The Effects of Task Interruption on Human Performance: A Study of the Systematic Classification of Human Behavior and Interruption Frequency

Byung Cheol Lee[1] and Vincent G. Duffy[1,2]

1 School of Industrial Engineering, Purdue University, West Lafayette, Indiana, USA
2 School of Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana, USA

Abstract

Interruptions are prevalent phenomena in modern working environments; yet, few interruption studies have been conducted on different types of human tasks. A study using computer-based human behavior tasks was carried out to investigate the effects of interruptions with 4 different primary-interrupting task sets. The tasks used in this study were determined by cognitive and motor skill processes based on human behavior classification theory. The results showed that interruption effects were most negative in cognitive/cognitive task sets, and skill/cognitive task sets showed the least amount of effects in task completion time, $F(3, 332) = 77.88$, $p < .001$. In error rate performance, skill/skill task sets showed the highest rate, and error rates for cognitive primary task sets (cognitive/cognitive and cognitive/skill task sets) were not significantly different, $F(1, 42) = 41.18$, $p < .001$; $F(1, 42) = 3.56$, $p = 0.0661$. Interruption frequency also negatively affected task performance, $F(1, 412) = 89.88$, $p < .001$, but skill tasks' quantitative performance did not show significant effects at different interruption frequencies, $F(1, 187) = 3.78$, $p = .0534$. The results showed that interruptions increase more time to complete in cognitive tasks and produce more errors in skill tasks. Also, similar types of primary-interrupting tasks were more susceptible from interruptions. Thus, based on task composition of work process, we can estimate different effects from the interruptions and memory load, and task similarity in primary-interrupting task relationship were considered a main factor.

Keywords: Interruption; Human task performance; Task classification; Interruption frequency; Memory load

1. INTRODUCTION

Interruptive environments are widespread in modern workplaces, and the negative effects from interruptions and multitasking are taken as more serious. The information systems, such as e-mail, instant messaging, and web assistants, provoke that the amount of information a human receives outweighs the amount of information a human can handle, and human performance can easily be overwhelmed by overloaded information from such systems (Cutrell, Czerwinski, & Horvitz, 2000). Also, such systems are increasingly competing for workers' attention and adversely affect task performance and emotional states. However, most workers argue that multitasking becomes a common and essential working strategy for dealing with numerous, simultaneous information inputs, but interruptive and multitasking environments in the workplace are unavoidable (Freedman, 1997). One study asserted that workers spend less than 3 min on any single task before switching to another one (Gonzalez & Mark, 2004).

Correspondence to: Byung Cheol Lee, School of Industrial Engineering, Purdue University, 315 N. Grant Street, West Lafayette, IN 47907. Email: lee102@purdue.edu

Received: 21 February 2012; revised 6 November 2012; accepted 21 November 2012

View this article online at wileyonlinelibrary.com/journal/hfm

DOI: 10.1002/hfm.20603

A growing body of studies have been discussed about the negative effects of interruptions. Generally, interruptions increase the task completion time, worsen decision making, and lead to more errors, frustration, annoyance, and anxiety (Carayon et al., 2007; Cutrell et al., 2000; Gillie & Broadbent, 1989). Freedman (1997) also claimed that the average time lost in U.S. companies is 2.1 hr of employee productivity per day as a result of work interruptions. Such detrimental effects of interruptions and task switching are proved not only on the task requiring motor skill processes such as activating, connecting, and pressings but also on the task requiring cognitive processes, such as analyzing, calculating, and estimating (Meister, 1985). Hembrooke and Gay (2003) supported negative effects that multitasking can interfere with the memory performance on lecture content in the classroom. This study has potential implications for healthcare and builds on the past research of researchers within our group (Seals & Duffy, 2005).

Although many studies have examined interruption effects in multitasking and task switching environment, most of them have been conducted without considering different types of task (Brixey, Walji, Zhang, Johnson, & Turley, 2004; Carvalho, Vidal, & de Carvalho, 2007; Czerwinski, Horvitz, & Wilhite, 2004). Multitasking is possible, without excessive cognitive efforts, based on types of task, and the effects of interruptions can vary by different tasks as well. For example, people can walk and talk at the same time without any trouble, but they cannot remember the taste of the chocolate they had yesterday while doing simple one-digit multiplication. We can drive a car while listening to music but cannot talk on two phone calls at the same time. In addition, even if multitasking may be possible, task performance may not be the same as the case without multitasking. Several studies mentioned the importance of task classification in interruption research, but few have been actually carried out (Burmistrov & Leonova, 1996; Keirn & Aunon, 1990).

This study mainly focused on different interruption effects on human task performance with systematic task classification. To investigate the interruption effects by different types of tasks, this study used mental arithmetic problems as cognitive process tasks and simple word-processing works as motor skill tasks. Since a cognitive process task requires more mental demands to complete than a motor skill task, it is likely that the former is more susceptible to interruptions than the latter. Other than task types, interruption frequency was also chosen as another factor to impact on task performance.

This study may help to build a strategy to minimize detrimental effects from interruptions. While prior research proposed the opportune timing of interruption or suggested different approaches of interruption coordinating (Adamczyk & Bailey, 2004), this study examined another aspect of multitasking working environment, type of task. Practically, considering that most tasks in work environments consist of a mixture between simple motor skill tasks and complicated cognitive process tasks, the investigation of interruptions impacts on each task performance can provide the foundation to build a viable solution to minimize detrimental effects. The next sections will discuss the tasks used in previous literatures and theoretical backgrounds for task classification. Then, the experiment will be introduced and the results will be discussed.

1.1 The Tasks Used in Previous Interruption Research

The tasks used in the experiment on interruption did not receive a lot of attention. In the first interruption research, traced back to the 1920s, a list of three-letter anagrams was given to subjects to solve (Weybrew, 1984). While solving the anagrams, the subjects were abruptly asked to estimate the amount of time it took them to solve the first 10 anagrams (an interrupting task). The study found that the tasks that were interrupted were recalled more often than those that were not interrupted, and that the subjects recalled the interrupted tasks first. It is called the "Zeigarnik effect" (Weybrew, 1984); however, her study did not consider task types and performance, and the overall effects of interruptions were not investigated as well.

The lack of consideration of task types has been continued until recently. Kreifeldt and McCarthy (1981) conducted an interruption experiment to compare the different interface designs of the reverse Polish notation (RPN) and algebraic notation (AN) calculators without a discussion on different task types. The study focused more on relative evaluation on different logic designs in interruption environments. Field (1987) used a database of traversal tasks as primary tasks and completing a sequence of numbers or looking up book titles as interrupting tasks. The study indicated a significant effect on users' behavior after interruption activities, but the study was not specific about which problems were interrupted by which tasks. Gillie and Broadbent

(1989) adopted computer-based game tasks and investigated the similarity between interruption and main task. They demonstrated that the interruption that is similar to the main task is disruptive, but they did not consider different types of task.

In recent research, Eyrolle and Cellier (2000) employed more realistic working environment and three different types of task: creation, regularization, and modification task. The study investigated the sensitivity of temporal constraints on task performance and error rates. However, the tasks were chosen without considering cognitive demands, which can be a distinctive feature of modern working tasks. Speier, Vessey, and Valacich (2003) also investigated the effects of interruptions on decision-making performance with college-level coursework in different information-presenting modes. They claimed that interruptions facilitate performance of simple tasks but impede performance of complex tasks.

Monk, Boehm-Davis, Mason, and Trafton (2004), using programming a VCR with a simulated interface as a primary task and tracking moving targets on a computer screen as an interrupting task, suggested the importance of interruption timing on task resumption and insisted that the middle of the task is the most critical moment for resuming interrupted tasks. Other studies applied various types of tasks such as text editing and phone calls, visual search tasks, call center tasks, complex resource allocation tasks, and more (Burmistrov & Leonova, 1996). Even though the tasks used in the study were well designed for examining various attributes of interruptions, they lacked a systematic approach to differentiating human tasks or behaviors. Such distinction of tasks can be one of the important variables in evaluating interruption effects on task performance.

1.2. The Theoretical Task Classifications and Rasmussen's Skill, Rule, and Knowledge-Based Classification

Many different types of task classifications have been proposed. The origin of task classification can be traced back to Taylor's work on method analysis (Taylor, 1911), and Miller (1953) proposed traditional human factors task analysis. In human–computer interaction domain, goals, operators, methods, selection rules (GOMS) analysis was also suggested as a task analysis or classification tool (Card, Moran, & Newell, 1983). The skill, rule, and knowledge (SRK)-based behavior classification was developed by Rasmussen (1983) and widely cited in cognitive engineering area. It was applied to develop a taxonomy of human performance models, stages of skill acquisition, theories of expertise effects in memory recall, and a framework for interface design for complex sociotechnical systems (Harvey & Koubek, 2000; Tian et al., 2014; Vicente, 1999).

TABLE 1. Definitions of human behavior classification and task examples

Skill-based Human behavior
Physical activities without cognitive efforts
e.g. copying prescription, taking phone calls
Rule-based Human behavior
Selecting and applying the formulas
e.g. expanding abbreviations
Knowledge-based Human behavior
Solving the questions with inferring
e.g. mathematical calculation for prescription

Rasmussen's SRK-based behavior classification is based on the traditional top-down approaches, which apply a set of predefined decision rules and theoretical models for human behavior (El-Gamal & Grether, 1995). It provides a useful framework to distinguish human behavior or task by the types of information processing demands, and it distinguishes categories of human behavior by different states of the constraints in working environments (Rasmussen, 1983; Reason, 1990). Brief definitions of each level of human behavior and task examples are shown in Table 1.

Rasmussen's framework gave the insights to design the structure of task types used in this study. Since skill-based tasks need motor skill sets or dexterity to complete and rule- or knowledge-based tasks require high demand of mental resources, Rasmussen's framework can turn into physical motor tasks and cognitive demand requiring task. In addition, rule- and knowledge-based tasks share many similar features. Both are applied in the problem-solving stages of human cognitive processing, and they are difficult to detect because they mostly operate internally. Both require conscious control of action with critical choice of cognitive demands. With theses common points, rule- and knowledge-based tasks were merged into one class of task in this study, which can be differed from skill-based tasks. For the sake of convenience, skill-based tasks are called motor skill tasks or skill tasks and rule-/knowledge-based tasks are called cognitive process tasks or cognitive tasks in this study.

1.3. Different Types of Task Interferences

Although some studies confirmed that multitasking can be achieved without detriment to task performance and interruption can even improve the performance of certain tasks, the negative effects of interruptions are generally accepted. Since interruptions happen in an environment in which tasks must be processed serially, the understanding of the serial task process is critical in interruption research; moreover, in the serial task processing environment, besides interruptions, there are other types of task interferences. To differentiate the effects of interruptions, other interferences need to be specifically addressed.

Broadly, task interferences in serial processing can be divided into four different types: interruption, distraction, task switching, and task interleaving. An interruption can be mainly defined as "an externally generated, randomly occurring, discrete event that breaks continuity of cognitive focus on a primary task" (Coraggio, 1990, p. 19). However, while a distraction happens when two different sensory channels receive two different types of information at the same time, an interruption happens when two different types of information are provided through a single sensory channel (Speier et al., 2003). Distractions and interruptions are similar in that they happen when the decision makers carry on a primary task, but they are different in the number of sensory channels involved. For example, some noise might be heard from outside while you are reading a book. This can be called a distraction. When you are reading a book in a room and you turn your attention to someone that enters the room, it can be defined as an interruption; thus, an ongoing task can continue in a situation with a distraction condition, but an ongoing task needs to be stopped in an interruption condition.

Further, a switching task occurs in the shift between an ongoing task and an interrupting task brought on by an interruption or a distraction, and task interleaving is a repeated form of task switching. They concern the circumstance of returning to one task after having dealt with another. In sum, interruptions and distractions are necessary components for task switching and task interleaving, and they are considered extended forms of single interruptions or distractions. Task switching and task interleaving are ultimately influenced by four factors: urgency, importance, duration, and switching or interruption cost (Freed, 1998).

Among the above types of work interferences, only interruptions were dealt with in the present study because interruption is a basic element of task switching and task interleaving and because the negative effects of interruption are more salient and easier to measure than other forms of interferences.

2. METHOD

2.1. Research Framework

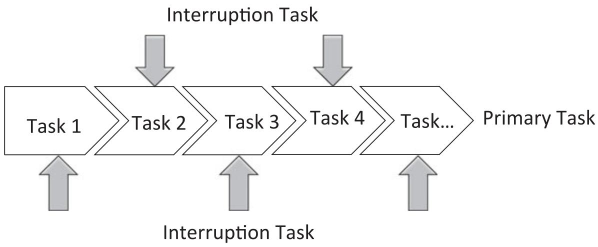
Controlled laboratory experiments were conducted to investigate the effects of interruptions on task performance by different types of tasks. In this study, the effects of interruption were measured in task completion time and the frequency of errors committed by participants. The participants performed a series of tasks, which consists of predetermined order of two task types: cognitive process tasks (cognitive tasks) and motor skill tasks (skill tasks). Mental arithmetic tasks, which are word problems in seventh-grade mathematics, were selected as cognitive process tasks and simple word processing tasks were chosen as motor skill tasks. The frequency of interruptions was also considered as another variable for investigating interruption effects. Figure 1 explains the experiment framework for this study.

2.2. Subjects

Thirty-nine subjects participated in the study. All subjects were college students who were taking a junior- or senior-level industrial engineering courses. The average age of participants was 22 years, and most were in their junior year (40%) or their senior year (56%). Thirteen were women, and 26 were men. All participants were familiar with typing sentences using a keyboard and a computer and had no problem solving basic mathematic questions. All participants received a maximum credit of 5% toward their final course grade. All participants were eligible for incentives if they completed in the experiment. The incentives were approved as part of the institutional review board (IRB) protocol reviewed prior to testing to ensure protection of human subjects in research.

2.3. Experiment Design

A 4×3 full factorial design of the experiment with four sets of task combination, which is described in Figure 1, and three levels of interruption frequency (none, once



Task% set%	Primary% Task%	Interrup1on% Task%
Cog-Skill	Cognitive task*	Skill task**
Cog-Cog	Cognitive task	Cognitive task
Skill-Cog	Skill task	Cognitive task
Skill-Skill	Skill task	Skill task

**: Cognitive task = Cognitive process task: Mental arithmetic task (Word problems in 7th grade mathematics)*

***:Skill task = Motor skill task: word processing (typing given sentences)*

Figure 1 Experiment framework.

per task, three per task) was used to investigate the interruption effects. Different task sets were applied to between subjects while within-subject design was used in different interruption frequencies. The experimental procedures and independent and dependent variables' operationalized levels had been tested and validated in pilot studies. Interruption frequency was also limited to three times per task because task performance had not changed significantly at more than three times per task and subjects showed unintended annoyance, which can possibly affect the task performance.

2.4. Independent Variables

2.4.1. Types of Tasks

This study used mental arithmetic problems and word processing tasks as different types of tasks. The cognitive process and motor process in task processing are the main criteria for selection, and tasks should be easy to measure their task completion time and errors as well; in addition, tasks needed to be suitable for embedding in computer applications. Considering these criteria, we decided word problems in seventh-grade mathematics as cognitive tasks and simple word processing as skill tasks.

Many psychology and psychophysiology studies used mental arithmetic problems and word processing tasks to measure reaction time and human performance (Barrett & Krueger, 1994; John & Newell, 1989; Keele, 1968). Word processing task is a major form of human-to-computer communication and is a basic task involving perceptual-motor processing. Comparing with a cognitive task such as stimulus-response (S-R) compatibility, word processing tasks require longer duration and a flow of behavior. While cognitive tasks should be performed in a sequential way—perceive the stimulus, do cognitive process, execute the response—word processing task is parallel in nature: look ahead at what is coming while executing the motor response for the current letters (John & Newell, 1989). Also, it is an essential and very common work for their academic activities in current college education environments. The number of words in the word processing task in this study was decided by the average typing speed of clerical workers, which, in the experiment, ranged between 42 and 48 words per minute (Ostrach, 1997).

The mental arithmetic tasks were employed in many studies to measure cognitive demands (Campbell & Clark, 1992; McCloskey, Caramazza, & Basili, 1985). Stein and Smith (1998) suggested low-level and high-level cognitive demands with mathematic question solving tasks. Low-level demand tasks include memorization and procedures without connections, and high level comprising procedures with connection and doing mathematics. In this study, they were used to evaluate different magnitudes of interruption effects from word processing tasks. Kreifeldt and McCarthy (1981) showed that writing down simple multiplication tables was negatively affected on the ability to return to the main task. Thus, mathematic questions are enough to test mental demands and prove disruptive effects from interruptions.

To maintain uniform difficulty of questions and to minimize the effects from different math competency by each subject is the main obstacle of mental arithmetic problems in a between-subjects laboratory experiment. To do so, three following criteria were applied: First, only word problems with similar question lengths were used. Reading the questions averages out different math skill competency in terms of task completion time and adds some amount of mental demands, such as understanding the questions and finding the appropriate rules for solving to all subjects. Second, only questions requiring simple mathematic operations were selected from a seventh grade-level math test book (Linderman, 1999). Third, calculators or computers were not allowed in the experiment to encourage cognitive efforts in task process. Table 2 shows some sample questions for cognitive and skill tasks.

TABLE 2. Sample questions for cognitive and physical task

Cognitive task: Math question solving[a]
Your school cafeteria makes its delicious tuna salad by adding 2 pounds of mayonnaise to every 3 pounds of canned tuna. Canned tuna costs \$1.50 per pound and mayonnaise costs \$0.75 per pound. How many pounds of tuna salad can the cooks prepare for \$100?
a. 88 1/3 b. 33 1/3 c. 55 d. 30
Joe found a battery-powered drill for 25% off the original price. At the checkout counter the clerk enters the sale price, adds 5% sales tax, and then tells Joe he owes \$189. What was the original price of the drill?
a. \$158.78 b. \$198.45 c. \$240 d. \$226.80
Physical task: Sentence copying[b]
Type the following sentences in given space.
Another approach to definitions of the quality of healthcare is directly connected with patient safety, which can evidently be confirmed by reviewing some definitions of patient safety. One study described healthcare quality as the eligibility of the recommended medical treatments and Berwick expounded.
Type the following sentences in given space.
Additionally, the performance measuring of healthcare providers can be divided into two dimensions. Technical performance that based on knowledge and judgment to diagnosis and care delivery and interpersonal performance between care providers and patients. Thus, measuring the performance and quality of healthcare requires additional effort.

[a]:All questions were selected from grade 7th math text book (Linderman, 1999). Only word problems were selected.
[b]:Sentences were arbitrary composed but the number of words was limited between 40 and 45. This number is based on average typing words per minute for clerical workers (Ostrach, 1997).

2.4.2. Interruption Frequency

Interruption frequency is defined as the number of tasks interrupting primary task performance. In the experiment, interruption frequency was set to zero, one, or three. In the pilot experiment, more than three interruptions per task was also tested, but too many interruptions in a task resulted in a severe decrease in task performance due to frustration and lack of motivation, not due to the effects of interruptions. Generally, many studies maintained that a high number of interruptions generate more detrimental effects on task performance (Bailey, Konstan, & Carlis, 2000; Czerwinski, Cutrell, & Horvitz, 2000); however, the effects of different interruption frequencies in cognitive tasks need not be the same as those in skill tasks. In fact, an increased number of interruptions in cognitive tasks more severely exhausts mental resources and produces more adverse effects on task performance (Wickens & McCarley, 2007).

2.4.3. Other Independent Variables

Individual variance in participants could have been considered another independent variable in cognitive tasks. However, since the participants were limited to college students who were enrolled in junior- or senior-level industrial engineering courses, we can assume that experience and skill proficiency were not that dissimilar. Further, because word processing tasks require simple mechanical motor skill and are very familiar with college students, we could assume that participants have similar capability to perform skill tasks.

The other possible independent variable is the timing of interruptions. The effects of interruptions could vary significantly based on when interrupting tasks occur (Li et al., 2006). In this experiment, the timing of interruption was set to 15 s after primary tasks began. Pilot tests for setting interruption timing suggested that 15 s indicated most participants were engaged in reading the problem sentences in cognitive tasks. Therefore, this study fixed the timing of interruption to minimize unnecessary variation in the results.

2.5. Dependent Variables

To identify and measure detrimental effects of interruptions, the completion time of the primary tasks was measured. In addition, the wrong answer rate for cognitive tasks and number of typographical errors in skill tasks were assessed. In this study, typographical errors were written as typos and measured for only motor skill tasks. The task completion time and wrong answers and typos can be considered as quantitative and qualitative measurements for task performance.

2.5.1. Task Completion Time

As quantitative measurements, task completion time was measured in seconds by total time to complete a primary task minus time taken by interrupting tasks. Compared to the task completion time of the noninterruption task, the task completion time of the task with interruption increased by two transition time intervals: interruption lag and resumption lag (Altmann & Trafton, 2002). The former is the wrap-up time for the primary task before engaging in the interruption task and is called "switching time" to interrupting tasks (Wickens & McCarley, 2007). The latter provides the magnitude of the disruptive effect from interruptions and is called "return time" to primary tasks. These two types of transition times are described in Figure 2.

As shown in the Figure 2, interruption lag and resumption lag are carryover effects from primary tasks and interrupting tasks, and they depend on the type of task and the amount of mental resources needed to perform the task. The elongated time due to the two lags not only increases total task completion time but also decays the memory capacity for retaining information relevant to the interrupted task.

2.5.2. Task Performance

Based on the nature of the tasks, task performance had to be measured differently for each task type; thus, the number of right answers for mental arithmetic problems (cognitive tasks) and the number of typos for word processing tasks (skill tasks) were the two main measurements for task performance. To minimize the chance of getting right answers accidently, the choice of "I don't know" was also given to participants. Typos in word processing tasks included not only spelling errors but also capitalization and punctuation errors. Compared to none-interruption cases, it was assumed that increased cognitive workload due to interruption would lead to more wrong answers in mental arithmetic questions and to more typos in word processing tasks.

2.6. Procedure

Each subject was asked to fill out a demographic questionnaire and consent form before the experiment, and a 10-minute training session with several sample tasks was provided.

The subjects then performed three different task sets with combinations of skill and cognitive tasks. Each subject experienced all three different interruption frequencies (none, once, and three times) randomly. A none-interruption scenario consisted of pairs of cognitive and skill tasks, and the results were used as performance reference for once and three times interruption scenarios. Once and three times interruption scenarios were assigned in a predetermined order. The order of task sets was fully counterbalanced.

When the experiment started, a primary task was placed on the computer screen. Then, interruption tasks were given, popping up on a new screen. The timing of interruptions was set to 15 s after primary tasks began. Once the subject answered the interruption question, he or she clicked an "OK" button and automatically returned to the interrupted primary task. The experiment included three scenarios per subject, and each scenario continued for 10 min.

3. RESULTS

3.1. Interruption Effects on Quantitative Task Performance: Task Completion Time

Our main research question was how two different types of interruption tasks affect the performance of two different types of primary tasks; cognitive (mental arithmetic questions) and skill (word processing) tasks were used as both primary and interruption tasks. We performed repeated measures analyses on task completion time as a quantitative task performance. To evaluate the performance of the two different task types on the same basis, time performance ratio (TPR), which is the standardized task completion time, was introduced, defined as the ratio between task completion time with interruptions and without interruptions. The formula

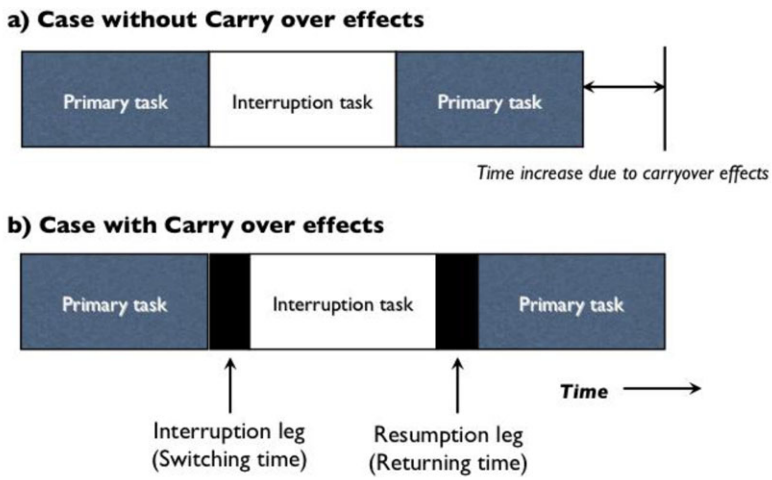


Figure 2 Carryover effects in interruption task environment and interruption and resumption legs.

for TPR is shown as follows:

$$Time\ Performance\ Ratio\,(TPR) = \frac{Task\ completion\ time\ with\ Interruptions}{Task\ completion\ time\ without\ Interruption} \quad [1]$$

The TPR was calculated for each subject. If the TPR value was 110, it meant that 10% more time was required to finish a task with an interruption. These standardized values made possible the direct comparison between cognitive and skill task performances. Table 3 shows the descriptive statistics of TPR by different task sets and interruption frequencies.

A two-way analysis of variance (ANOVA) with repeated measures showed a significant effect of task sets on TPR, $F(3, 331) = 42.45$, $p < .001$. It also showed a significant effect of interruption frequency on TPR, $F(1, 331) = 85.87$, $p < .001$. No interaction effect was observed.

Paired comparisons showed that the TPR of cognitive/cognitive task sets (mean = 177.73, standard deviation [*SD*] = 17.29) was significantly higher than those of any other task sets. Cognitive/skill task sets had the next highest TPR (mean = 162.85, *SD* = 14.26). Regarding interruption frequency effect, three time interruptions per task (mean = 148.12, *SD* = 33.63) negatively impacted more than one time interruption (mean = 140.16, *SD* = 27.16).

3.2. Interruption Effects on Qualitative Task Performance: Wrong Answer Rate and Typo Rate

As we described in the previous section, wrong answer rates in mental arithmetic questions and typo rates in word processing tasks were chosen as qualitative measurements for task performance. To compare two different task types, we have to use the ratio of two measurements: wrong answer rate ratio (WARR) and the typo rate ratio (TRR), which are the standardized values for two measurements. WARR is applied to cognitive primary tasks: cognitive/Cognitive task sets and cognitive/Skill task sets, and TRR is used on skill primary tasks: skill/Cognitive task sets and skill/skill sets. WARR and TRR can be defined by following formulas:

$$Wrong\ Answer\ Rate\ Ratio\,(WARR) = \frac{Wrong\ Answer\ Rate\ with\ Interruptions}{Wrong\ Answer\ Rate\ without\ Interruption} \quad [2]$$

$$Typo\ Rate\ Ratio\,(TRR) = \frac{Typo\ Rate\ with\ Interruptions}{Typo\ Rate\ without\ Interruption} \quad [3]$$

Table 4 shows descriptive statistics for each subject's WARR and TRR. A two-way ANOVA with repeated measures showed a significant effect of task sets on WARR and TRR, $F(3, 74) = 25.44$, $p < .001$. It also showed a significant effect of interruption frequency

TABLE 3. Descriptive statistics of Time Performance Ratio (TPR) in different task sets and frequencies

Time Performance Ratio (TPR) [a](Primary /Interrupting task)	Mean	Median	Standard Deviation
Cognitive/Cognitive Task Set			
One time interruption per task	166.34	169	12.35
Three times interruptions per task	193.30	194	8.79
Cognitive/Physical Task Set			
One time interruption per task	156.43	156	12.16
Three times interruptions per task	172.00	173	11.95
Physical/Cognitive Task Set			
One time interruption per task	110.43	111	8.58
Three times interruptions per task	115.43	115	7.75
Physical/Physical Task Set			
One time interruption per task	136.27	138	8.78
Three times interruptions per task	141.61	141	7.95

[a] $Time\ Performance\ Ratio\,(TPR) = \dfrac{Task\ Compltion\ Time\ with\ Interruptions}{Task\ Compltion\ Time\ without\ Interruption}$

TABLE 4. Descriptive statistics of typo rate ratio and wrong answer rate ratio in different task sets and frequencies

Wrong Answer Rate Ratio (TPR) [a,c] Typo Rate Ratio (TRR) [b,c] (Primary /Interrupting task)	Mean	Median	Standard Deviation
Cognitive/Cognitive Task Set			
One time interruption per task	133.10	131.5	5.86
Three times interruptions per task	144.11	143	3.37
Cognitive/Physical Task Set			
One time interruption per task	126.30	127	5.79
Three times interruptions per task	142.70	143	5.17
Physical/Cognitive Task Set			
One time interruption per task	145.9	148	6.90
Three times interruptions per task	174.11	174	4.01
Physical/Physical Task Set			
One time interruption per task	163.80	162.5	7.45
Three times interruptions per task	194.60	195	6.36

[a] $Wrong\ Answer\ Rate\ Ratio\,(WARR) = \dfrac{Wrong\ Answer\ Rate\ with\ Interruptions}{Wrong\ Answer\ Rate\ without\ Interruption}$

[b] $Typo\ Rate\ Ratio\,(TRR) = \dfrac{Typo\ Rate\ with\ Interruptions}{Typo\ Rate\ without\ Interruption}$

[c] WARR's were measured on cognitive primary tasks (cognitive/cognitive task sets and cognitive/physical sets) and TRR's were measured on physical primary tasks (physical/cognitive task sets and physical/physical sets)

on WARR and TRR, $F(1, 74) = 17.85$, $p < .001$. No interaction effect was observed.

Paired comparisons showed that the TRR of skill/skill task sets (mean = 179.20, $SD = 17.18$) was significantly higher than that of skill/cognitive task sets (mean = 159.26, $SD = 15.51$), and the WARR of cognitive/cognitive task sets (mean = 138.32, $SD = 7.36$) was not significantly higher than that of cognitive/skill task sets (mean = 134.50, $SD = 9.97$). In addition, three time interruptions per task (mean = 157.12, $SD = 23.63$) were more negatively impacted than one time interruption (mean = 138.79, $SD = 17.16$).

4. DISCUSSION

4.1. Interruption Effects on Quantitative Performance

Interruptions were more detrimental to cognitive tasks than skill tasks. Figure 3 shows that TPR's of

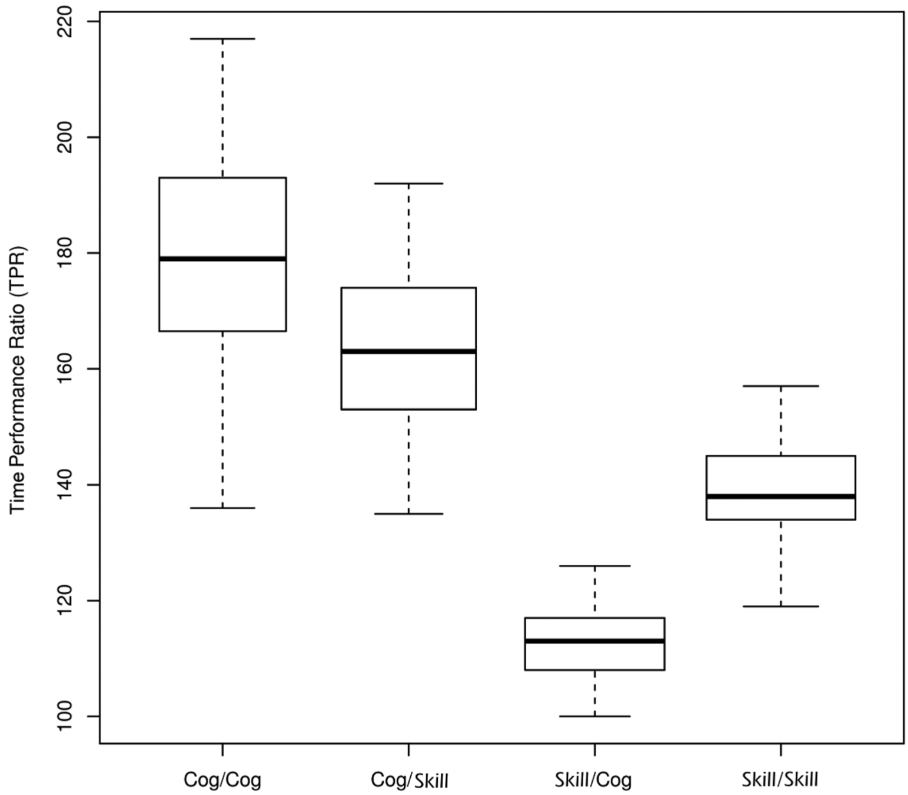


Figure 3 Quantitative task performances by four different primary/interrupting task sets.

cognitive/cognitive and cognitive/skill task sets were higher than those of skill/cognitive and skill/skill task sets, $F(1, 334) = 80.36, p < .001$. It means that the interruption lags and resumption lags in cognitive tasks are longer than those of skill tasks. Comparing task completion times of four different task sets, as also shown in Figure 3, cognitive/cognitive task sets were most influenced from interruptions and skill/cognitive task sets are most resilient, $F(3, 332) = 77.88$, $p < .001$. Since cognitive tasks require more mental resources for task completion, more switching time to interrupting tasks and returning time to interrupted tasks were required. The long switching time and returning time made it difficult to retrieve the cues for the interrupted goal and delayed the transition to the new task (Altmann & Trafton, 2002).

Single channel theory, which is based on the paradigm of the psychological refractory period (PRP), explains well the long switching and returning time given two simultaneous inputs on a single sensory or mental resource (Ferreira & Pashler, 2002). The theory also asserts that the prolonged reaction time or waiting time for second stimuli that are examples of PRP originates in the waiting time for first stimuli processing (Ferreira & Pashler, 2002; Wickens & McCarley, 2007). In this study, there were two PRPs situated in the experiment: before and after the interrupting task. To switch to the interrupting task, PRP was required, and to return to the interrupted primary task, another PRP was needed. These PRPs were much longer in cognitive tasks than in skill tasks because more information stimuli had to be processed. This could explain why task completion time in cognitive/cognitive task sets was longer than in other task sets.

The effects of interruptions on skill tasks showed different results. Interestingly, according to the data, the typing speeds in skill/skill task sets were lower than those in skill/cognitive task sets. The feasible cause could be task similarity. Some research has already confirmed that similar tasks generate more inadvertent effects on task performance than dissimilar tasks (Eyrolle & Cellier, 2000; Gillie & Broadbent, 1989). This also could explain the high interruption effects on the cognitive/cognitive task sets.

Accordingly, the main factors of interruption effects on quantitative performance were memory load and task similarity. Cognitive tasks such as mental arithmetic problems require more memory load to complete than simple skill tasks that rely on sensory information for finishing. According to task completion time data, the role of memory load was more critical than task similarity in performance degradation, and cognitive tasks were more vulnerable to interruptions. In the skill task environment, task similarity was a more important factor in task performance.

4.2. Interruption Effects on Qualitative Performance

Interruption effects on qualitative performance showed different results than on quantitative performance. In the skill task environment, skill/skill task sets showed a higher TRR than skill/cognitive task sets, $F(1,42) = 41.18$, $p < .001$, but the WARR in cognitive/cognitive task sets and in cognitive/skill task sets was not statistically different, $F(1,42) = 3.56$, $p = .0661$. Figure 4 shows the WARR for cognitive/cognitive task sets and cognitive/skill task sets as well as the TRR for skill/cognitive task sets and skill/skill task sets.

Similar to the quantitative performance results, task similarity was regarded as a main cause for the high interruption effects on skill/skill task sets. In the skill task environment, even though tasks did not require a high level of mental resources or effort, interruptions brought disconnections and disruptions to well-trained consistent motor behaviors, and it made it very difficult to retrieve cues to activate the interrupted goal (Altmann & Trafton, 2002; Wickens & McCarley, 2007); moreover, as mentioned earlier, the same type of primary/interruption task environment brings more confusion and deteriorates the connection between cue and goal (Altmann & Trafton, 2002). We suggest that this is the principal factor in the high typo rates for skill/skill task sets, which consist of word processing tasks in both primary and interrupting tasks.

Surprisingly, the number of wrong answers due to the interruptions was not significantly increased in mental arithmetic tasks; TRRs, qualitative measures for skill tasks, were much higher than WARRs, qualitative measures for cognitive tasks, $F(1, 85) = 75.44$, $p < .001$, and variations in TRRs for different task sets were not significant, $F(1, 42) = 3.31$, $p = .076$.

Many previous studies asserted that task interruptions lead to errors, mistakes, or work failures (Boehm-Davis & Remington, 2009; Burmistrov & Leonova, 1996); however, the Swiss cheese model (Reason, 2000) describes work failure causation as a series of events that must take place in a specific order and manner for a failure to happen. Therefore, we suggest that interruptions in the mental arithmetic questions increased the cognitive workload and required more mental resources, but the level of workload and mental resources was not enough to induce a noticeable or statistically significant difference in the number of errors or mistakes induced for the mental arithmetic questions.

According to our experiment results, interruptions negatively affected the performance on cognitive tasks, but slight changes in mental disruption could not be measured by qualitative performance, such as a wrong answer rates for the mental arithmetic questions.

4.3. Interruption Frequency Effects on Task Performance

Generally, the effects of interruption frequency were significant, and higher interruption frequency had a greater negative effect on performance for both quantitative and qualitative measurement, $F(1, 412) = 89.88$, $p < .001$. More specifically, in qualitative performance measurements, the interruption frequency impacted the cognitive task performance more than it impacted the skill task performance, $F(1, 334) = 8.77$, $p < .01$, but differences in skill task performance were not statistically significant at different interruption frequencies, $F(1, 187) = 3.78$, $p = .0534$. In the case of skill task performance at different interruption frequencies, task similarity did not affect the performance. As mentioned earlier, skill tasks do not depend much on mental resources, and relatively low workloads are required; therefore, the transition to interrupting tasks and the transition from interrupting tasks were very efficient in the skill task environment in the present study.

In applying single channel theory, Reynolds (1966) found that the PRP is lengthened if tasks involved a more complex choice rather than a simple response. The word processing tasks used in the present experiment could be considered simple response tasks. The subjects in the experiment, junior- or senior-level college students, performed the tasks with much fewer cognitive demands and automatically responded as tasks initiated; thus, the PRP in the word processing tasks was very short, and task switching between primary and interrupting tasks was efficiently executed.

As shown in Figure 5, there were significant differences in qualitative task performance between the low- and high-interruption frequency conditions. Interestingly, the subjects in the high-frequency interruption condition had significantly higher typo rates

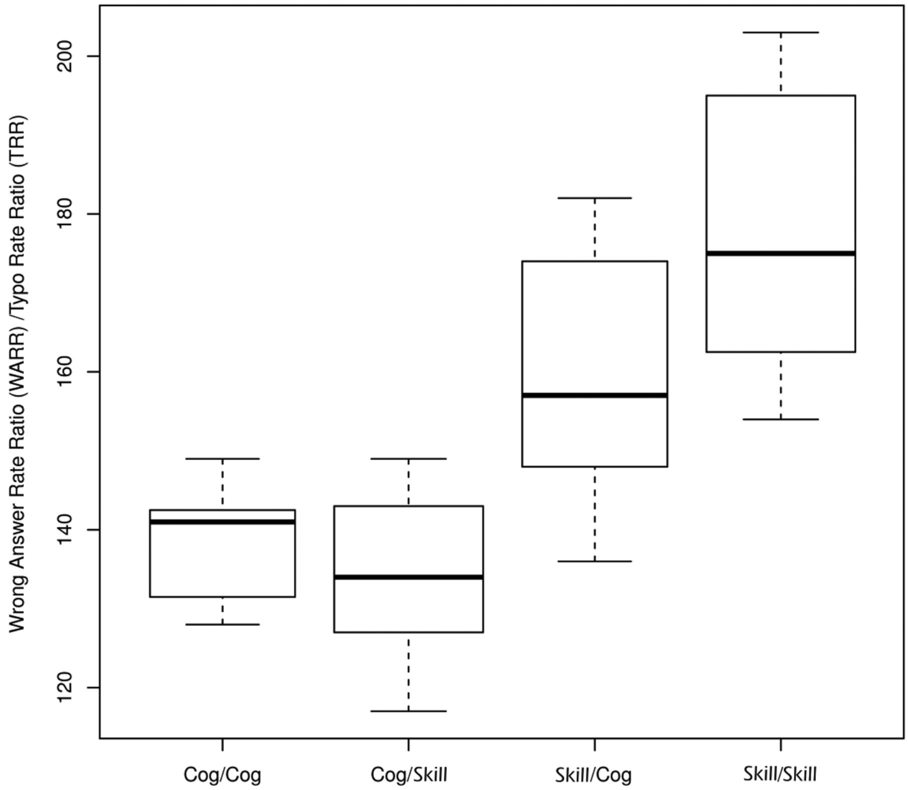


Figure 4 Qualitative task performances by four different primary/interrupting task sets.

and longer task completion times than the subjects in the low-frequency conditions. These results clearly resonate with precedent research on detrimental interruption effects; in that research, subjects experiencing high-frequency interruptions paid much less attention to the task and performed worse on the task than those experiencing low-frequency interruptions (Speier et al., 2003).

4.4. Combined Results from Task Types and Interruption Frequency

Figure 6 shows the combined results of interruption effects on task performance by different types of tasks and different frequency. It indicates that interruptions elongate task completion time in cognitive tasks and produce more errors in skill tasks. Comparing to those of cognitive primary task sets such as cognitive/cognitive task sets and cognitive/skill task sets, task performances of skill primary task sets are scattered, which means skill primary task sets are more sensitive to interruption effects. Also, task similarity is more important than task types or memory load because both cognitive/cognitive task sets and skill/skill task sets demonstrated lowest task performances.

5. LIMITATION

Potential concerns about this study are the increased control given by a laboratory experiment and the representativeness of the subject. Taking the study's constraints into consideration, the value of the findings from any study must be evaluated. In this study, the controlled experiment was inevitable to achieve the goal of study objective and must be balanced against the fundamental limitations of generalizability. In addition, the subject recruiting strategy used in this study was intended, in that participants were specifically targeted for uniform competency and familiarity of experiment tasks. While the wide range of the task proficiency and the lack of motivation of student subjects can be other concerns, we mitigated them by ensuring that the tasks were designed with sufficiently similar levels of mental demand and by providing performance incentives.

Moreover, we acknowledge that the nature of the tasks indicated in Section 1.2 may not be

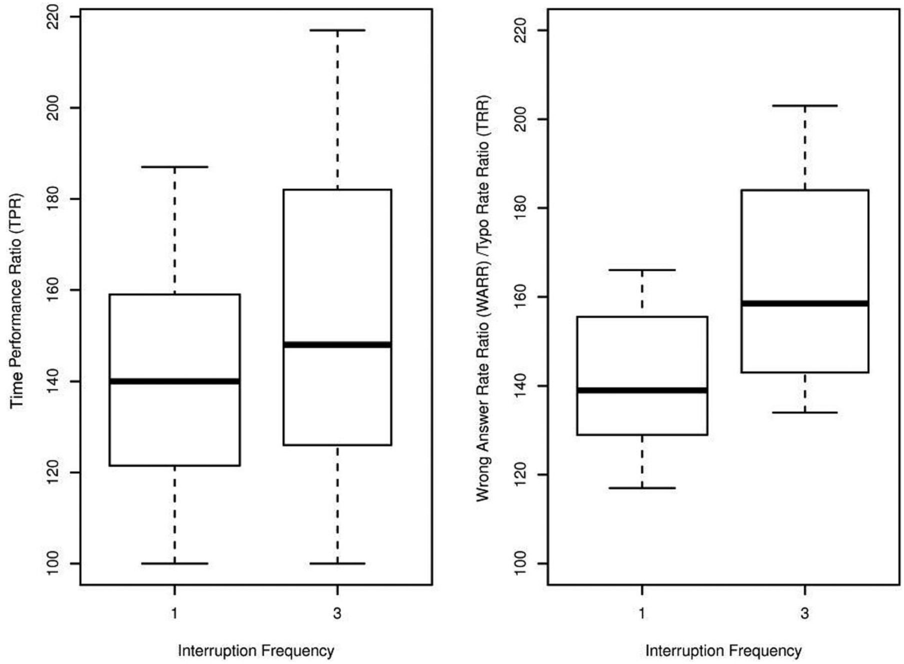


Figure 5 Quantitative task performance (TPR) and qualitative task performance (WARR, TRR) by interruption frequency.

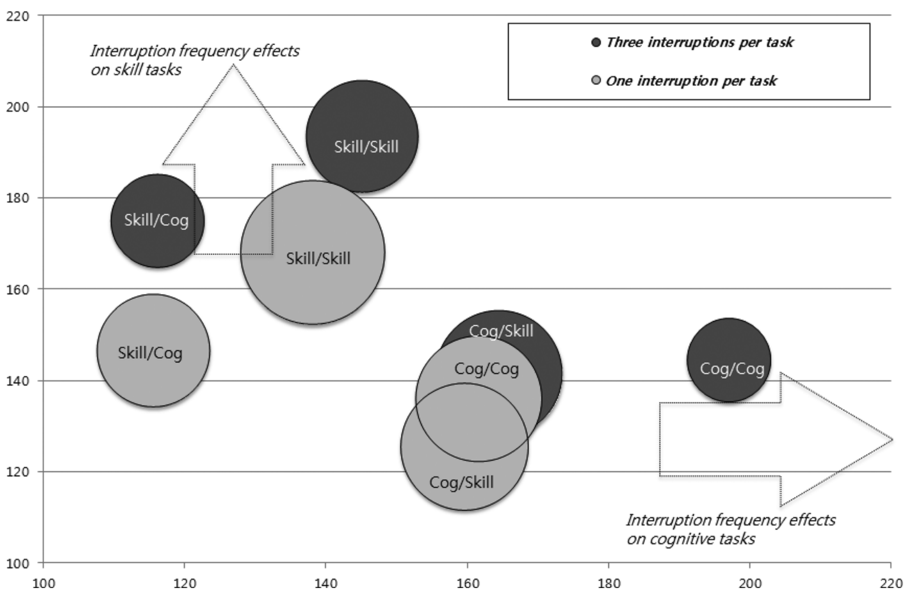


Figure 6 Interruption effects on task performance by task types and interruption frequency.

comprehensive for cognitive demand requiring and physical motor skill tasks, but they were chosen to clarify the distinction between cognitive process and motor process and their performance can be easily and reliably measured (Meister, 1985). Furthermore, we are unable to provide the optimal timing and frequency of interruption per task. Even though we fully considered the timing of interruption to be in the middle of the task, this timing can be varied by each subject's task proficiency and skill. The data showed that any task was not finished in 15 s, which is the first onset timing of interruption. So, we can reasonably assume that interruption occurred at appropriate timing to measure the effect on task performance.

Finally, the interruptions employed in this study were lacking social characteristics in task environments, and thus the results of the study should be considered before generalizing into other work environments.

6. FUTURE RESEARCH

First, we could expand the experiments with different types of interruption modes. In this study, only a single type of interruption mode, immediate interruption type, was reviewed. Furthermore, the timing of interruptions was fixed. Other aspects of interruptions, such as different interruption timings and time constraint environments, could also be considered for further study (McFarlane, 2002; McFarlane & Latorella, 2002).

Second, the independent variables that were fixed in this experiment could be tested. A different level of education could provide a different level of task experience and skill. In addition, if we were able to carefully classify cognitive tasks into rule based and knowledge based, we could garner more detailed results about interruption effects on specific human cognitive tasks.

Finally, using the theoretical results from this study, practical application is possible. Currently, there are many work environments prone to interruptions, such as aviation control, driving, and healthcare. In many cases, interruptions are considered a part of work, not a source of work interference. Using the theoretical results from this study, systematic analysis of interruptions happening in the workplace is possible and can prevent detrimental effects from interruptions.

7. CONCLUSION

The study investigated the effects of interruptions on different task types and interruption frequency. Quantitative results showed that interruptions had greater effects on cognitive primary task performances than on skill primary task performances. Task similarity also played a notable role in more negative interruption effects in cognitive/cognitive task sets and skill/skill task sets than in cognitive/skill task sets and skill/cognitive task sets.

In qualitative performance measurements, interruptions brought more errors to skill tasks than to cognitive tasks. Skill task performance easily deteriorated with divided attention from an interruption, but cognitive task performance was quite resilient to such temporal attention deprivation. Instead, some level of cumulated, increased cognitive workload could be considered the main cause for errors or mistakes in cognitive tasks. In addition, task similarity clearly indicated higher typo rates in skill/skill task sets than in skill/cognitive task sets.

The effects of interruption frequency suggested that higher frequency of interruption brings more unfavorable effects on both qualitative and quantitative performance; however, qualitative performance in cognitive tasks did not result in significant difference.

References

Adamczyk, P. D., & Bailey, B. P. (2004). If not now, when? The effects of interruption at different moments within task execution. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 271–278). New York: ACM Press.

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

Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2000). Measuring the effects of interruptions on task performance in the user interface. In Systems, Man, and Cybernetics, 2000 IEEE International Conference on (Vol. 2, pp. 757–762). Piscataway, NJ: IEEE.

Barrett, J., & Krueger, H. (1994). Performance effects of reduced proprioceptive feedback on touch typists and casual users in a typing task. Behaviour & Information Technology, 13(6), 373–381.

Boehm-Davis, D. A., & Remington, R. (2009). Reducing the disruptive effects of interruption: A cognitive framework for analyzing the costs and benefits of intervention strategies. Accident Analysis and Prevention, 41, 1124–1129.

Brixey, J., Walji, M., Zhang, J., Johnson, T., & Turley, J. (2004). Proposing a taxonomy and model of interruption. Enterprise Networking and Computing in Healthcare Industry, 2004. Proceedings of 6th International Workshop on Enterprise Networking and Computing in Healthcare Industry (Healthcom 2004) (pp. 184–188), IEEE.

Burmistrov, I., & Leonova, A. (1996). Interruptions in the computer aided office work: Implications to user interface design. Human-Computer Interaction: Human Aspects of Business Computing. Proceedings of EWHCI, 96.

Campbell, J. I. D., & Clark, J. M. (1992). Cognitive number processing: An encoding-complex perspective. Advances in Psychology, 91, 457–491.

Carayon, P., Wetterneck, T. B., Hundt, A. S., Ozkaynak, M., DeSilvey, J., Ludwig, B., . . . Rough, S. S. (2007).

Evaluation of nurse interaction with bar code medication administration technology in the work environment. Journal of Patient Safety, 3(1), 34.
Card, S. K., Moran, T. P., & Newell, A. (Eds). (1983). The psychology of human-computer interaction. Hillsdale, NJ: Lawrence Erlbaum.
Carvalho, P. V., Vidal, M. C., & de Carvalho, E. F. (2007). Nuclear power plant communications in normative and actual practice: A field study of control room operators' communications. Human Factors and Ergonomics in Manufacturing & Service Industries, 17(1), 43–78.
Coraggio, L. (1990). Deleterious effects of intermittent interruptions on the task performance of knowledge workers: A Laboratory investigation (Unpublished doctoral dissertation). University of Arizona, Tucson.
Cutrell, E. B., Czerwinski, M., & Horvitz, E. (2000). Effects of instant messaging interruptions on computing tasks. In Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2000 Extended Abstracts), (pp.99–100). New York: ACM.
Czerwinski, M., Cutrell, E., & Horvitz, E. (2000). Instant messaging and interruption: Influence of task type on performance. In Proceedings of OZCHI (Vol. 2000, pp. 356–361).
Czerwinski, M., Horvitz, E., & Wilhite, S. (2004). A diary study of task switching and interruptions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 175–182). New York: ACM.
El-Gamal, M. A., & Grether, D. M. (1995). Are people Bayesian? Uncovering behavioral strategies. Journal of the American Statistical Association, 90(432), 1137–1145.
Eyrolle, H., & Cellier, J.-M. (2000). The effects of interruptions in work activity: Field and laboratory Results. Applied Ergonomics, 31, 537–543.
Ferreira, V. S., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(6), 1187.
Field, G. E. (1987). Experimentus interruptus. ACM SIGCHI Bulletin, 19(2), 42–46.
Freed, M. (1998). Managing multiple tasks in complex, dynamic environments. Proceedings of the National Conference on Artificial Intelligence (pp. 921–927). Menlo Park, CA: AAAI Press.
Freedman, D. (1997). What's next: The cost of competence. Inc.com. Retrieved January 11, 2014, from http://www.inc.com/magazine/20070901/whats-next-the-cost-of-competence.html
Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. Psychological Research, 50(4), 243–250.
Gonzalez, V. M., & Mark, G. (2004). "Constant, constant, multi-tasking craziness": Managing multiple working spheres. Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data (pp. 668–668).
Harvey, C. M., & Koubek, R. J. (2000). Cognitive, social, and environmental attributes of distributed engineering collaboration: A review and proposed model of collaboration. Human Factors and Ergonomics in Manufacturing, 10(4), 369–393.
Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. Journal of Computing in Higher Education, 15(1), 46–64.
John, B. E., & Newell, A. (1989). Cumulating the science of HCI: From SR compatibility to transcription typing. ACM SIGCHI Bulletin, 20(SI), 109–114.
Keele, S. W. (1968). Movement control in skilled motor performance. Psychological Bulletin, 70(6, Pt. 1), 387–403.
Keirn, Z. A., & Aunon, J. I. (1990). A new mode of communication between man and his surroundings. IEEE Transactions on Biomedical Engineering, 37(12), 1209–1214.
Kreifeldt, J., & McCarthy, M. (1981). Interruption as a test of the user-computer interface. In JPL Proceedings of the 17th Annual Conference on Manual Control (pp. 655–667).
Li, S. Y. W., Cox, A. L., Blandford, A., Cairns, P., Young, R. M., & Abeles, A. (2006). Further investigations into post-completion error: The effects of interruption position and duration. In Proceedings of the 28th Annual Meeting of the Cognitive Science Society. Mahwah, NJ: Lawrence Erlbaum.
Linderman, B. (1999). Word problems homework booklet, grades 7–8. Columbus, OH: Instructional Fair.
McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. Brain and Cognition, 4(2), 171–196.
McFarlane, D. (2002). Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. Human-Computer Interaction, 17(1), 63–139.
McFarlane, D. C., & Latorella, K. A. (2002). The scope and importance of human interruption in human-computer interaction design. Human-Computer Interaction, 17(1), 1–61.
Meister, D. (1985). Behavioral analysis and measurement methods. New York: Wiley.
Miller, R. B. (1953). Handbook on training and training equipment design. Wright-Patterson Air Force Base, OH: Wright Air Development Center, Air Research and Development Command, United States Air Force.
Monk, C. A., Boehm-Davis, D. A., Mason, G., & Trafton, J. G. (2004). Recovering from interruptions: Implications

for driver distraction research. Human Factors, 46(4), 650–663.

Ostrach, T. R. (1997). Typing speed: How fast is average: 4,000 typing scores statistically analyzed and interpreted. Orlando, FL: Five Star Staffing.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man and Cybernetics, 13(3), 257–266.

Reason, J. (1990). Human error (1st ed.). Cambridge: Cambridge University Press.

Reason, J. (2000). Human error: Models and management. BMJ, 320(7237), 768–770.

Reynolds, D. (1966). Time and event uncertainty in unisensory reaction time. Journal of Experimental Psychology, 71(2), 286–293.

Seals, A. B., & Duffy, V. G. (2005). Toward development of a computer-based methodology for evaluating and reducing medication administration errors. Ergonomics, 48(9), 1151–1168.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. Decision Sciences, 34(4), 771–797.

Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. Mathematics Teaching in the Middle School, 3(4), 268–275.

Taylor, F. W. (1911). Shop management. New York: Harper & Brothers.

Tian, R., Lee, B. C., Yucel, G., Abel, S. R., Hultgren, K. E., & Duffy, V. G. (2012). Evaluating bar coding-aided medication administration through identification of nursing work deficiencies. Human Factors and Ergonomics in Manufacturing & Service Industries, 24(4), 468–478.

Vicente, K. J. (1999). Cognitive work analysis: Toward safe, productive, and healthy computer-based work. Mahwah, NJ: Lawrence Erlbaum.

Weybrew, B. B. (1984). The Zeigarnik phenomenon revisited: Implications for enhancement of morale. Perceptual and Motor Skills, 58, 223–226.

Wickens, C. D., & McCarley, J. S. (2007). Applied attention theory (1st ed.). Boca Raton, FL: CRC Press.