrupted task be trained? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*, 368–371.

Chisholm, C. D., Dornfeld, A. M., Nelson, D. R., & Cordell, W. H. (2001). Work interrupted: A comparison of workplace interruptions in emergency departments and primary care offices. *Annals of Emergency Medicine*, *38*(2), 146–151.

Cutrell, E., Czerwinski, M., & Horvitz, E. (2001). *Notification, disruption, and memory: Effects of messaging interruptions on memory and performance*.

Eyrolle, H., & Cellier, J.-M. (2000). The effects of interruptions in work activity: Field and laboratory results. *Applied Ergonomics*, *31*(5), 537–543.

González, V. M., & Mark, G. (2004). Constant, constant, multi-tasking craziness: Managing multiple working spheres. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 113–120.

Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied*, *6*(4), 322.

Hettinger, A. Z., & Fairbanks, R. J. T. (2012). Recognition of Patient Selection Errors in a Simulated Computerized Provider Order Entry System. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*, 1743–1747. http://pro.sagepub.com/content/56/1/1743.short

Hodgetts, H. M., & Jones, D. M. (2006). Contextual cues aid recovery from interruption: The role of associative activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1120.

Kim, G. R., Chen, A. R., Arceci, R. J., Mitchell, S. H., Kokoszka, K. M., Daniel, D., & Lehmann, C. U. (2006). Error reduction in pediatric chemotherapy: Computerized order entry and failure modes and effects analysis. *Archives of Pediatrics & Adolescent Medicine*, *160*(5), 495–498.

Koppel, R., Metlay, J. P., Cohen, A., Abaluck, B., Localio, A. R., Kimmel, S. E., & Strom, B. L. (2005). Role of computerized physician order entry systems in facilitating medication errors. *Jama*, *293*(10), 1197–1203.

Monk, C. A. (2004). The effect of frequent versus infrequent interruptions on primary task resumption. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *48*, 295–299.

Ratwani, R. M., McCurry, J. M., & Trafton, J. G. (2008). Predicting postcompletion errors using eye movements. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 539–542.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Harvard University Press.

Trafton, J. G., Altmann, E. M., & Brock, D. P. (2005). Huh, what was I doing? How people use environmental cues after an interruption. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*, 468–472.

Trafton, J. G., Altmann, E. M., Brock, D. P., & 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*(5), 583–603.