Applied Ergonomics 59 (2017) 94–103

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

The effects of interruption similarity and complexity on performance in a simulated visual-manual assembly operation

Carl Pankok Jr., Maryam Zahabi, Wenjuan Zhang, Inchul Choi, Yi-Fan Liao, Chang S. Nam, David Kaber*

North Carolina State University, Raleigh, NC, USA

ARTICLE INFO

Article history:
Received 2 February 2016
Received in revised form
18 August 2016
Accepted 26 August 2016

Keywords:
Task interruptions
Manual assembly
Worker productivity

ABSTRACT

The objective of the study was to assess the effects of interruption task similarity and complexity on performance of a simulated industrial assembly operation. Eighteen participants performed a simulated industrial assembly operation, including one trial with no interruption and eight others presenting an interruption task. Interruption conditions comprised a full crossing of task similarity to the primary assembly operation (similar, dissimilar) and complexity (simple, complex) with replication for each participant. Order of condition presentation was randomized. Findings revealed greater time to return to primary visual-manual assembly performance after a similar task interruption. Results also indicated complex interruptions may promote cognitive arousal that increases productivity following assembly interruptions. The majority of results are explained in terms of the Activation-Based Memory for Goals model. Findings provide some guidance for interruption management protocol design for workers engaged in procedural visual-manual assembly operations.

1. Introduction

With increasing technology in the workplace, workers are often faced with multitasking situations and greater levels of distraction that can have negative side effects on primary task performance, such as decreased productivity or increased errors. For example, in an observational study in which operators captured and updated telephone-line data on a computer screen, Eyrolle and Cellier (2000) reported significant increases in processing time as the number of task interruptions increased. Westbrook et al. (2010) found that nurses who were interrupted while administering medications exhibited a 12–13% increase in error rate and error severity with more interruptions. In an analysis of daily work logs of 21 employees, Murray and Khan (2014) found that office workers experienced, on average, seven interruptions per day, while Gonzalez and Mark (2004) observed workers to spend less than 3 min on any task before switching to another. These studies reveal serious potential negative effects of interruptions on human performance.

1.1. Effects of interruptions on task performance

In a review of interruption literature, Li et al. (2011) categorized primary tasks used in interruption research as: procedural (e.g., Gillie and Broadbent, 1989), problem-solving (e.g., Hodgetts and Jones, 2006), and decision-making tasks (e.g., Hodgetts et al., 2014). (These three types of tasks are reviewed more in-depth in the following sections, including specific examples of each). Interruptions have been manipulated in terms of time of presentation (e.g., Monk et al., 2004), number of occurrences (e.g., Eyrolle and Cellier, 2000), complexity (e.g., Ziljstra et al., 1999), and similarity to the primary task (e.g., Ledoux and Gordon, 2006). The general objective of these studies has been to assess relative effects on primary task performance. Response measures have commonly included primary task completion time (e.g., Edwards and Gronlund, 1998), primary task error rate (e.g., Monk et al., 2008), interruption lag (defined as the time taken between acknowledging a pending interruption and beginning the interruption task; e.g., Czerwinski et al., 2000), and resumption lag (defined as the time taken to resume the primary task after completing the interruption task; e.g., Cades et al., 2008). While findings in the existing literature suggest interruptions of all types have negative effects on primary tasks of all types, the vast majority of primary tasks that have been studied are cognitive in nature, with examples

* Corresponding author. Dept. ISE, North Carolina State University, 400 Daniels Hall, Raleigh, NC 27695-7906, USA.
E-mail address: dbkaber@nscu.edu (D. Kaber).

http://dx.doi.org/10.1016/j.apergo.2016.08.022

including: essay writing (Foroughi et al., 2014), aircraft radar screen monitoring for targets (Hodgetts et al., 2014), or completing a series of tasks in a video game (Gillie and Broadbent, 1989; Edwards and Gronlund, 1998).

As Wickens et al. (2013) suggest, existing literature reports conflicting results regarding the effect of similarity of an interruption to a primary task; some studies indicate similar interruptions cause interference in working memory (WM), leading to degraded performance (e.g., Lee and Duffy, 2012). However, other research claims interruptions have a negative effect on performance regardless of how similar demands may be to the primary task (e.g., Speier et al., 1999). Inconsistent findings have also been found regarding interruption complexity. Some research claims that more complex interruptions degrade primary task performance more so than simpler interruptions (Hodgetts and Jones, 2006; Cades et al., 2007). However, other research suggests there is no significant difference (e.g., Ziljstra et al., 1999). One possible explanation for these inconsistencies is the use of different experimental paradigms across studies. Another explanation might be the use of different definitions of similarity or complexity by researchers. For example, level of complexity of interruption tasks has been defined based on the nature of operations (perceptual, cognitive and motor) and operation counts as well as task time. However, task completion time is actually an outcome of complexity and does not represent a work design parameter. Related to this, task time limits represent artificial constraints that are not normally imposed in real-world assembly operations.

Given the inconsistent research findings, we conducted a review of existing literature on interruption similarity and complexity with a focus on static tasks (i.e., tasks that do not evolve during an interruption, as in assembly work) and explain results in terms of Activation-Based Memory for Goals (MFG; Altmann and Trafton, 2002) theory. MFG theory has three facets that predict cognition: (1) an interference level, the theoretical level above which a goal needs to be in order to direct behavior; (2) a strengthening constraint, which is required to make a goal higher than the theoretical interference level and consequently direct behavior; and (3) a priming constraint, which predicts the relationship between cue strength and goal encoding. MFG models the cognition involved in switching between goals, and the relative interference that multiple goals pose on each other, making it a suitable theory for explaining the effects of interruption similarity and complexity.

1.2. Similarity of interruption and primary task demands

Lee and Duffy (2012) conducted an experiment with two types of interruption tasks and two types of primary tasks, including math word problems and simple word processing, in order to assess effects of similarity of the interrupting task on primary task performance. They found combinations of similar tasks (e.g., a word problem task being interrupted by another word problem) to produce longer task completion times and higher error rates than combinations of dissimilar tasks. Similarly, Eyrolle and Cellier (2000) also demonstrated that similar interruptions degraded task completion time and increased error rates in a rule-based perception task. Regarding procedural tasks, Gillie and Broadbent (1989) and Edwards and Gronlund (1998), using a similar paradigm, found similar interruptions in a procedural “to-do list” task to degrade task completion time and increase error rates. Since similarity can be defined in multiple ways (e.g., Wickens et al., 2013), it should be noted that the studies reviewed in this section manipulated the similarity of interruption task operations, as opposed to similarity of information coding or modalities of information presentation. MFG theory suggests that when people switch from one task/goal to another, the residual goal from the previous task interferes with the new goal, leading to the negative effects demonstrated by these experiments. Coupled with memory theory (Wickens, 1992, pp. 227–228), it is possible that similarity-induced confusion of goals among tasks might lead to greater negative effects on performance than under dissimilar goal conditions.

1.3. Complexity of the interruption task

Eyrolle and Cellier (2000) defined complexity as the amount of information processed during the interruption task, and investigated the effects of interruption complexity on rule-based task performance. They reported a marginally significant effect on error rate, but no effect on task completion time. Hodgetts and Jones (2006), in an experiment utilizing a problem-solving Tower of London task, found resumption lag to be shorter following a simple interruption (completion of a “mood checklist”) than following a complex interruption (a verbal reasoning task). Regarding the effects of complex interruptions on procedural tasks, Gillie and Broadbent (1989) concluded interruption complexity was an indicator of a disruptive interruption; complex interruptions were demonstrated to increase primary task completion time compared to simple interruptions. Finally, in a series of procedural VCR-programming tasks, several studies found more complex interruptions increased resumption lag (Monk et al., 2004; Cades et al., 2008) and error rate (Monk et al., 2008). However, Cades et al. (2007) found that resumption times in the complex (3-back recall) and simple (1-back recall) conditions were not statistically different, concluding that interruption complexity may not be the only reason for disruptiveness. Despite the differing results, complex tasks require greater engagement than simple tasks, possibly leading to the complex-task goal being more active than goals for simple interruptions, as explained by MFG. Being more active, the complex-interruption goal may interfere to a greater extent with the primary task goal upon return to the primary task, leading to increased resumption lag, error rate, and task completion time.

1.4. Other measures of interruption effects

Although error rates and response times have been used extensively for assessing the effects of interruptions on primary task performance, little work has reported on physiological responses to interruptions (e.g., Katidioti et al., 2014). Related to this, there is a substantial body of work on the use of physiological responses as proxy measures of workload with advantages of unobtrusiveness and high resolution. For example, Young and Stanton (2005) said that heart rate (HR) provides a simple method of monitoring of workload state without being intrusive on primary task performance. During task performance, participants expend mental effort, which has been associated with increased HR. In a review article of pilot workload assessment, Roscoe (1992) reported HR responses to be a reasonably accurate and reliable indicator of workload changes. Paxion et al. (2014) identified HR as a sensitive indicator of high- and low-complexity driving situations. A review by Scerbo et al. (2001) revealed the use of HR for measuring pilot workload in a flight simulator, automobile driver workload, and workload for electrical equipment operators conducting a visual-manual task. These studies suggest HR may be sensitive to workload associated with interruptions in various domains.

1.5. Problem statement

Existing work in the field of interruption effects has identified potential negative effects, including degradations in productivity (e.g., task completion time), increased error rates, and increased resumption lag. Furthermore, there are conflicting findings regarding the effects of complexity and similarity of interruption tasks on primary task performance. The vast majority of research in the field focuses on cognitive tasks with little work having investigated interruptions in procedural visual-manual tasks, such as worker assembly of products in a manufacturing environment. Since there are over 12 million manufacturing workers in the United States (Bureau of Labor Statistics (2015)), it is important to assess the effects of interruptions on this population with potential generalizability to other domains involving similar types of tasks. Given these shortcomings in the existing literature, we conducted an experiment to analyze effects of similar and complex interruptions on performance in a simulated industrial assembly operation. In order to simulate a manual assembly task as closely as possible, we defined a Lego assembly task involving block grasping, repositioning, and fastening actions, similar to the behaviors required to perform an occupational assembly task. This type of task differs from computer simulations (requiring mouse, keyboard, and/or touchscreen interaction), as used in previous interruption experiments. From a practical perspective, the research was aimed at addressing questions such as: what types of interruptions in assembly work would be least disruptive to operators? And what types of interruptions should be delayed (if possible) to the end of an assembly cycle?

1.6. Hypotheses

Based on the literature review, Table 1 presents hypotheses for each response in our experiment.

2. Method

2.1. Participants

Eighteen undergraduate and graduate students with varied academic backgrounds (11 male, 7 female) between 19 and 32 years of age ($\mu = 22.6$, $\sigma = 3.42$) were recruited for this study. The sample size was based on a calculation considering pilot test data (response means for experiment condition and variance) and a Type I error of 0.05 and Type II error of 0.2. All participants self-reported normal or corrected 20/20 vision, no color vision impairment, or upper extremity disability, which were required for performance of the assembly operation. Participants were compensated at a rate of $15/hour for their time.

2.2. Independent variables

The independent variable manipulations included complexity (simple vs. complex) and similarity (similar vs. dissimilar) of the interruption task to the primary assembly operation. The primary task was a 36-layer Lego assembly operation with a procedural, step-by-step manual. To validate the task to ensure it was comparable to a real-world assembly task, we conducted a fundamental motion analysis (Freivalds, 2014, pp. 151–154) and compared it to a motion analysis for iPhone assembly (Ma, 2014) since both involve manual assembly directed by a procedural, step-by-step manual. Fig. 1 reveals the Lego assembly to be comparable to the basic iPhone assembly operation in terms of fundamental motion behavior (e.g., search, select, reach, grasp, move, etc.).

The first three pages from an example manual are presented in Fig. 2a. The numbers located in top-left corner indicate the current assembly layer and the bricks in the box indicate required pieces for the layer. WM is required to encode the task goal, store the color and shape of Lego pieces and to reposition the Lego and/or assembly in order to correctly fasten the Lego to the assembly. Limitations in WM capacity could lead to repeated viewing of the instructions (and, consequently, increased completion time) and/or errors in the assembly. When switching to the interruption task, it was expected that WM chunks dedicated to Lego shape, color, location, etc. would decay until return to the primary task. The encoded goal of completing the primary assembly was also expected to decay (as predicted by MFG), since a new goal associated with the interruption task was encoded, but was expected to be reinstantiated upon return to the primary task.

The similar interruption task required participants to perform another Lego assembly operation while the dissimilar interruption tasks were a series of pencil-and-paper math problems, unrelated to the primary Lego assembly operation (Geary et al., 1986). The simple-similar interruption task was a 12-layer Lego assembly, requiring only one type of Lego brick, presented with a procedural, step-by-step manual (see Fig. 2b). The complex-similar level was a five-layer assembly operation involving a variety of Lego blocks with the assembly reference being a one-page static "3D" computer image (see Fig. 2c), requiring participants to mentally rotate the assembly in order to correctly complete layers. The four combinations of interruption tasks are detailed in Table 2.

The simple-dissimilar interruption task involved basic addition problems and the complex-dissimilar task involved higher-difficulty multiplication problems requiring carry-over of a digit to the hundredths place. All problems within each level of complexity were formulated to ensure use of the same number of WM chunks. Goals, Operators, Methods, and Selection rules (GOMS) models were developed to determine the maximum WM chunk counts required by each task (Kieras, 1997), an approach used by Magrabi et al. (2010) to justify complexity levels of interruption tasks. The models revealed that the complex tasks imposed greater WM loads than the simple tasks. Pilot testing corroborated the models, as they required significantly more time ($p < 0.05$) to complete than the simple tasks, across both similar and dissimilar tasks. A similar manipulation-check procedure was used by Speier et al. (1999).

2.3. Apparatus

The experiment setup consisted of two desktop computers presenting instructions for the primary assembly and similar interruption tasks, along with space in the interruption task area to complete the paper-based dissimilar interruption tasks. Two tables, oriented perpendicular to each other, were used to administer the main and interruption tasks separately (see Fig. 3). Each unique Lego type was placed in a separate part bin on each table. Bins were

Table 1
Hypotheses formulated for the experiment.

Response	Interruption (compared to non-interruption conditions)	Complex interruptions (compared to simple interruptions)	Similar interruptions (compared to dissimilar interruptions)
Error rate	Increased	Increased	Increased
Post-interruption productivity	Decreased	Decreased	Decreased
Resumption lag	N/A	Longer	Longer
Workload	N/A	Increased	Increased

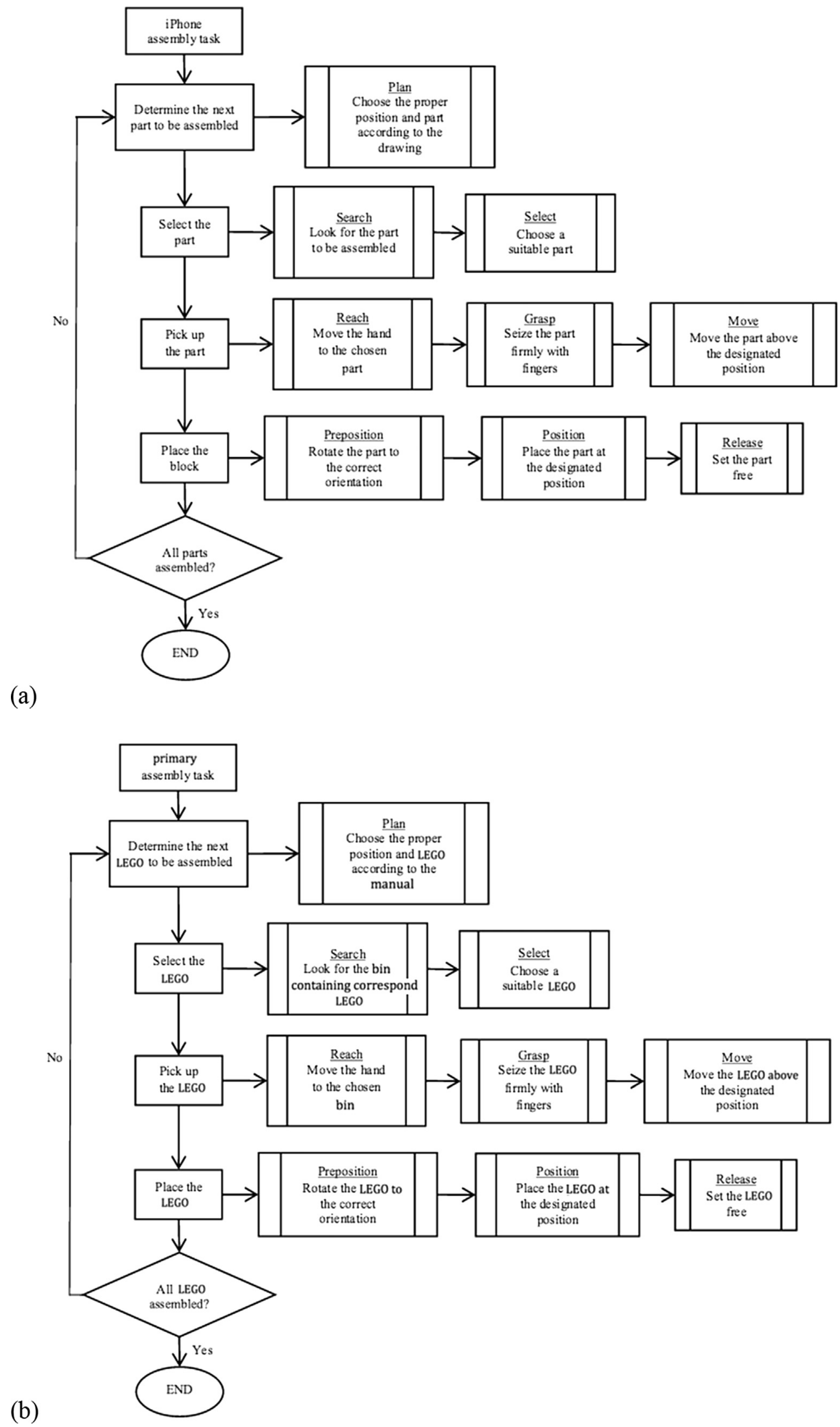

Fig. 1. Motion analysis of (a) iPhone assembly and (b) Lego assembly tasks.

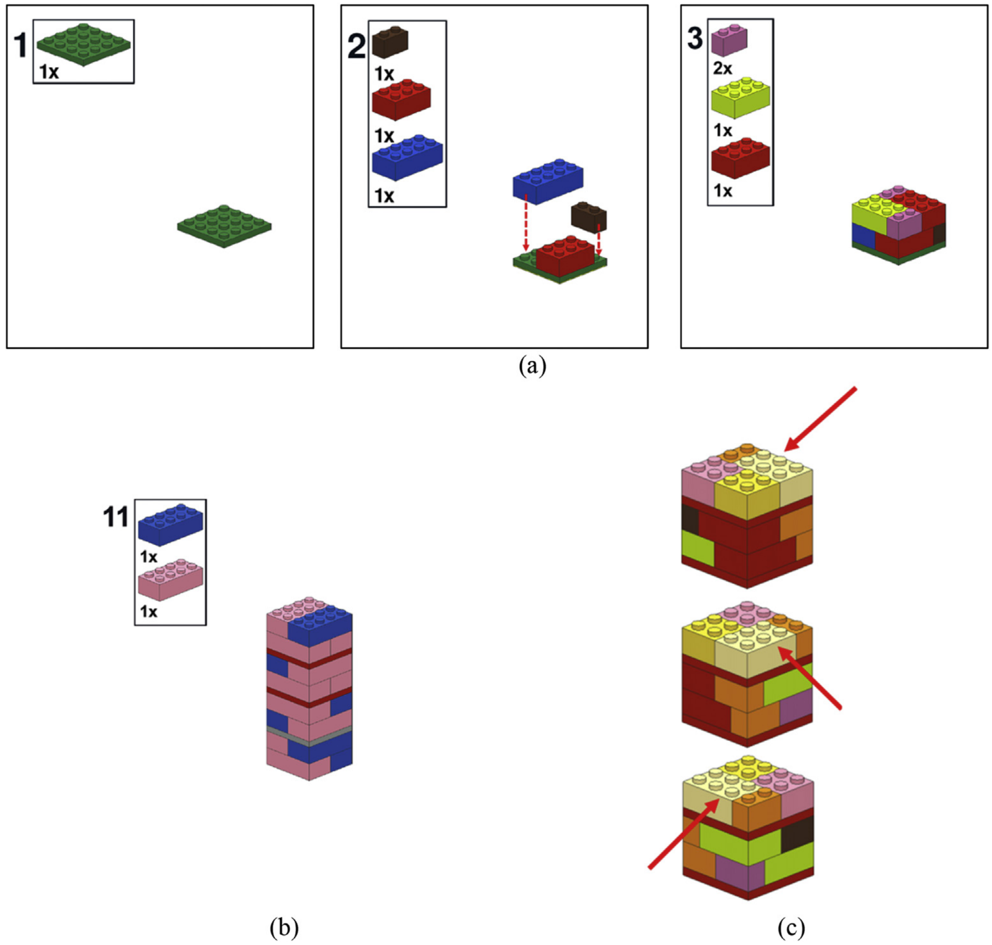

Fig. 2. (a) An example step-by-step primary task assembly manual; (b) an example 11th-layer similar-simple interruption assembly manual; and (c) the manual for the similar-complex interruption task.

Table 2
Descriptions of the interruption tasks.

Complexity level	Similarity to primary task	Description	WM chunk counts
Simple	Similar	Assembly operation with Legos of the same shape (i.e., 2*4) and a step-by-step manual	4
	Dissimilar	Addition problems; A + B where 1 < A<9, 1 < B < 9, and A*B < 25	3
Complex	Similar	Assembly operation with Legos of different shapes and static 3D instructions.	7
	Dissimilar	Multiplication problems which require carrying over a digit to the hundredths place; A*B where A<99, B>9, and A! = B	6

located at a distance of 547.7 mm from participants, which was calculated based on the 5th percentile of zone of convenience reach for the US population (Bridger, 2008).

2.4. Procedure

Prior to the scheduled experiment time, potential participants were asked to disclose any color vision impairment, upper extremity disability, and prior experience with Legos. We excluded participants who indicated recent and frequent Lego use in order to recruit a uniform set of participants in terms of Lego assembly skill. Upon arrival to the experiment, participants read and signed an informed consent document and were subsequently asked to don a HR chest strap monitor (Polar S810i HR Monitoring system, Polar Electro Oy, Finland). They were then given an overview of the experiment, and a training session was provided to familiarize participants with the Lego assembly operation and interruption tasks.

For trials with interruption, participants heard an automated bell, indicating a stop for the primary task with immediate execution of the interruption task at the other table until completion. The interruption occurred at random times between 3 and 5 min after the start of the primary task to allow participants to cognitively immerse themselves in the work before the interruption occurred. (The uninterrupted primary task lasted approximately 6 min) The interruption time frame was chosen based on situation awareness studies indicating participants require ~3–5 min to learn system states before a first assessment should be made (Endsley, 2000). Furthermore, Murray and Khan (2014) reported that interruptions occurring near the end of a primary task were more impactful than

Fig. 3. Experiment setup, including instructions, primary task assembly area, and interruption task area.

those near the beginning, potentially promoting the sensitivity of our study.

Once the interruption task was completed, participants were asked to immediately return to the primary task. Although MFG theory predicts that the duration of an interruption can affect post-interruption performance, we required participants to finish the interruption task in order for the experiment to more closely replicate a real-life interruption situation. In industrial assembly operations, for example, if a worker is pulled from a production line for rework of a unit, the rework is completed prior to return to the primary assembly operation. Once participants in our study finished the primary assembly operation, they were given a break before proceeding to the next trial. During this break, the part bins containing specific types of Legos were repositioned to prevent participants from learning part locations. This procedure was repeated eight times, lasting approximately 2.5 h per participant.

2.5. Experiment design

The experiment followed a 2 × 2 within-subjects design representing a full crossing of the levels of interruption complexity and similarity. Each condition was replicated for each participant in order to quantify any performance variability due to individual differences, yielding eight treatment trials. Participants were also administered a baseline trial that included no interruptions, resulting in nine total trials per participant. The order of administration of the nine trials was randomized in order to mitigate potential order effects.

2.6. Dependent variables

Dependent variables included error rate in the primary

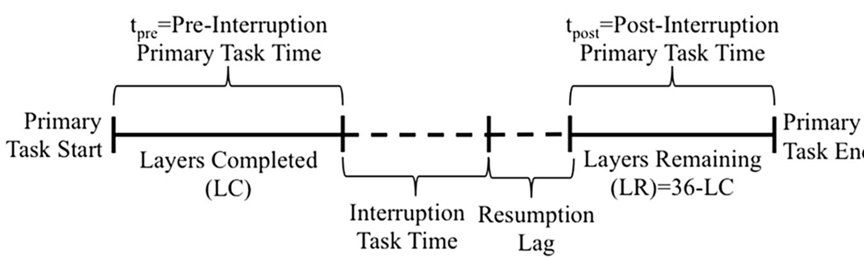

Fig. 4. Timeline of events for each trial and terminology used as a basis for response calculations (each assembly contained 36 total layers).

Table 3
Dependent variable definitions.

Dependent variable	Description
Error rate	The number of primary-task layers in which an assembled layer was not identical to the manual with respect to the color, shape, and orientation, divided by the total number of layers of the assembly.
Post-interruption time per layer	The average time to complete the remaining layers of the primary assembly after interruption task performance; t_{post}/LR in Fig. 4 (LR equals the total number of layers, 36, minus the number of pre-interruption layers completed). To assess the effect of interruptions across all manipulations, analysis was performed on the time per layer in the treatment tasks minus the equivalent time per layer in the baseline task; ($t_{post,treatment}$/LR)-($t_{post,baseline}$/LR). Higher productivity equates to a smaller time per layer (e.g., if LR = 10 in the treatment scenario, the $t_{post,baseline}$ was a measure of the time to complete the last 10 layers in the baseline scenario). This measure was calculated in order to mitigate any potential learning or fatigue effects associated with the baseline assembly (e.g., the participant was unfamiliar with the bin locations at the beginning of the baseline scenario, which could have skewed the results had we compared the time-per-layer of the post interruption time to the time-per-layer of the total baseline assembly). This was calculated post-experiment via video recordings of each participant.
Resumption lag	The time to resume the primary task after interruption, measured as the time between the end of the interruption task and a participant placing a first Lego on the primary assembly. Resumption lag was calculated based on timestamps recorded by experimenter keystrokes at a timing software application when a participant (1) finished the interruption task, and (2) when they placed the first Lego on the primary assembly.
Increase in HR	The percent increase in HR during the interruption task, compared to the participant's HR before interruption. Although HR variability has previously been used as an indicator of cognitive workload in low physical demand tasks, the duration of our tasks was too short (1.5–3 min) to accurately determine the variability measure (Wilson, 1992; Jorna, 1992). Furthermore, the physical activity of the Lego assembly was expected to corrupt any indication of cognitive load.

assembly operation, post-interruption time per layer, resumption lag time, and percent increase in HR from pre-interruption performance to post-interruption performance. Fig. 4 provides information on how each response was calculated, and Table 3 reports the dependent variables and their descriptions.

2.7. Data analyses

Prior to any inferential statistical analyses, all response data sets were screened for outlier data points. Outliers were defined as extreme points as identified by residual plots against fitted values (Netter et al., 1990). Regarding the post-interruption time per layer analysis, 45 data points were excluded from the *t*-test analysis and 33 data points from the Analysis of Variance (ANOVA) due to experiment equipment issues (e.g., missing video recordings), failure of participants to follow instructions, or identified as outliers. For the resumption lag analysis, 7 data points were excluded, and for the HR analyses, 26 data points were excluded. Since interruption task duration was not controlled (see Table 4 for descriptive statistics on the interruption times) and has been shown to significantly alter behavior in prior research (e.g., Monk et al., 2008), the effect of interruption duration was included in the ANOVA for post-interruption productivity and resumption lag. However, the term did not prove significant in either response model ($p = 0.12$ and 0.72, respectively) and was, therefore, removed for all reported statistical analyses. Similarly, trial number was initially added to the response models as a covariate but was removed from the HR analysis due to a lack of significance ($p = 0.78$). There was a significant trend of decreasing time per layer ($p < 0.01$) and decreasing resumption lag ($p = 0.03$) as the experiment progressed.

A paired *t*-test was used to analyze the effect of interruptions across conditions on post-interruption time per layer (productivity) while ANOVA was used to analyze the effects of similarity and complexity of interruptions on the dependent variables. Constant variance and residual normality were assessed using accepted statistical diagnostic procedures to ensure all ANOVA assumptions were met. If there was evidence of any assumption violation, either a square-root transform was applied to the raw responses (Netter et al., 1990) or the responses were ranked in order to conduct a nonparametric ANOVA. All plots of response measures also include error bars representing one standard deviation from the mean. For all statistical tests, we also report observed or post-hoc power ($1-\beta$) values.

Table 4
Descriptive statistics for interruption duration.

Independent variable	Level	Mean (s)	Standard deviation (s)
Similarity	Similar	88.34	24.04
	Dissimilar	138.72	94.20
Complexity	Simple	82.88	16.32
	Complex	144.90	93.21
Interaction	Similar-Simple	84.66	15.36
	Similar-Complex	92.02	31.76
	Dissimilar-Simple	81.14	17.26
	Dissimilar-Complex	196.31	104.31

Table 5
Descriptive statistics for number of assembly operation errors.

Condition	Mean	Standard deviation
Similar-simple	0.057	0.236
Similar-complex	0.056	0.232
Dissimilar-simple	0.000	0.000
Dissimilar-complex	0.083	0.280

3. Results

3.1. Error rates

Shown in Table 5, error rates in the primary assembly operation were uniformly low across participants in all conditions. There were no significant effects of interruption similarity or complexity on error rates ($p > 0.05$).

3.2. Post-interruption time per layer

A paired *t*-test revealed a significant effect of interruptions ($t(98) = -4.353$, $p < 0.001$, $d = 0.101$) on mean time per layer. Reported in Table 6, the mean time per layer decreased significantly (i.e., productivity increased) after an interruption ($\mu = 10.290$ s, $\sigma = 2.476$ s), as compared to the baseline trial ($\mu = 11.154$ s, $\sigma = 2.986$ s).

An ANOVA was performed on the square-root-transform of mean time per post-interruption layer in order to assess the effect of different types of interruptions. Results revealed a significant effect of complexity ($F(1,91) = 4.2903$, $p = 0.041$, $\omega^2 = 0.013$, $1-\beta = 0.4243$), but no effect of similarity ($F(1,91) = 1.5508$, $p = 0.216$, $\omega^2 = 0.002$, $1-\beta = 0.1097$) or an interaction effect ($F(1,91) = 0.326$, $p = 0.570$, $\omega^2 = 0.000$, $1-\beta = 0.05$). Complex interruptions significantly decreased time per layer (i.e., increased productivity) after the interruption, as compared to simple interruptions.

3.3. Resumption lag

An ANOVA was performed on resumption lag z-scores, revealing a significant effect of similarity ($F(1,114) = 7.4546$, $p = 0.007$, $\omega^2 = 0.050$, $1-\beta = 0.7032$), but no significant effect of complexity ($F(1,114) = 0.2030$, $p = 0.653$, $\omega^2 = 0.000$, $1-\beta = 0.05$) or an interaction effect ($F(1,114) = 0.0006$, $p = 0.981$, $\omega^2 = 0.000$, $1-\beta = 0.05$). Shown in Fig. 5, resumption lag was significantly longer when the interruption task was similar to the main assembly operation vs. the dissimilar arithmetic task.

3.4. Heart rate increase

A nonparametric ANOVA was performed on the ranked percent increase in HR response, revealing a significant effect of similarity ($F(1,98) = 11.3832$, $p = 0.001$, $\omega^2 = 0.061$, $1-\beta = 0.8838$), but no effect of complexity ($F(1,98) = 0.7452$, $p = 0.390$, $\omega^2 = 0.000$, $1-\beta = 0.05$) or an interaction effect ($F(1,98) = 0.9448$, $p = 0.333$,

Table 6
Descriptive statistics for time per layer in interruption trials.

Condition	Mean (s)	Standard deviation (s)
Difference between interruption and baseline	−0.863	1.973
Similar-simple	10.002	5.095
Similar-complex	10.915	2.800
Dissimilar-simple	10.600	2.155
Dissimilar-complex	10.339	2.611

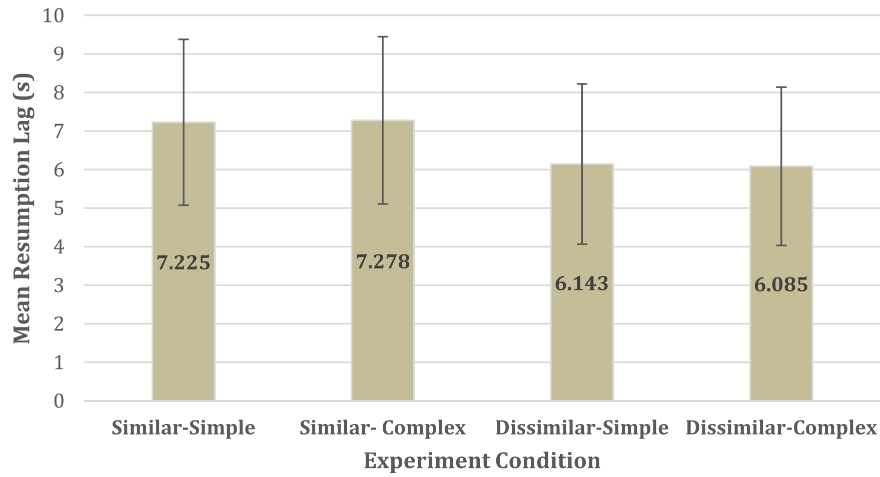

Fig. 5. The effect of experiment condition on resumption lag.

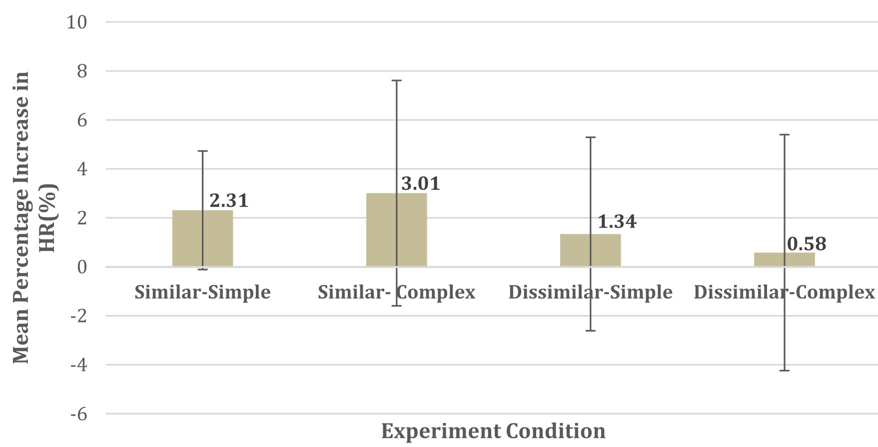

Fig. 6. The effect of experiment condition on percent increase in HR.

$\omega^2 = 0.000$, $1-\beta = 0.05$). Shown in Fig. 6, similar interruptions increased HR more than dissimilar interruptions.

4. Discussion

In general, the hypotheses formulated on the potential influence of interruption similarity and complexity on the various response measures were based on trends reported in the prior literature. We expected error rates to increase in the presence of interruptions, in the presence of more complex interruptions, and in the presence of similar interruptions, accordingly. Results did not support these hypotheses, as error rates were consistently low across all experimental conditions. When participants resumed their previously-suspended goal of primary task completion, they were presented with color-coded instructions on a large computer screen with instructions beginning where they left off prior to the interruption, representing a very strong visual cue. From the viewpoint of MFG theory, these strong visual cues supported performance to the point that there were very few assembly errors. Some existing work has reported no effect of interruptions on errors, particularly when task cues are strong, suggesting that error rate may simply not be sensitive to the effect of interruptions (Hodgetts et al., 2014; Hodgetts and Jones, 2006; Monk et al., 2008; Magrabi et al., 2010). Beyond this, Monk et al. (2004) and Li et al. (2008) reported that the timing of the interruption had a significant effect on error rate. Since we randomized the timing of the interruptions relative to the primary task, it is also possible that this procedure diminished any effect on error rates following primary task resumption.

Contrary to our expectation that productivity would decrease following an interruption, post-interruption productivity in the primary task increased. We also expected that complex and similar interruptions would degrade post-interruption productivity; neither of these hypotheses were supported by the results. There was a significant effect of complexity, but productivity increased after the interruption, contrary to expectation. Since complex interruptions took longer to complete, participants may have felt the need to compensate for time away from the primary task by working more quickly after the interruption (despite the fact that participants were instructed to complete the primary task as quickly as possible, under all circumstances). Most of the existing interruption research has studied primary tasks requiring high-level critical thinking, including procedure planning. Even in the VCR programming task used by Monk et al. (2008), participants were not guided step-by-step through the process and had to formulate a procedure. Our primary Lego assembly task required participants to simply follow the instructional manual and the participants might have been able to develop some degree of automaticity in task performance. This task design differentiates our research from other previous studies. In regard to our complex interruptions, both the difficult multiplication problems and complex Lego assembly required a higher-degree of cognition, including extensive WM use and/or mental rotation of an assembly in WM, as compared to the primary Lego assembly task. Therefore, our results suggest that arousal via cognitively-demanding interruptions may have a positive carry-over effect on performance of more mundane, procedural assembly tasks. Referring to MFG theory, it is possible that increased complexity required greater participant engagement, as well as increased time to complete the tasks, leading to decay of the primary task completion goal below the "interference level"; that is, the theoretical threshold above which goals need to be activated to direct behavior. Upon completion of the complex interruptions, participants were required to encode a new goal and, in combination with the anticipation of being close to finishing the task, might have led to increased arousal and higher post-interruption productivity.

There was no significant resumption lag difference between simple/complex interruption tasks, but lag was significantly greater when a similar interruption was administered than a complex one. The second finding suggests that there may be interference in WM between the interruption task and the primary task in the 3–8 s it took to transition back to the primary assembly operation. As explained by MFG theory, our similar interruption conditions might have led to a higher rate of decay of the primary task goal than dissimilar interruptions, providing enough interference with the primary task goal to increase workload (reflected in the HR response) and delay return to the primary task. It is also possible that the differences in media (i.e., colorful computer-based instructions for the similar interruptions vs. paper and pencil for the dissimilar interruptions) might have contributed to these results, as there may have been less goal interference switching between different mediums than switching to a task presented in the same medium.

The cardiac response revealed no significant difference in interruption complexity, contrary to our hypothesis. However, there was a significant difference in HR due to interruption similarity, supporting our expectation that similar interruptions would increase cognitive workload. Our results indicate that participants experienced higher levels of workload (as indicated by the increased HR) during the similar interruptions, as compared with dissimilar interruptions. Participants may have experienced additional cognitive loading due to the interference associated with switching between goals of similar tasks, as predicted by MFG. The large standard deviations shown in Fig. 6 are likely due to individual differences in HR, since HR is dependent on many factors, including physical fitness, caffeine intake, age, gender, etc. (Individual differences were accounted for in our statistical analyses by normalizing the response measures for each participant.) Our results on interruption complexity, while unexpected, are similar to Ziljstra et al. (1999), who found that a more complex interruption did not have a significant effect on psychological state.

5. Conclusions

The primary objective of this study was to assess the effect of interruptions, including similarity and complexity characteristics, on performance in a simulation of a procedural visual-manual assembly task. Findings indicate that complex interruptions may cause arousal that increases productivity in a primary assembly operation following an interruption. Similar interruptions both increase workload and time required to resume a primary assembly operation following an interruption. Although some of the results we obtained suggest differential effects of interruption complexity and similarity, the behavior exhibited by participants can generally be explained by MFG theory. The goals associated with the different tasks need to be activated and reactivated throughout the task and differences in similarity and complexity of primary and interruption tasks lead to differences in how the goals are encoded, decay, and interfere with one another. In regard to the practical research questions we identified in our Problem Statement, findings suggest that while it is not ideal to interrupt a worker during an assembly operation, status updates or performance of mental calculations may be less disruptive than assembly rework or repairs, which should be saved for the end of a workshift or an ongoing assembly operation.

5.1. Limitations

The primary task in our experiment was a completely manual assembly operation; that is, no tools were required for the assembly, which may not be representative of many industrial jobs. Another limitation was that participants were interrupted by an automated bell, which may not be representative of actual job environments where interruptions may come in the form of requests for information by other employees or supervisors. Finally, the primary task that we studied might have been relatively simple, as reflected by the consistently low error rates across experimental conditions; however, the task design was intended to be a realistic representation of assembly operation performance and to provide a basis for generalizing interruption research results to the assembly domain, in general.

5.2. Future work

Future work should examine the use of different, more complex assemblies potentially including tool use. Further work could involve a paradigm in which participants choose when to stop the primary task and begin the interruption task rather than requiring them to immediately address the interruption. Finally, a follow-on experiment should investigate the utility of interrupting a mundane/monotonous task for boosting post-interruption productivity.

Acknowledgements

This research was supported by the Edward P. Fitts Department of Industrial & Systems Engineering at North Carolina State University. Lego is a trademark of the Lego Group of companies, which did not sponsor, authorize or endorse this research. None of the authors of this study has any financial relationship with the Lego Group.

References

Altmann, E.M., Trafton, J.G., 2002. Memory for goals: an activation-based model. Cogn. Sci. 26, 39–83. http://dx.doi.org/10.1016/S0364-0213(01)00058-1.
Bridger, R., 2008. Introduction to Ergonomics. CRC Press, Boca Raton, FL.
Bureau of Labor Statistics, 2015. Industries at a Glance: Manufacturing. Retrieved from. http://www.bls.gov/iag/tgs/iag31-33.htm.
Cades, D., Davis, D., Trafton, J., Monk, C., 2007. Does the difficulty of an interruption affect our ability to resume?. In: Hum Fac Erg Soc P. Human Factors and Ergonomics Society, Santa Monica, CA, pp. 234–238. http://dx.doi.org/10.1177/154193120705100419.
Cades, D.M., Werner, N., Boehm-Davis, D.A., Trafton, J.G., Monk, C.A., 2008. Dealing with interruptions can be complex, but does interruption complexity matter: a mental resources approach to quantifying disruptions. In: Hum Fac Erg Soc P. Human Factors and Ergonomics Society, Santa Monica, CA, pp. 398–402. http://dx.doi.org/10.1177/154193120805200442.
Czerwinski, M., Cutrell, E., Horvitz, E., 2000. Instant messaging and interruption: influence of task type on performance. In: Proc OZCHI Conf. Association for Computing Machinery, New York, NY, pp. 356–361.
Edwards, M., Gronlund, S.D., 1998. Task interruption and its effects on memory. Memory 6 (6), 665–687. http://dx.doi.org/10.1080/741943375.
Endsley, M.R., 2000. Direct measurement of situation awareness: validity and use of SAGAT. In: Endsley, M.R., Garland, D.J. (Eds.), Situation Awareness Analysis and Measurement. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 147–173.

Eyrolle, H., Cellier, J.M., 2000. The effect of interruptions in work activity: field and laboratory results. Appl. Ergon. 31, 537–543. http://dx.doi.org/10.1016/S0003-6870(00)00019-3.
Foroughi, C.S., Werner, N.E., Nelson, E.T., Boehm-Davis, D.A., 2014. Do interruptions affect quality of work? Hum. Factors 56 (7), 1262–1271. http://dx.doi.org/10.1177/1541931213571035.
Freivalds, A., 2014. Methods, Standards, & Work Design, thirteenth ed. WCB/McGraw-Hill, Boston.
Geary, D.C., Widaman, K.F., Little, T.D., 1986. Cognitive addition and multiplication: evidence for a single memory network. Mem. Cogn. 14 (6), 478–487. http://dx.doi.org/10.3758/BF03202519.
Gillie, T., Broadbent, D., 1989. What makes interruptions disruptive? A study of length, similarity, and complexity. Psychol. Res. 50, 243–250. http://dx.doi.org/10.1007/BF00309260.
Gonzalez, V.M., Mark, G., 2004. Constant, constant multi-tasking craziness: managing multiple working spheres. In: Proc SIGCHI Conf Hum Factors Comp Sys. ACM, New York, NY, pp. 113–120. http://dx.doi.org/10.1145/985692.985707.
Hodgetts, H.M., Jones, D.M., 2006. Interruption of the tower of london task: support for a goal-activation approach. J. Exp. Psychol. Gen. 135 (1), 103–115. http://dx.doi.org/10.1037/0096-3445.135.1.103.
Hodgetts, H.M., Vachon, F., Tremblay, S., 2014. Background sound impairs interruption recovery in dynamic task situations: procedural conflict? Appl. Cogn. Psych. 28, 10–21. http://dx.doi.org/10.1002/acp.2952.
Jorna, P.G., 1992. Specral analysis of heart rate and psychological state: a review of its validity as a workload index. Biol. Psychol. 34, 237–257.
Katidioti, I., Borst, J.P., Taatgen, N.A., 2014. What happens when we switch tasks: pupil dilation in multitasking. J. Exp. Psychol. Appl. 20 (4), 380.
Kieras, D.E., 1997. A Guide to GOMS model usability evaluation using NGOMSL. In: Helander, M., Landauer, T., Prabhu, P. (Eds.), Handbook of Human-computer Interaction. North-Holland, Amsterdam, pp. 733–766.
Ledoux, K., Gordon, P.C., 2006. Interruption-similarity effects during discourse processing. Memory 14 (7), 789–803. http://dx.doi.org/10.1080/09658210600679915.
Lee, B.C., Duffy, V.G., 2012. The effects of task interruption on human performance: a study of the systematic classification of human behavior and interruption frequency. Hum. Factor Ergon. Man. 25 (2), 137–152. http://dx.doi.org/10.1002/hfm.20603.
Li, S.Y.W., Blandford, A., Cairns, P., Young, R.M., 2008. The effect of interruptions on postcompletion and other procedural errors: an account based on the activation-based goal memory model. J. Exp. Psychol. Appl. 14 (4), 315–328. http://dx.doi.org/10.1037/a0014397.
Li, S.Y.W., Magrabi, F., Coiera, E., 2011. A systematic review of the psychological literature on interruption and its patient safety implications. J. Am. Med. Inf. Assoc. 19 (6), 6–12. http://dx.doi.org/10.1136/amiajnl-2010-000024.
Ma, W., 2014. An Application of Quantitative Methods for Motor Ability Level Classification, Performance Prediction and Training Protocol Selection. North Carolina State University. http://repository.lib.ncsu.edu/ir/handle/1840.16/9546.
Magrabi, F., Li, S.Y.W., Day, R.O., Coiera, E., 2010. Errors and electronic prescribing: a controlled laboratory study to examine task complexity and interruption effects. J. Am. Med. Inf. Assoc. 17, 575–583. http://dx.doi.org/10.1136/jamia.2009.001719.
Monk, C.A., Boehm-Davis, D.A., Trafton, J.G., 2004. Recovering from interruptions: implications for driver distraction research. Hum. Factors 46 (4), 650–663. http://dx.doi.org/10.1518/hfes.46.4.650.56816.
Monk, C.A., Trafton, J.G., Boehm-Davis, D.A., 2008. The effect of interruption duration and demand on resuming suspended goals. J. Exp. Psychol. Appl. 14 (4), 299–313. http://dx.doi.org/10.1037/a0014402.
Murray, S.L., Khan, Z., 2014. Impact of interruptions on white collar workers. Eng. Manag. J. 26 (4), 23–28. Retrieved from. http://search.proquest.com/docview/1630248496?accountid=12725.
Netter, J., Wasserman, W., Kutner, M.H., 1990. Applied Linear Statistical Models, third ed. Irwin, Homewood, IL.
Paxion, J., Galy, E., Berthelon, C., 2014. Mental workload and driving. Front. Psychol. 5, 1–11. http://dx.doi.org/10.3389/fpsyg.2014.01344.
Roscoe, A.H., 1992. Assessing pilot workload. Why measure heart rate, HRV and respiration? Biol. Psychol. 34 (2), 259–287. http://dx.doi.org/10.1016/0301-0511(92)90018-P.
Scerbo, M.W., Freeman, F.G., Mikulka, P.J., Parasuraman, R., Di Nocero, F., Prinzel, L.J., 2001. The Efficacy of Psychophysiological Measures for Implementing Adaptive Technology (NASA TP-2001-211018). NASA Langley Research Center, Hampton, VA.
Speier, C., Valacich, J.S., Vessey, I., 1999. The influence of task interruption on individual decision making: an information overload perspective. Decis. Sci. 30 (2), 337–360.
Westbrook, J.I., Woods, A., Rob, M.I., Dunsmuir, W.T.M., Day, R.O., 2010. Association of interruptions with an increased risk and severity of medication administration errors. Arch. Intern. Med. 170 (8), 683–690. http://dx.doi.org/10.1001/archinternmed.2010.65.
Wickens, C.D., 1992. Engineering Psychology and Human Performance, second ed. Harper Collins, New York.
Wickens, C.D., Santamaria, A., Sebok, A., 2013, September. A computational model of task overload management and task switching. In: Hum Fac Erg Soc P. Human Factors and Ergonomics Society, Santa Monica, CA, pp. 763–767. http://dx.doi.org/10.1177/1541931213571167.
Wilson, G.F., 1992. Applied use of cardiac and respiration measures: practical considerations and precautions. Biol. Psychol. 34, 163–178. http://dx.doi.org/10.1016/0301-0511(92)90014-L.
Young, M.S., Stanton, N.A., 2005. Mental workload. In: Stanton, N.A., Hedge, A., Brookhuis, K., Salas, E., Hendrick, H.W. (Eds.), Handbook of Human Factors and Ergonomics Methods (Pp. 39–1 to 39-9). CRC Press, Boca Raton, FL.
Ziljstra, F.R.H., Roe, R.A., Leonora, A.B., Krediet, I., 1999. Temporal factors in mental work: effects of interrupted activities. J. Occup. Organ Psych. 72, 163–185. http://dx.doi.org/10.1348/096317999166581.