INTERRUPTION PRACTICE REDUCES ERRORS

Kevin Zish[1] J. Gregory Trafton[2]

George Mason University[1]
Fairfax, VA

Naval Research Laboratory[2]
Washington, DC

Mitigating the effects of interruptions is important for tackling the increasing number of possible disruptions at home, at work, and online. Previous work has shown that the benefits of practice can decrease the amount of time it takes to resume a task after an interruption. This paper demonstrates that the same benefit can be extended to error rates at the post-completion step showing that a general increase in interruptions leads to a decrease in error rates for the last step of a form-filling task.

INTRODUCTION

Interruptions are disruptive because they distract from the goals of a task (González & Mark, 2004). The disruption can manifest itself in a slower time to completion, a reduction in accuracy, or delayed time to resume the task (B. Edwards & Gronlund, 1998; Cutrell, Czerwinski, & Horvitz, 2001; Eyrolle & Cellier, 2000; Hodgetts & Jones, 2006; Monk, 2004; Ratwani, McCurry, & Trafton, 2008; Trafton, Altmann, Brock, & Mintz, 2003).

Even short interruptions lasting less than three seconds can lead to loss of information and errors (Altmann, Trafton, & Hambrick, 2013). One place that interruptions can be especially costly is in the medical domain: medical professionals are interrupted on average 9.7 times per hour (Chisholm, Dornfeld, Nelson, & Cordell, 2001).

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). MFG is an activation-based model where memory elements with higher activations should be retrieved faster with a lower likelihood of errors than memory elements with lower activations (there is, of course, noise in the system that can obscure this relationship).

MFG makes three fundamental predictions about resuming a task after an interruption. First, goals decay (details of the equations that enter into this decay can be found in Altmann & Trafton, 2002; Trafton et al., 2003). MFG suggests that there are two primary ways of reducing or slowing down decay: rehearsal and using environmental cues (the two other components to the theory). Finally, there are (learned) associative links between steps in a task (Altmann & Trafton, 2007).

One obvious way to reduce the disruptiveness of interruptions is to provide practice on the primary task. In fact, the long term working memory theory predicts that practice with the primary task should greatly reduce the disruptiveness of interruptions (Ericsson & Kintsch, 1995; Oulasvirta & Saariluoma, 2004, 2006). However, evidence from Trafton et al. (2003) shows, somewhat surprisingly, that there is almost no benefit from practicing the primary task alone. Understanding how to reduce the disruptiveness of interruptions is thus a primary focus of this work.

MFG suggests another method of reducing the disruptiveness of interruptions: additional practice with interruptions on a specific task. As long as the interrupting task and the resuming task are the same, additional interruption practice should facilitate resumption.

As an example, previous work has shown that practicing a task with interruptions can reduce the amount of time to resume by several hundred milliseconds (Cades, Boehm-Davis, Trafton, & Monk, 2011). Time to resume after an interruption is more formally known as resumption lag and is a measure that has helped researchers define the cost of an interruption and discuss ways to mitigate its effect (Monk, 2004; Trafton, Altmann, & Brock, 2005; Trafton et al., 2003; Altmann & Trafton, 2004).

To investigate the specificity of the practice effect for task/interruption pairs, Cades et al. (2011) used a clever paradigm where they used two different interrupting tasks but only a single primary task. Some participants saw the same interrupting task throughout the experiment, while other participants switched interrupting tasks. They found that participants got faster and better at resuming only when they received practice with the interruption; when the interruption was switched, there was no benefit to practice. Cades et al. (2011) explained this effect by suggesting that the interrupting task primed the primary task. Resumption suffered because the priming from the interruption task is reduced when switched.

Interestingly, Cades et al. (2011) only focused on resumption lag effects. As important as time effects are (Gray & Boehm-Davis, 2000), errors typically have a larger impact in real-world settings. However, because environmental cues provide so much priming and are so important in applied tasks, it could be that errors only increase when activation is especially low and are not likely to occur often. It is an open question whether the priming from the environment will mitigate the effects of associative priming.

Research on procedural tasks has shown that some steps are particularly error prone. The post-completion step (PCS) is the last goal of a task and has a high error rate when interrupted compared to other steps (Bailey & Konstan, 2006; Ratwani et al., 2008).

In the medical field there is a plethora of evidence demonstrating the dangers of errors at the PCS. Electronic health record systems are used to reduce certain errors related to poor-handwriting and dosage miscalculations (Koppel et al., 2005). There are cases where the user (medical staff, MD, Nurse, etc.) forgets to complete the PCS which is to log off or close the record of an active patient file. This has resulted in serious errors where medications for patients have been

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2014**	2. REPORT TYPE	3. DATES COVERED **00-00-2014 to 00-00-2014**
4. TITLE AND SUBTITLE **Interruption Practice Reduces Errors**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Research Laboratory ,Washington,DC,20375**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES **Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 58, 1 vol. pp. 265-269, 2014.**		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a REPORT **unclassified**	b ABSTRACT **unclassified**	c THIS PAGE **unclassified**	**Same as Report (SAR)**	**5**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

switched resulting in severe harm (Ash, Berg, & Coiera, 2004; Hettinger & Fairbanks, 2012; Kim et al., 2006).

The current study explores the practice effect with interruptions at the PCS. MFG makes a straightforward prediction: the more practice a user receives on an interrupting task/resuming task pair, the higher the activation should be. This higher activation should result in faster resumption lags and fewer errors regardless of the number of iterations the practice is presented.

METHOD

Participants

Fifty-seven George Mason University psychology undergraduates participated for course credit.

Tasks

Primary Task. The primary task consisted of a simulated system to help medical practitioners fill out patient order forms. To complete the computerized physician order entry (CPOE) form, participants had to fill in all widgets and bubbles using information provided on the screen (Figure 1). The orders were presented three at a time to the right of the screen with a status of Urgent, Priority, or Normal. Participants were asked to fill out the patient order that had the most urgent status.

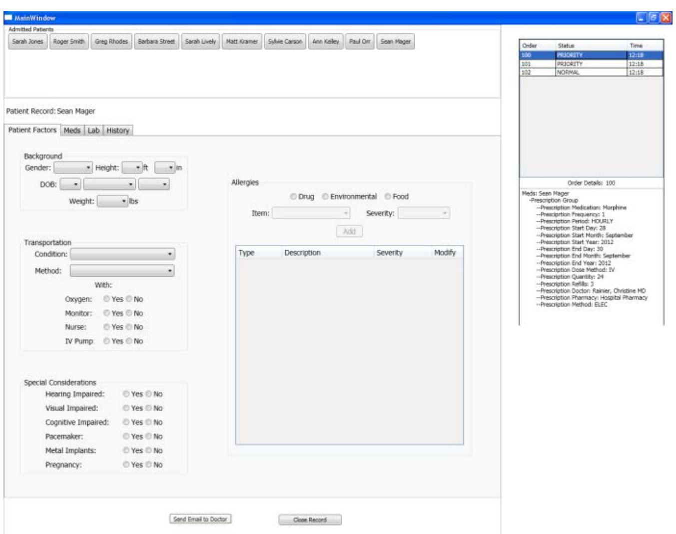

Figure 1. Screenshot of the CPOE system.

In starting the task, the participant chose a patient order based on its status. Clicking on an order opened a set of details below which provided all of the information that would be needed to complete the form. The participant then chose the patient mentioned in the details from a list on the top of the screen. Each patient had four tabs associated with their name. These tabs included Patient Factors, Labs, Medication, and History. Only the first three tabs were used in this experiment. The correct tab was listed in the Order Details and needed to be selected to continue. Afterwards the participant would fill out all of the information in the CPOE system by clicking on the widgets.

Once the order was complete and all of the information filled in, participants were instructed to click the “Send Email to Doctor” button that would send an email to the doctor. This button was to be clicked only one time after the order was fully finished (otherwise the doctor would get multiple emails and be unsure which order was correct). After this button was pressed, a pop-up box appeared confirming that the email had been sent. Participants accepted this step by clicking “OK” on the screen.

The final step of a trial was to click “Close Record” in the CPOE task which brought participants to the next order. Because of the high error rate associated with the PCS this step was analyzed for this paper.

Secondary Task. At times participants would receive a series of math problems to solve. The screen associated with this interrupting task occluded all information from the primary task. Two-digit subtraction problems appeared with four answers in the center of the screen. Participants chose one of four answers to move onto the next problem.

Design

Participants were assigned to one of three conditions in a 2 (within: interruption/non-interruption) x 3 (between: number of interruptions) design. Each condition had six, twelve, or eighteen interruptions throughout the task. The interrupted widgets were equally divided into PCS and non-PCS interruptions such that participants randomly had math problems at the PCS three, six, or nine times. The other interruptions were randomly displayed at other steps and trials throughout the task. This was to reduce the chance that participants would consciously prepare for an interruption at the PCS. Because only the post-completion step was analyzed, interrupted trials are considered to be those where the participant was interrupted just prior to the PCS.

Of the 24 patient orders that were completed, participants had both interruption and non-interruption trials for a within-participants design. Filling in all of the information and clicking “Close Record” was recorded as completing a trial. The placement of which trials were interrupted was randomized. Math problem interruptions persisted for a predetermined 15s; participants were instructed to answer as quickly and accurately as possible.

Procedure

All participants filled out two copies of an approved IRB consent form. Biographical information was taken before each participant was instructed on the CPOE task. Participants were seated approximately 47 cm from the computer monitor. The experimenter explained the task using screenshots of the CPOE system and example math problems. Three practice trials were used prior to the main experiment. Practice trials included examples of interruptions. Participants were asked to complete the orders as quickly and accurately as possible. The experiment was completed without the experimenter being present in the room. Once finished, all participants were debriefed and dismissed.

Measures

Behavioral data based on mouse clicks was collected for all participants in addition to screen recordings. Only the behavioral data was analyzed in this study. An error at the post-completion step was defined as clicking any other widget of the CPOE interface that was not the “Close Record” button.

A percentage of errors at the post-completion step were calculated between interrupted and non-interrupted control trials. This percentage represented the ratio of errors at the post-completion step over the number of total possible errors.

RESULTS

Error Rates

Fifty-seven participants made a total of 106 errors at the post-completion step. The proportion of errors were analyzed using a repeated-measures ANOVA model using the mean number of errors for interrupted and non- interrupted trials with the number of interruptions at the PCS as a covariate.

Participants had a significantly longer resumption lag if they were interrupted (M = 4,827.54) at the PCS than control trials (M = 1,326.43), F(1, 55) = 218.865, MSE = 34,834,7915, $p < .05$, $\eta^2 = 3.98$.

Somewhat surprisingly, resumption lag was statistically flat across interruption trials, F(1, 55) = 1.65, MSE = 4,147,218, $p > .05$ (Figure 2). Additionally there was a marginal effect between the PCS being interrupted and the number of interruptions throughout the task, F(1, 55) = 3.43, MSE = 5,480,615, p =.07.

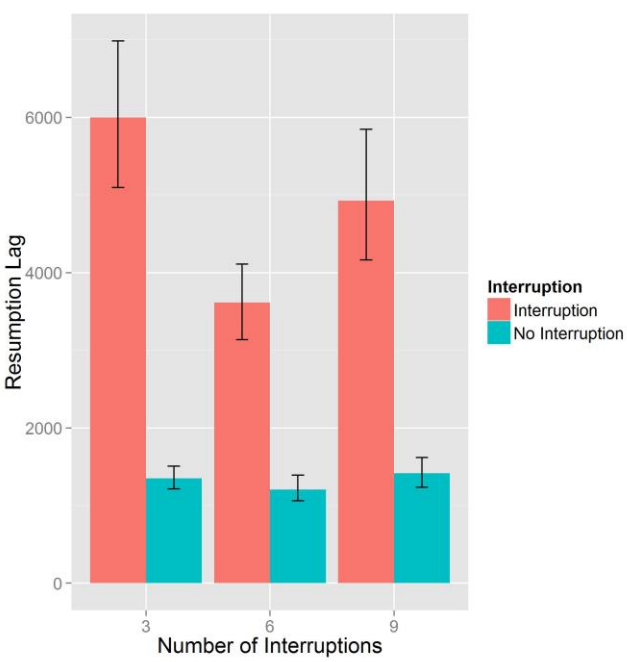

Figure 2. The average resumption lag (milliseconds) by number of interruptions at the PCS. Error bars are 95% confidence intervals.

Participants made significantly more errors at the post-completion step during interruption trials (M = 27.1%) than control trials (M = .01%), F(1, 55) = 84.31, MSE= 2.61, $p < .05$, $\eta^2 = 1.53$. The very low error-rate for control trials suggests that the task was well-learned.

As suggested by Figure 3, error rates at the PCS decreased as the number of interruptions increased F(1, 55) = 10.16, MSE =.31, $p< .05$, $\eta^2 = .18$ A significant interaction between the number of interruptions and interrupted trials suggests that trials without interruptions did not show the same decrease in error rates F(1, 55) = 10.87, MSE = .34, $p< .05$, $\eta^2 = .20$.

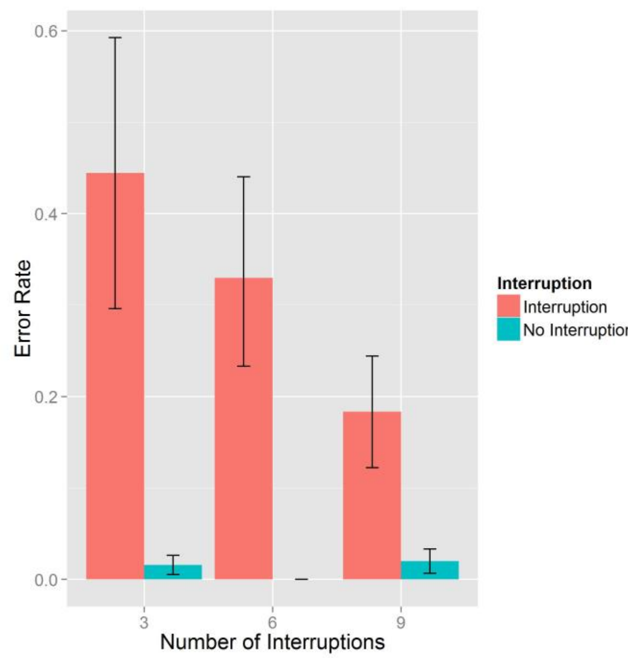

Figure 3. The average proportion of errors by number of interruptions at the PCS. Error bars are 95% confidence intervals.

DISCUSSION

The results of this study show two measures of the effect of interruptions at the PCS. These measures include errors and resumption lag. Participants practiced with increasing number of interruptions across three groups in a CPOE task. Interruptions lasted for 15s and made it difficult for the participant to rehearse the most recent step of the primary task. A low error rate for trials that were not interrupted suggests that the task was well learned and errors were not a result of misunderstanding the task.

Consistent with MFG we found that more practice with interruptions reduces errors. However resumption lag did not follow the expected pattern which raises a theoretical question. MFG (and most other memory theories) predicts that error and resumption lag measures are based on the same theoretical construct (activation) and therefore should show similar patterns. There are several possible reasons why the measures differed.

Theoretical

The first idea for why errors and resumption lag did not show the same decrease is that activation as understood in MFG is not the only contributor to whether a goal is retrieved. It is possible that activation accounts for a large, but not total, amount of the variance when calculating whether a memory has a higher signal than interference. If something in addition to activation contributes to goal retrieval then the difference between the pattern for errors and resumption lag is a result of how practice affects their potentially distinct sources.

Another possibility may be the existence of two processes of goal selection using activation. The current model of MFG suggests that higher activation leads to preferential and faster selection for a goal (Altmann & Trafton, 2002, 2007). Thus selection of a goal can be achieved by situations where there are more memory representations of one action on the task than another.

Our study provided an equal number of interruptions both at the PCS and non-PCS widgets in the CPOE task. An equivalent number of memory traces for resuming at both types of widgets may have created interference when trying to retrieve the correct goal. This interference could create a selection bias for a process related to accuracy over a second process for speed. The result of this bias could be error rates and resumption lag measures that do not operate with the same directionality.

A third possibility is that resumption lag as a measure is noisy and may need larger numbers of participants to develop a distinct effect. Therefore it is possible that the relatively flat nature of resumption lag at the PCS is due to noise. Considering that measures of error are as noisy as resumption lag and yet were still significantly different, it is unlikely that our flat result is due to lack of power. Of course, it is difficult to interpret a null effect, so we do this cautiously.

Applied

In 2000 a report by the institute of medicine showed that medical errors account for 90,000 preventable deaths a year (Kohn, Corrigan, & Donaldson, 2000). Training protocols are used in hospitals to respond to some of these errors because they are less costly than system redesigns and are easier to implement.

In combination with Cades et al. (2011) our results strongly suggest that current training regimens need to consider the effects of interruptions on performance. Even for well-trained tasks, the introduction of interruptions increases error rates (Cades et al., 2011). Interruptions can even be momentary to decrease performance (Altmann et al., 2013). Our results suggest that training protocols are most effective when they utilize a large number of interruptions. Because practice is specific to the task-interruption pair, it is most effective to train with the interruption most likely to occur or the most likely to reduce performance (Cades et al., 2011).

As an example, infusion pumps have been used to allow consistent dose of a medication over a period of time. New technology in the form of “Smart” infusion pumps have medication libraries and controls that are marketed to aid medical professionals in administering medication. A study of these pumps found that they reduced errors that typically were caught by other means before reaching the patient (Rothschild et al., 2005). The rate of events that did reach the patient did not change with pump use. Considering that factors such as an incorrect decimal place can result in death, this study would suggest that training with a large number of the interruptions most likely to occur while programming the pump would be most effective at reducing errors.

This study also shows that there are noticeable differences in performance for a single session of training in addition to the multiple sessions used by others (Cades et al., 2011). We thus suggest that current practitioners incorporate a sizable number of interruptions into their training regime.

ACKNOWLEDGEMENTS

This work was supported by the Office of Naval Research.

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. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA480333

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., & Hambrick, D. Z. (2013). Momentary Interruptions Can Derail the Train of Thought. Retrieved from http://psycnet.apa.org/psycinfo/2013-00033-001

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. doi:10.1016/j.chb 2005 12.009

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.

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. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.26.418

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211.

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. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp.

113–120). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=985707

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. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 56, pp. 1743–1747). SAGE Publications. Retrieved from 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.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (2000). *To err is human building a safer health system* (Vol. 627). National Academies Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=JInZiZnUyicC&oi=fnd&pg=PA1&dq=to+err+is+human+institute+of+medicine&ots=hpKPPDbGR6&sig=7lEcQexi3gK6AHxHXgkwzsp7dJQ

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 The Journal of the American Medical Association*, *293*(10), 1197–1203. doi:10 1001/jama.293.10.1197

Monk, C. A. (2004). The effect of frequent versus infrequent interruptions on primary task resumption. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 48, pp. 295–299). SAGE Publications. Retrieved from http://pro.sagepub.com/content/48/3/295.short

Oulasvirta, A., & Saariluoma, P. (2004). Long-term working memory and interrupting messages in human–computer interaction. *Behaviour & Information Technology*, *23*(1), 53–64.

Oulasvirta, A., & Saariluoma, P. (2006). Surviving task interruptions: Investigating the implications of long-term working memory theory. *International Journal of Human-Computer Studies*, *64*(10), 941–961.

Ratwani, R. M., McCurry, J. M., & Trafton, J. G. (2008). Predicting postcompletion errors using eye movements. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 539–542). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1357141

Rothschild, J. M., Keohane, C. A., Cook, E. F., Orav, E. J., Burdick, E., Thompson, S., … Bates, D. W. (2005). A controlled trial of smart infusion pumps to improve medication safety in critically ill patients*. *Critical Care Medicine*, *33*(3), 533–540.

Trafton, J. G., Altmann, E. M., & Brock, D. P. (2005). Huh, what was I doing? How people use environmental cues after an interruption. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 49, pp. 468–472). SAGE Publications. Retrieved from http://pro.sagepub.com/content/49/3/468.short

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.