IISE TRANSACTIONS ON OCCUPATIONAL ERGONOMICS AND HUMAN FACTORS
https://doi.org/10.1080/24725838.2020.1770898

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

Transitions Between Low and High Levels of Mental Workload can Improve Multitasking Performance

Shannon Patricia Devlin[a], Nadine Marie Moacdieh[b], Christopher D. Wickens[c]*, and Sara Lu Riggs[a]

[a]Engineering Systems and Environment, University of Virginia, Charlottesville, Virginia, USA; [b]Industrial Engineering and Management, American University of Beirut, Beirut, Lebanon; [c]Alion Science and Technology, Boulder, Colorado, USA

OCCUPATIONAL APPLICATIONS

Complex and dynamic environments including military operations, healthcare, aviation, and driving require operators to transition seamlessly between levels of mental workload. However, little is known about how the rate of an increase in workload impacts multitasking performance, especially in the context of real-world tasks. We evaluated both gradual and sudden workload increases in the dynamic multitasking environment of an Unmanned Aerial Vehicle (UAV) command and control testbed and compared them to constant workload. Workload transitions were found to improve response time and accuracy compared to when workload was held constant at low or high. These results suggest that workload transitions may allow operators to better regulate mental resources. These findings can also inform the design of operations and technology to assist operators' management of cognitive resources, which include negating the adverse effects of vigilance decrements during low workload periods and data overload during high workload periods.

TECHNICAL ABSTRACT

Background: High workload and workload transitions can affect performance; however, it is not clear how the rate of transition from low to high workload influences performance in a multitasking setting.

Purpose: We investigated the effect of workload transition rate on performance in a multitasking environment that is akin to the expectations of operators in complex, data-rich work domains.

Method: An Unmanned Aerial Vehicle (UAV) command and control testbed was used to vary workload between low, high, gradually transitioning from low to high, and suddenly transitioning from low to high. Performance measures consisted of the response time and accuracy of one primary task and three secondary tasks. Analyses compared: (a) performance differences between gradual and sudden increases in workload; (b) performance during the low workload phases of the workload transitions; and (c) performance during the high workload phases of the workload transitions.

Results: Overall, there were limited performance differences between gradual and sudden workload transitions. However, both types of transitions led to better performance than constant workload, lending some support for the effort regulation explanation which suggests that participants actively evaluated the amount of mental resources necessary to successfully complete a task after a workload transition.

Conclusions: Gradual and sudden workload transitions benefit primary and secondary task performance, suggesting that the applicability of existing theoretical explanations depend on the context. For example, varying task demands can be a means to assist operators in the appropriate regulation of mental resources in domains with interdependent tasks. These findings can inform occupation and technology design to support task management.

ARTICLE HISTORY
Received 8 June 2019
Accepted 14 May 2020

KEYWORDS
Mental workload; multitasking; workload transitions; UAV

1. Introduction

Complex and dynamic environments – such as healthcare (Amin, 2011), driving (Morgan & Hancock, 2011), air traffic control (Edwards et al., 2012), and emergency response (Parush & Rustanjaja, 2013) – require operators to transition seamlessly between varying levels of workload. Workload transitions refer to a continuous period during which workload levels

CONTACT Sara Lu Riggs sriggs@virginia.edu University of Virginia, Charlottesville, VA, USA.
*Adjunct Professor of Psychology, Colorado State University; Professor Emeritus of Psychology and Aviation, University of Illinois, USA

change (Prytz & Scerbo, 2015). To date, most research has focused on the effects of low or high workload; however, transitions from low to high workload are more realistic and similar to what operators experience while on the job (Hooey et al., 2017; Huey & Wickens, 1993). For example, when the Apollo 12 spacecraft was struck by lightning, "all the [alarm] lights came on" the telemetry stream (Murray & Cox, 1989, pg 374), and there was a sudden, significant increase in workload. In the present study, we investigated the effects of gradual and sudden transitions from low to high workload in a multitasking environment, to better understand the effect of such transitions on performance in complex and dynamic domains.

Workload is a multi-dimensional construct for which there are several definitions in the literature (Van Acker et al., 2018). Wickens (1992) defined workload as the gap between the attentional resources of and mental demands placed on the user. Existing evidence has shown that workload at either extreme (i.e., low or high) can be problematic. For example, very low workload can lead to vigilance decrements, while very high workload can lead to fatigue, frustration, errors, and delays (Grier et al., 2003; Helton & Russell, 2012; Warm et al., 2008). In contrast, there is no consensus on the associated performance effects for workload transitions. When compared to constant workload levels, previous researchers have found that workload transitions can result in either performance costs (Bowers et al., 2014; Cox-Fuenzalida, 2007; Cumming & Croft, 1973; Goldberg & Stewart, 1980), improvements (Cox-Fuenzalida, 2007; Cumming & Croft, 1973; Goldberg & Stewart, 1980; Kim et al., 2019; Matthews, 1986; Prytz & Scerbo, 2015), or no differences (Bowers et al., 2014; Helton et al., 2008; Matthews, 1986; Voorheis et al., 2005). Given that the impact of workload transitions may be dynamic, further examination of transition effects on performance is needed (Cox-Fuenzalida, 2007; Jansen et al., 2016). For example, little is known on how the rate of a workload transition impacts multitasking performance. To our knowledge, only one study to date has compared gradual and sudden shifts from low to high workload, and found that gradual shifts led to a lower percentage of correct detections (Moroney et al., 1995). More work is needed to further understand these performance trends, especially for domains requiring multitasking.

In addition, comparisons of performance between workload transitions and constant workload can shed further light on the effects of workload transitions and prior theoretical explanations. Previous work has shown decrements in accuracy during the low and high phases of workload transitions, compared to when workload is held constant at low and high levels, respectively (Cox-Fuenzalida, 2007; Cox-Fuenzalida et al., 2006; Cox-Fuenzalida & Angie, 2005). Other studies have found this decrement is only apparent for constant low workload, but not constant high workload, (Matthews, 1986; Prytz & Scerbo, 2015). A separate body of research has shown no significant performance differences during the low and high phases of a workload transition compared to when workload is held constant at low and high (Helton et al., 2008; Moroney et al., 1993; 1995).

Several theoretical explanations have been presented to explain the findings associated with workload transitions in single-task environments, including expectancy effects (Cumming & Croft, 1973); strategic persistence (Matthews, 1986); and adaptation models (Cox-Fuenzalida, 2007). However, there exists no single explanation that can account for all findings. Cox-Fuenzalida (2007) found that models based on the adaptation process, such as state control theory (Hockey, 1986) and the dynamic model of stress adaptation (Hancock & Warm, 1989), were the most successful in explaining performance findings across studies. Two explanations have been primarily used to explain performance when transitioning between a dual- and single-task paradigm:

1. **Resource depletion** – after a transition, there will be a shortage of mental resources during the second phase of workload (i.e., the single task) and performance will suffer. However, resources will eventually recuperate during the second phase, due to the compensatory regeneration component, and performance will eventually rebound (Gluckman et al., 1993). In other words, there will be a decrement in performance immediately after a shift, but eventually resources will recuperate, and performance will improve.
2. **Effort regulation** – after a transition, people will need to actively evaluate the amount of mental resources necessary to perform successfully in the second workload phase. In other words, people can perform successfully as long as their appraisal is correct and workload does not reach levels of overload (Hockey, 1997).

Previous studies support either resource depletion (Cox-Fuenzalida & Angie, 2005; Gluckman et al., 1993; Moroney et al., 1995), effort regulation (Cox-Fuenzalida, 2007; Jansen et al., 2016; Matthews & Desmond, 2002), or both (Ungar et al., 2005). Ungar et al. (2005) investigated the validity of these explanations by comparing performance between dual- to single-task transition conditions and found that the validity of each explanation was dependent on primary task difficulty (i.e., the task in both the dual- and single-task paradigm). For example, resource depletion was supported with a more difficult primary task, as primary task performance was worse compared to its single-task baseline condition. However, effort regulation was supported with an easier primary task, as primary task performance was superior compared to the baseline.

Our study aimed to address some of the knowledge gaps in workload transition research. Understanding the effect of workload transitions is critical because operators in most complex environments – including military operations, aviation, healthcare, and driving – rarely experience constant levels of workload (Huey & Wickens, 1993). For example, military unmanned aerial vehicle (UAV) command and control operations may require operators to multitask during workload transitions, especially when managing multiple UAVs (Cummings et al., 2007; Winnefeld & Kendall, 2013).

There is a need to study factors that have been underexamined historically in the workload transition literature in the context of real-world domains. For example, there is a need to examine the effects of workload transitions relevant to complex and dynamic work environments (Cox-Fuenzalida & Angie, 2005; Huey & Wickens, 1993). The workload transition literature to date has examined multitasking in either highly-controlled laboratory settings (Cox-Fuenzalida & Angie, 2005), realistic testbeds (Jansen et al., 2016; Morgan & Hancock, 2011), or thoroughly validated multitasking environments (e.g., Air Force Multi-Attribute Task Battery; Bowers et al., 2014). Further exploration with different multitasking environments, especially as it relates to primary and secondary task performance (Cox-Fuenzalida & Angie, 2005; Jansen et al., 2016; Morgan & Hancock, 2011), is also of interest in this study. For example, operators in these environments are expected to continuously manage various interdependent tasks and responsibilities that fluctuate in their demand for operator attention. Depending on the situation, tasks that may be secondary to the mission may be critical to the overall viability of the UAV, so capturing this nuance in task prioritization is important (Clare et al., 2010). Previous researchers have suggested that findings from single-task studies may not hold true in dual-task or multitasking environments (Gluckman et al., 1993), so these environments warrant their own investigation. Understanding the effects of secondary tasks is important to expand the current workload transition literature, and from a systems safety standpoint given work environments are becoming increasingly complex.

Also, there is a need to understand other underexamined factors that may affect performance for both the workload transition literature and real-world occupational applications. For example, the number of workload transitions experienced by an operator, and the rate of workload transitions in these specific kinds of multitasking environments, are relevant to the domain, but have not been thoroughly explored. The majority of existing research has employed one or two workload transitions in a single-task environment (an exception is the work by Morgan & Hancock, 2011); however, findings from one or two workload transitions may not be representative of the multiple transitions operators may experience while on the job. Further, the rate of transition has been underexplored. Most research has focused on sudden transitions (i.e., immediate shifts between low and high workload), whereas there has been less reported on gradual transition (Matthews, 1986; Moroney et al., 1995).

There is also a need to understand the applicability and validity of the resource depletion and effort regulation explanations in multitasking domains that are not just transitioning between dual- and single-task paradigms (Cox-Fuenzalida, 2007). Existing evidence supports both explanations (Ungar et al., 2005), but most investigations have been limited to studying a dual- to single-task transition. There has been limited work that explores the applicability of these explanations in other types of multitasking environments (Bowers et al., 2014; Cox-Fuenzalida, 2007; Gluckman et al., 1993; Jansen et al., 2016).

Our aim in the current study was to address two main research gaps: the first as it pertains to occupational applications, and the second as it pertains to theoretical explanations. From these gaps, three specific research questions (RQ) emerged:

RQ I. How does performance compare between gradual versus sudden transitions from low to high workload? We expected that when workload transitions from low to high, sudden transitions will result in better primary and secondary task performance compared to gradual transitions (Moroney et al., 1995).

RQ II. How does performance compare between the low workload phases of the gradual and sudden workload transitions and constantly low workload? Previous work has shown that both primary and secondary task performance were better when workload was constantly low (Bowers et al., 2014; Cox-Fuenzalida, 2007; Cox-Fuenzalida & Angie, 2005; Matthews, 1986; Ungar et al., 2005; Wickens et al.,

Table 1. Summary of the research questions and expectations based on the theoretical explanations.

Research Question (RQ)	Resource Depletion Explanation	Effort Regulation Explanation
RQ II: Low workload (L) vs. low workload phases of gradual and sudden workload transitions (G and S, respectively)	If L outperforms S or G: • Resources depletion **IS** a function of workload transitions. Workload transitions deplete more mental resources than constant low workload (Bowers et al., 2014; Cox-Fuenzalida, 2007; Cox-Fuenzalida & Angie, 2005; Matthews, 1986; Ungar et al., 2005). If S or G is equal to or outperforms L: • Resources depletion **IS NOT** a function of workload transitions. Workload transitions do not deplete mental resources differently than low workload (Matthews, 1986).	n/a
RQ III: High workload (H) vs. high workload phases of gradual and sudden workload transitions (G and S, respectively)	n/a	If H is equal to or outperforms S or G • Effort regulation **IS NOT** a function of workload transitions. Workload transitions do not impact the effective regulation of mental resources (Gluckman et al., 1993). If S or G outperforms H • Effort regulation **IS** a function of workload transition. Workload transitions can help actively regulate mental resources under high workload (Hockey 1997; Matthews, 1986; Matthews & Desmond, 2002).

1985). Additionally, the resource depletion explanation predicts that both primary and secondary task performance will be better when workload is constantly low, compared to the low workload phases of gradual and sudden workload transitions, because resources will have not been depleted from any transitions during constant low workload.

RQ III. How does performance compare between the high workload phases of gradual and sudden workload transitions and constantly high workload? We expected that performance will be worse for both primary and secondary tasks when workload is constantly high (Bowers et al., 2014; Cox-Fuenzalida & Angie, 2005; Matthews, 1986; Prytz & Scerbo, 2015; Wickens et al., 1985). The effort regulation explanation predicts that primary task performance will be better during the high workload phases of the gradual and sudden workload transition than constant high workload, because the workload transitions will signal for the adjustment of resources accordingly.

Table 1 summarizes how each research question expands on each theoretical explanation. Of note, RQ I is not included because it does not provide further insights into the theoretical explanations.

The application of the current research includes informing the design of safer and more effective systems in complex and dynamic domains. The domain selected for this study was unmanned aerial vehicle (UAV) command and control. Previous research in this domain has examined the effects of workload in general, but there has been limited work investigating the effects of workload transitions (Cummings et al., 2007; Gateau et al., 2016; Jasper et al., 2016). Workload transitions are of importance given the goal of the U.S. Department of Defense to have a single operator handling all mission, flight, and sensor management tasks for multiple UAVs (Winnefeld & Kendall, 2013). This goal presents a major challenge due to: (1) the limited mental resources of human operators, and (2) the unpredictability of UAV missions, largely due to workload shifts caused by monitoring various sources of data (Cummings et al., 2019; Sibley et al., 2015).

2. Method

2.1. Participants

This study used the same methodology and experimental testbed from Moacdieh et al. (2020), but focused on the performance results. Twenty-one students participated in this study [13 males and 8 females; mean (SD) age = 20.9 (1.5) years]. Participants self-reported normal or corrected-to-normal vision and were compensated $10/hour. The study was approved by the Clemson University Institutional Review Board (IRB2015-217) and all participants provided informed consent.

2.2. Experimental Setup

The testbed was developed using the Unity game development platform and was based on the 'Vigilant Spirit Control Station' (VSCS) used by the Air Force to develop interfaces to control multiple UAVs (Feitshans et al., 2008). The VSCS platform has been used to test how interface design could aid UAV operators. Tasks in the testbed were typical of an

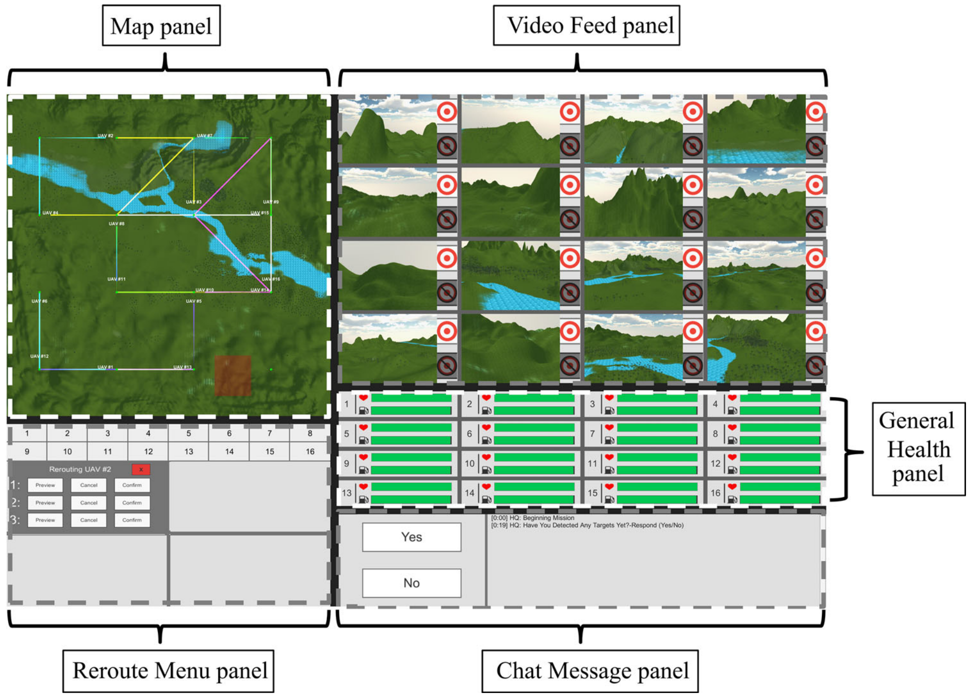


Figure 1. Screenshot of the UAV testbed with panels labeled.

UAV command and control environment, such as target detection and route planning (Feitshans et al., 2008), which require operators to employ perceptual, cognitive, and motor resources. The testbed ran on a desktop computer with an 81 cm monitor (2560 × 1440 resolution) and a standard mouse.

2.3. UAV Command and Control Testbed and Tasks

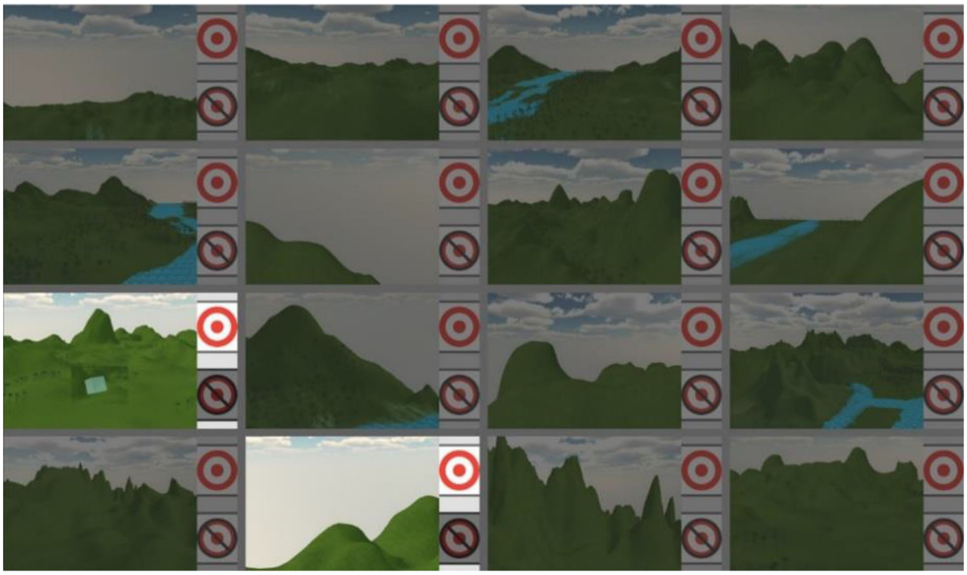
Participants were responsible for controlling and managing UAVs under four 15-minute scenarios in the UAV testbed (see Figure 1). There were four tasks in each scenario, one primary task and three secondary tasks, and each will be discussed in turn. For each scenario, the rate at which the primary task occurred varied (see section 2.5), while one of the three secondary tasks occurred every 20 s, on average, in a pseudo-random order.

2.3.1. Target Detection Task (Primary Task)

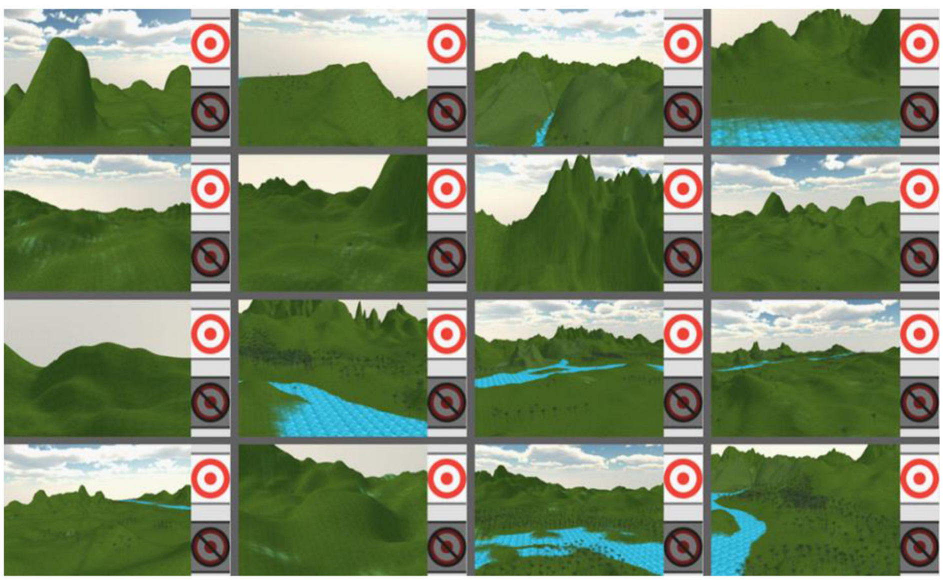
Participants were instructed that this task had the highest priority. They were tasked to monitor up to 16 UAV video feeds on the Video Feed panel for a target, presented as a semi-transparent cube (see Figure 2). Targets could only be detected when a UAV video feed was active (highlighted). When a UAV video feed was active and a target was present, participants were instructed to press the "target" button. Otherwise, they were instructed to leave the default "no target" button selected. The "no target" button was the default, as pilot testing suggested this was necessary in order to assess the participant's search abilities and not their ability to click quickly. UAV video feeds were active for 10 s, and targets could appear during this time. Video feeds cycled between active and inactive throughout the scenario. If a target was present in an active UAV, but the participant did not select the target button within the 10 s, the participant missed the opportunity to detect that specific target. On average, 20% of active UAVs detected a target. The number of simultaneously active UAVs determined the workload level (see section 2.5).

2.3.2. Reroute Task (Secondary Task #1)

Participants were tasked to reroute a UAV when it was projected to enter a no-fly-zone (i.e., red square on the Map panel in Figure 3). If a UAV was projected to enter the no-fly-zone, its route and label would turn orange and participants had 15–20 s to reroute the UAV. To reroute a UAV, a participant clicked on the respective UAV's numbered square in the Reroute Menu panel and chose from one of three new routes. For each new route, participants could select "Preview" to see it, "Confirm" to reroute the UAV to the specific new route, or "Cancel" to exit from the window. There was no limit to how many times a UAV could be rerouted. If a UAV was not



(a)



(b)

Figure 2. Screenshots of the Video Feed panel where any of the 16 video feeds could be active (i.e., video feed is highlighted). (a) UAVs 9 and 14 are active (numbering from left to right, top to bottom); all other UAVs are inactive; UAV 9 has a target (semi-transparent cube). (b) All 16 UAVs are active, but none of them are detecting a target.

rerouted to avoid the no-fly-zone, it became non-operational for the remainder of the scenario. The rerouting task occurred 18 times in each scenario.

2.3.3. Fuel Leak Task (Secondary Task #2)

Participants were also tasked to monitor for fuel leaks using the General Health panel (Figure 4). When a fuel leak occurred, the color of the health status bar (top bar denoted with a heart) changed from green to yellow with a "FIX LEAK" warning. Participants then had 10 s to click on the bar; otherwise, it would change from yellow to orange and read "FATAL FUEL LEAK" for that specific UAV. A fuel leak occurred 14 times in each scenario.

2.3.4. Chat Message Task (Secondary Task #3)

Participants were tasked with responding to chat messages by selecting between the two options on the left-hand side of the chat message panel (Figure 1). Responding to chat messages consisted of

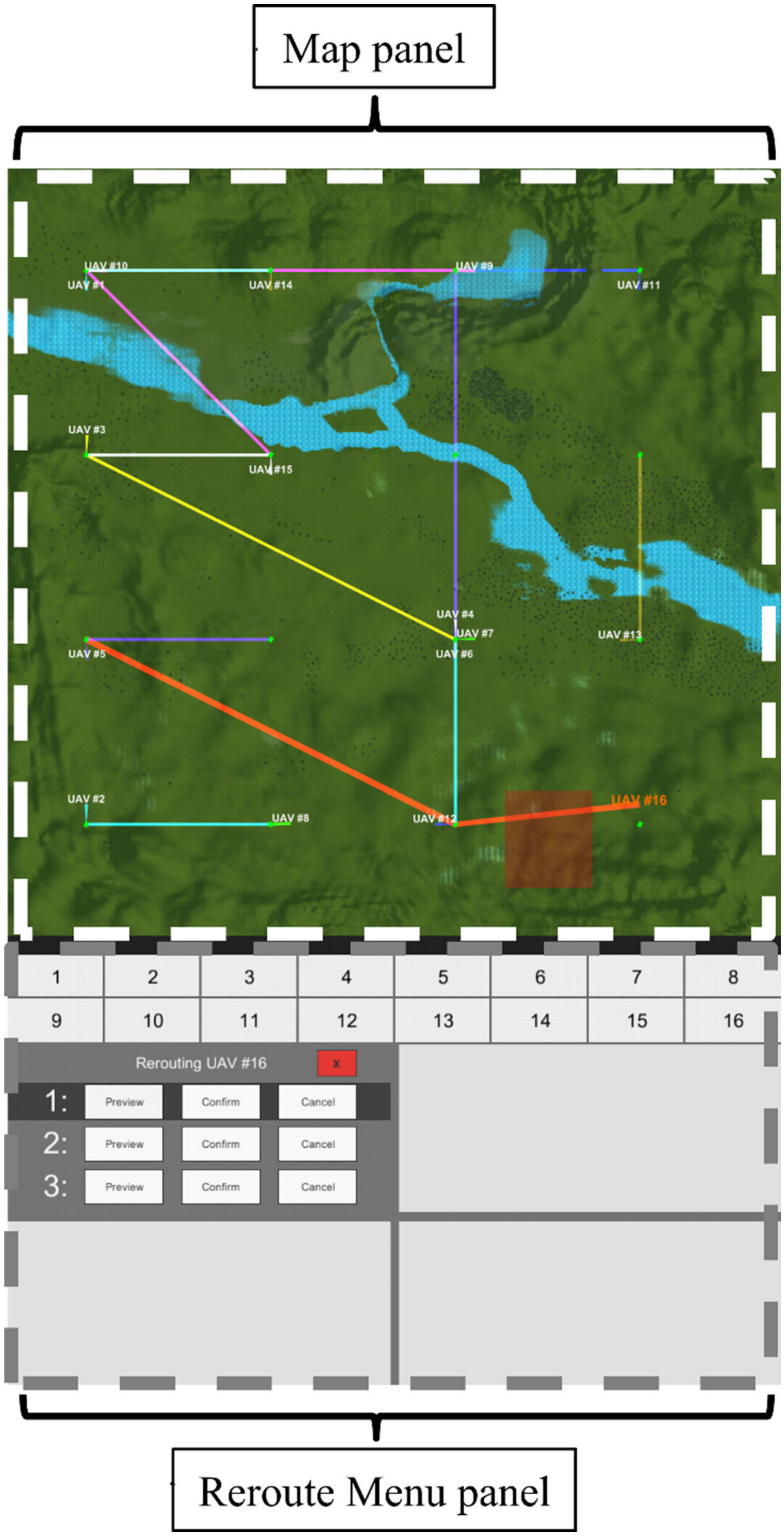


Figure 3. Screenshot of the Map panel (top half) and Reroute Menu panel (bottom half). After clicking on the respective UAV number (buttons numbered 1–16), a menu of route options was presented (dotted area with options 1–3 where "Preview," "Confirm," or "Cancel" could be selected).

selecting from one of two options (e.g., selecting 'yes' or 'no' to yes/no questions). Participants could respond to a chat message until another message appeared and were instructed to accurately answer as many questions as possible. There were 19 chat messages in each scenario.

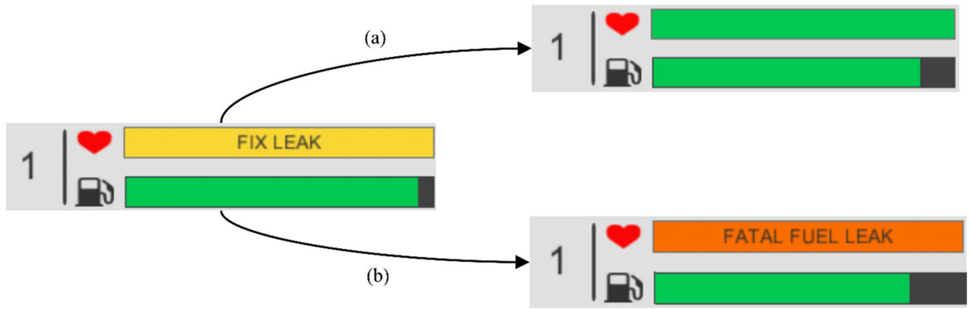


Figure 4. Screenshot of the health status bar for UAV #1. Participants were tasked to press the health status bar when a fuel leak occurred. (a) When a fuel leak was fixed in time, the health status bar changed from yellow to green and the "FIX LEAK" warning disappeared. (b) When a fuel leak was not fixed in time, the health status bar changed to an orange "FATAL FUEL LEAK".

Table 2. Point system for the UAV command and control testbed.

Task	Points per Response
Correctly recognizing a target	+100
Correctly recognizing a non-target	+50
All secondary tasks (reroute, fuel leak, and chat message)	+30
Any incorrect or lack of response (false positive or negative to target detection task, UAV flies through no-fly-zone, or "FATAL FUEL LEAK" condition)	−100

2.4. Point System and Dependent Measures

Table 2 shows the point system implemented to encourage participants to prioritize the primary task and avoid task shedding. This point system reinforced the need to prioritize searching for targets in the Video Feed panel, as successfully detecting a target earned the most points. Losing a UAV in the no-fly-zone not only resulted in an immediate loss of points, but also the loss of the opportunity to gain points from that UAV's target detection task. Whenever the UAV became inoperable, the corresponding UAV video feed became inactive for the remainder for the scenario. The highest scoring participant also earned a bonus $10 gift card.

Response time for the primary target detection task was calculated from the appearance of the target to when participants clicked the "Target" button. For all secondary tasks, response times were calculated from the onset of the event to when the participant responded. Accuracy was calculated as the percentage of correct responses within the time limit for each task.

2.5. Workload Scenarios

Workload was manipulated by varying the number of active UAVs (i.e., highlighted video feeds) in the target detection task. This approach is consistent with previous studies, where workload was manipulated by directly manipulating the task load of the operator (e.g., Hancock, Williams, Manning, & Miyake, 1995). Additionally, this approach was considered appropriate given that the long-term goal of UAV command and control is to increase the load per operator (Winnefeld & Kendall, 2013). The four workload scenarios used in this study are as follows:

1. *Low workload scenario*. There were 3–5 UAVs active for the entirety of the scenario.
2. *High workload scenario*. There were 13–16 UAVs active for the entirety of the scenario.
3. *Gradual transition scenario*. The number of active UAVs increased gradually. The scenario started at low workload for 20 s, and then one to three active UAVs were added every 10 s until high workload was reached (13–16 active UAVs). The scenario would remain at high workload for two minutes, before immediately returning to low workload. This cycle repeated five times for this scenario. The solid black line in Figure 5 depicts the theoretical number of active UAVs over the course of the gradual transition scenario.
4. *Sudden transition scenario*. The number of active UAVs increased instantaneously. One minute of low workload (3–5 UAVs) was followed by an instantaneous increase to high workload (13–16 UAVs) that lasted for two minutes. After the two minutes of high workload, there was an immediate return to low workload. This cycle repeated five times for this scenario. The dotted gray line in Figure 5 depicts the theoretical number of active UAVs over the course of the sudden transition scenario.

Workload level thresholds (i.e., low and high) were based on pilot testing data using both performance and NASA-TLX measures (Hart & Staveland, 1988). Mean target detection task

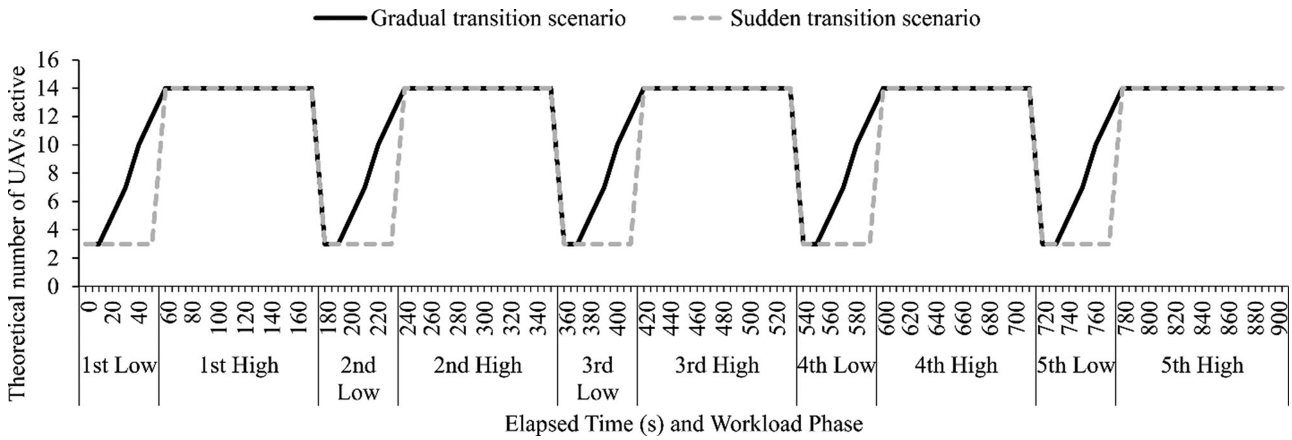


Figure 5. The theoretical number of active UAVs throughout the gradual and sudden transition scenarios. The horizontal axis denotes workload phase (e.g., 1st low, 1st high, etc.).

(primary task) accuracy was approximately 30% higher in low workload compared to high workload. NASA-TLX dimensions of interest included mental demand, temporal demand, and performance. Significant differences between low and high workload were found for all dimensions ($p<.05$; analysis was done with a Friedman test and pairwise comparisons were performed using Mann-Whitney tests). A range of UAVs was used for the low and high workload thresholds because it was possible for the participant to lose a UAV by not rerouting it in time. As such, the range of UAVs allowed the requisite workload level throughout the scenario to be maintained. There was never a situation where the intended workload was not reached due to the loss of UAVs. Of note, the gradual and sudden transition scenarios only included transitions from low to high workload, as this transition direction is emblematic of situations likely to occur in data-rich, dynamic domains (e.g., Apollo 12, Murray & Cox, 1989) and was therefore the focus of our work. Also, the high workload periods within the gradual and sudden transition scenarios were longer than the low workload periods as extended periods of high workload is another challenge of data-rich, dynamic domains (e.g., UAVs command and control; Arrabito et al., 2010; Williams, 2006). In order to have the gradual and sudden transition scenarios include the same number of transitions from low to high workload, the time spent in low workload varied slightly between the two scenarios. Although this presents a limitation of this work, the subsequent analyses and discussion are unaffected, since our overarching goal was to examine how the rate of workload transitions affects performance during the low and high workload periods.

2.6. Procedures

The experiment took place over two consecutive days at approximately the same time of day. On the first day, participants gave consent and were briefed about the study goals and expectations. Participants then completed a five-minute training session where 4–6 UAVs were always active. By the end of the training session, participants had to demonstrate proficiency by having a minimum accuracy of 70% across all tasks. If not, they could ask questions and reattempt the training session; otherwise, they were excused from the study. Only four participants completed the training session twice and each succeeded on their second attempt. Participants then completed the two constant workload scenarios (low and high workload scenarios) in a randomized order (i.e., 10 participants completed the low workload then high workload scenario and vice versa for the other 11 participants). This approach allowed for the constant workload scenarios to serve as baseline comparisons. On the second day, participants completed the two transition scenarios (gradual and sudden transition scenarios) in a randomized order (i.e., 11 participants completed the gradual then sudden transition scenario and vice versa for the other 10). The study lasted about 2.5 h over two days, with the first day lasting 1.5 h and the second day lasting 1 h.

2.7. Analysis

Three analyses were performed to address each research question. A sub-goal of all three research questions was to compare primary and secondary task performance. For all analyses, response time and accuracy of the primary task (target detection) was compared to the aggregated response time and

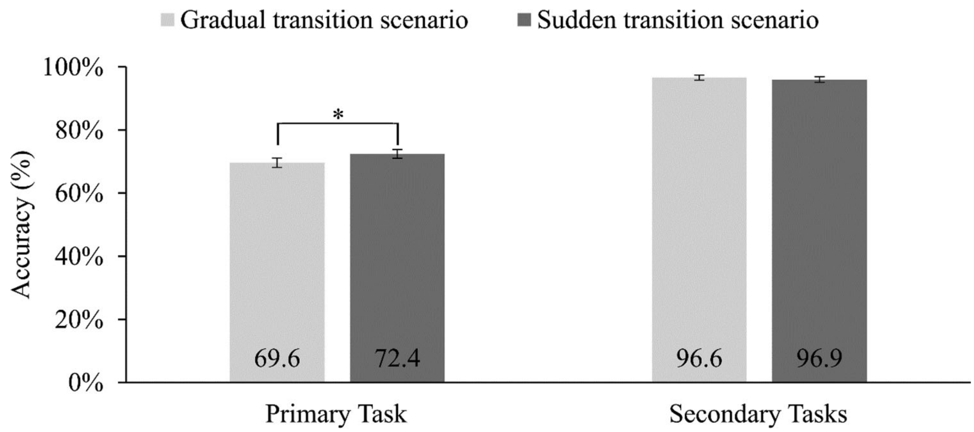


Figure 6. Mean accuracy for both task types for the gradual and sudden transition scenario. Asterisks (*) denote significant differences between scenarios.

accuracy of all the secondary tasks (reroute, fuel leak, and chat message). Analysis for RQ I compared performance between the different workload transition scenarios (i.e., gradual and sudden transition scenarios). Specifically, response time and accuracy were analyzed using separate 2 × 2 repeated-measures analyses of variance (ANOVA: two transition scenarios, two task types). Analysis for RQ II compared performance in the low workload scenario to the low workload phases of the gradual (the five 20 s phases aggregated) and sudden (the five 60 s phases aggregated) workload scenarios. Analyses for RQ III compared the high workload scenario to the high workload phases of the gradual and sudden transition scenario (the five 120 s phases aggregated, respectively). Response time and accuracy for RQ II and III were analyzed using separate 3 × 2 repeated-measures ANOVAs (three scenarios, two task types). Bonferroni corrected, Fisher's protected LSD post-hoc tests were performed to test differences between means. Significance was set at α=.05. Before completing the analyses associated with each research question, 2 × 2 × 2 (two order types, two transition scenarios, and two task types) ANOVAs were completed to assess whether the scenario order affected primary and secondary task performance. There were no significant main effects of order (p>.05), so analyses for RQ I, II, and III proceeded.

Violations of normality were assessed prior to analysis, and Greenhouse-Geisser corrections were used when sphericity was violated (using Mauchly's test). SPSS 24 was used for all analyses. Partial eta-squared (η_p^2) is reported as a measure of effect size for the omnibus test, and values of .01, .06, and .14 are interpreted as small, medium, and large effect sizes, respectively (Cohen, 1988). For all post-hoc pairwise comparisons, effect sizes were measured using Cohen's d for repeated measures (Lakens, 2013), and values of .2, .5, and .8 are interpreted as small, medium, and large effect sizes (Cohen, 1988). For all graphs, error bars indicate the standard error of the mean, curly brackets indicate groupings of scenarios, straight brackets indicate a specific comparison of scenarios, and asterisks denote significant differences between scenarios.

3. Results

3.1. RQ I: Gradual vs. sudden Workload Transitions

3.1.1. Response Time

When comparing gradual to sudden transitions, there was no main effect of transition type (F(1,20)= 1.497, p=.235, η_p^2 =.070), but there was a main effect of task type (F(1,20)=110.400, p<.001, η_p^2=.847). The mean response time for the primary task (M = 2.85 s, SE=.026 s) was significantly faster than the secondary tasks (M = 3.87 s, SE =.098 s; p<.001, d = 2.908). There was no transition scenario × task type interaction effect (F(1,20)=1.581, p=.223, η_p^2 =.073).

3.1.2. Accuracy

There was no main effect of scenario (F(1,20)=1.552, p=.227, η_p^2=.220), but there was a significant main effect of task type (F(1,20)=579.583, p<.001, η_p^2=.967). There was a significant scenario × task type interaction effect (F(1,20)=6.056, p=.023, η_p^2=.232). For the primary task, the gradual transition scenario had significantly worse accuracy than the sudden transition scenario (p=.048, d=.421), but this was not true for the secondary task (Figure 6).

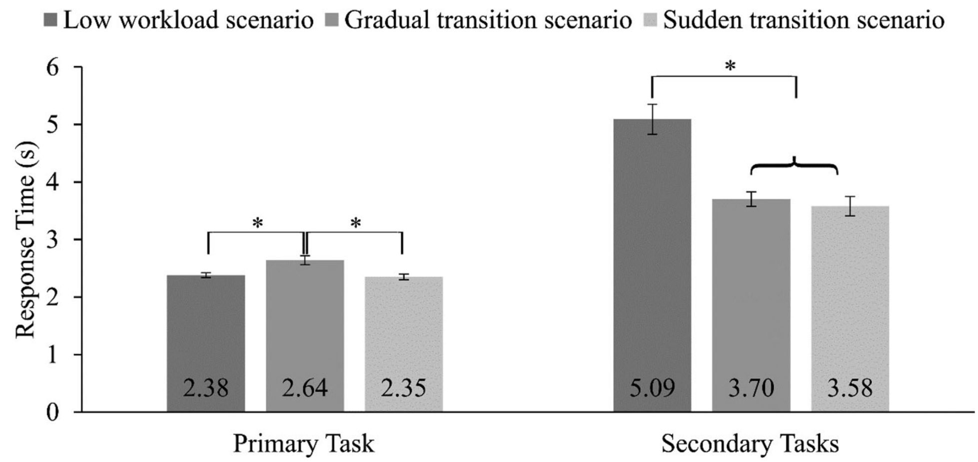


Figure 7. Mean response times for both task types during low workload phases. Asterisks (*) denote significant differences between scenarios.

3.2. RQ II: Performance during Low Workload Phases

3.2.1. Response Time

There was a significant effect of scenario ($F(2,40)=20.968$, $p<.001$, $\eta_p^2=.512$) and task type ($F(1,20)=155.535$, $p<.001$, $\eta_p^2=.886$), as well as a significant scenario × task type interaction ($F(2,40)=32.238$, $p<.001$, $\eta_p^2=.617$). For the primary task, response time for the low workload phases in the gradual transition scenario was significantly slower than the low workload scenario ($p<.009$, $d=.844$) and the low workload phases in the sudden scenario ($p<.002$, $d=.917$), whereas the latter two did not differ from each other. For the secondary tasks, response time was significantly slower for the low workload scenario than for the low workload phases in the gradual ($p<.001$, $d=1.445$) and sudden transition scenario ($p<.001$, $d=1.509$) as seen in Figure 7.

3.2.2. Accuracy

There was a significant effect of scenario ($F(1.369,27.384)=32.363$, $p<.001$, $\eta_p^2=.618$). Across both tasks, mean accuracy for all scenarios were significantly different from each other. Accuracy in the low workload scenario ($M=85.8\%$, $SE=1.5\%$) was significantly lower than the low workload phases in the gradual ($M=92.7\%$, $SE=1.1\%$; $p=.001$, $d=1.138$) and sudden transition scenario ($M=95.8\%$, $SE=0.7\%$; $p<.001$, $d=1.591$) and accuracy for the low workload phases in the gradual transition scenario was significantly lower than the low workload phases in the sudden transition scenario ($p=.022$, $d=.700$). There was also a significant effect of task type ($F(1,20)=77.774$, $p<.001$, $\eta_p^2=.795$), with the mean accuracy for the primary task ($M=87.8\%$, $SE=0.9\%$) being significantly lower than the secondary task ($M=95.1\%$, $SE=1.0\%$; $p<.001$, $d=1.671$). There was no significant scenario × task type interaction on accuracy ($F(1.476,29.514)=3.22$, $p=.068$, $\eta_p^2=.139$).

3.3. RQ III: Performance during High Workload Phases

3.3.1. Response Time

There was a significant effect of scenario($F(2,40)=38.882$, $p<.001$, $\eta_p^2=.660$), task type ($F(1,20)=109.988$, $p<.001$, $\eta_p^2=.846$), as well as a significant scenario × task type interaction ($F(2,40)=14.253$, $p<.001$, $\eta_p^2=.416$). Primary task response time for the high workload scenario was significantly slower than the high workload phases of the gradual ($p<.001$, $d=1.549$) and sudden transition scenarios ($p<.001$, $d=1.553$). For the secondary tasks, response time for the high workload scenario was slower than the high workload phases in the gradual ($p<.001$, $d=1.247$) and sudden transition scenario ($p<.001$, $d=1.490$) as seen in Figure 8.

3.3.2. Accuracy

There was a significant effect of scenario ($F(2,40)=44.028$, $p<.001$, $\eta_p^2=.688$). Across both tasks, mean accuracy for the high workload scenario ($M=73\%$, $SE=1.3\%$) was significantly worse than the high workload phases in the gradual ($M=82.3\%$, $SE=1.0\%$; $p<.001$, $d=1.507$) and sudden transition scenario ($M=81.6\%$, $SE=1.0\%$; $p<.001$, $d=1.459$). There was also a significant effect of task type ($F(1,20)=484.172$, $p<.001$, $\eta_p^2=.960$). Mean accuracy for the primary task ($M=63.7\%$, $SE=1.3\%$) was significantly lower than the secondary task ($M=94.1\%$, $SE=0.9\%$; $p<.001$, $d=5.777$). There was no significant scenario × task type interaction on accuracy ($F(2,40)=3.22$, $p=.200$, $\eta_p^2=.077$).

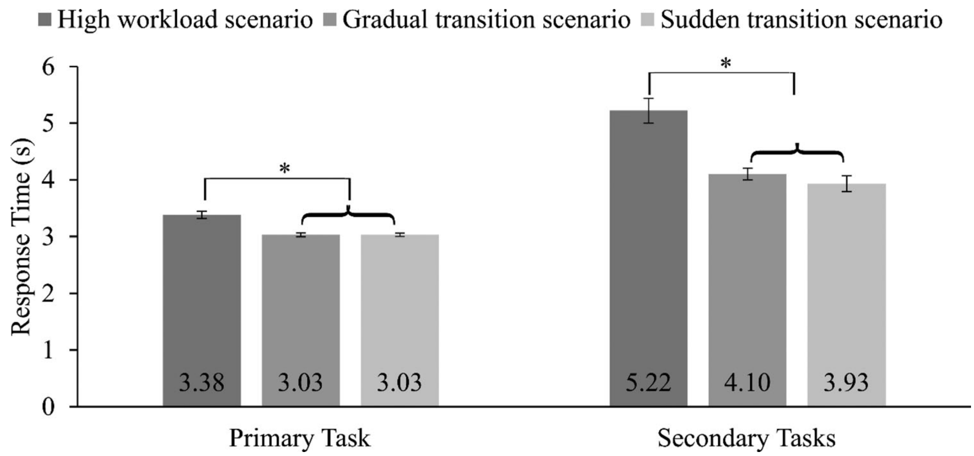


Figure 8. Mean response times for both task types during high workload phases. Asterisks (*) denote significant differences between scenarios.

4. Discussion

Our goal was to better understand how workload transitions influence performance in a multitasking, real-world domain, and to assess the applicability of the resource depletion and effort regulation explanations. With respect to RQ I, the main difference between gradual and sudden workload transitions was that sudden transitions had higher primary task accuracy compared to gradual transitions (Figure 6); however, the associated effect size was small. There were no other performance differences between the two workload transition types, and this was somewhat expected based on previous work. The findings here suggest that multitasking—the main experimental difference between this line of work and the single-task environment used in Moroney et al. (1995)—may affect workload transition performance differently. Although addressing RQ I provides further understanding on workload transitions, it does not elucidate the applicability of either the resource depletion or effort regulation explanations. To further understand the applicability of these explanations, comparing the low and high workload phases of gradual and sudden workload transitions to constant workload is needed.

Regarding RQ II, the majority of findings indicated faster response times and higher accuracy rates during the low workload phases of both the gradual and sudden transition scenarios compared to the low workload scenario, which is consistent with some prior work (e.g., Jansen et al., 2016; Krulewitz et al., 1975; Figure 7). The only instance when this was not the case was primary task response time. Specifically, low workload phases of the gradual transition scenario resulted in longer response times compared to the low workload phases of the sudden transition scenario and constant low workload scenario. Our results are consistent with previous work that found performance improves during low workload phases of workload transitions in the presence of a secondary task (Jansen et al., 2016; Matthews & Desmond, 2002). Our findings also show that fluctuations in workload result in superior performance than when workload is held at a constant low level. In sum, our findings do not support the resource depletion explanation, which predicts that low workload phases in workload transitions will result in worse performance, because resources are depleted during high workload phases. It appears that different kinds of multitasking environments may not cause operators to experience resource depletion during a workload transition in the same way (Gluckman et al., 1993). There needs to be more work to examine the impacts of specific contextual factors, as our work suggests that diversified task demands can help thwart resource depletion effects in complex work environments.

For RQ III, we found that primary and secondary task performance was better during the high workload phases of the gradual and sudden transition scenarios, compared to the high workload scenario (Figure 8). However, this improvement in performance is in stark contrast with some other work that has found performance decrements during high workload phases after workload transitions (Cox-Fuenzalida, 2007; Cox-Fuenzalida & Angie, 2005; Gluckman et al., 1993; Krulewitz et al., 1975; Moroney et al., 1995). Previous work, however, did not include secondary tasks and/or multiple workload transitions. The results here suggest that the effort regulation explanation, which suggests that participants can effectively redistribute their mental resources as task demands change, may be dependent not only on the presence of a workload transition, but also multiple workload transitions—the main difference that sets apart this study from previous ones. Further research, especially studies that modulate the number of workload transitions, is needed to corroborate this finding and would shed further understanding on how to improve performance during high workload (Grier et al., 2008).

Overall, the findings here demonstrate that existing theoretical explanations cannot fully explain the effects of workload transitions on performance. Instead, our work further supports the nuanced nature of workload transitions, as we explored contexts that have been underexamined to-date. For instance, existing explanations do not distinguish between primary and secondary task performance. Our results showed this to be an important consideration, since secondary task performance improved with both types of workload transitions, but this was not always the case for the primary task. This finding highlights the importance of reconsidering the secondary task(s) in complex domains, because historically it has been assumed that primary task performance would be prioritized over secondary task performance (Wickens et al., 2015). In reality, task prioritization may vary depending on context and environment, especially in dynamic environments such as UAV operations where operators are juggling multiple, interrelated tasks and responsibilities.

Our work demonstrates this notion empirically, as participants attended successfully to secondary tasks, even if at times it led to an immediate cost to primary task performance. This satisficing approach supports previous work that has shown that participants may change how they prioritize tasks as they recognize task interdependencies can change over time (Jansen et al., 2016). Therefore, it is important for future work to investigate workload transitions with interdependent, dynamic tasks to fully understand: (a) the effect of workload transitions on performance; (b) the applicability of existing theoretical explanations; and (c) contextual factors from different occupational environments. Overall, our results suggest that occupational factors in these complex and dynamic domains impact workload transition performance in different ways than expected and the findings can be used to inform design. For example, it may be beneficial to constantly, but strategically, engage and reengage operators in UAV command and control with diverse tasks so they can manage mental resources more effectively over time. Such a strategy may help negate the effects due to vigilance decrements during constant low workload and data overload during constant high workload. As a result, performance may improve because mental resources could be more effectively employed. This strategy could be taken into consideration occupational and technology design.

Our results show that the impact of workload transitions is nuanced. One limitation of the current work is that only two workload transition rates were examined. Transition rate was not thoroughly explored, because it would have affected either the number of transitions per scenario if scenario length was held constant, or the lengths of each scenario if the number of transitions was held constant. Although minimal differences were found between the workload transitions used here, it may be of interest to consider different rates in order to fully understand its effects. While outside of the scope of this study, scenario duration is another limitation worth mentioning as there is evidence that workload transitions may have longer-term ramifications (Cox-Fuenzalida, 2007). Another potential limitation was that low and high workload thresholds were set equivalent across participants. As a result, high workload may have been deemed to be more difficult for some participants than others (Prytz & Scerbo, 2015). Future work could individually tailor workload thresholds to account for individual differences, however such an experimental design comes with its own challenges (see Bowers et al., 2014). Additionally, future work should investigate whether expectancy of workload transitions affects performance. Finally, there is also a need to understand other occupationally-relevant effects, such as task type (e.g., manual vs. verbal tasks) and individual differences, (e.g., expertise and personality; Cox-Fuenzalida et al., 2006; Cox-Fuenzalida et al., 2004).

In summary, our results showed minimal differences in performance between gradual and sudden workload transitions, but transitions in general resulted in faster responses and higher accuracy compared to when workload was held constant at low and high workload levels. Our findings provide further support for the effort regulation explanation (Hockey, 1997); however, future work should investigate the applicability of existing theoretical explanations as they relate to different occupational factors and settings. Although our findings have implications on the design of systems for operators in various complex domains, future work needs to address how to best integrate them to ensure operators can safely cope with workload transitions, such as other aspects of multi-UAV (Cummings et al., 2019; Ramchurn et al., 2015).

Acknowledgments

The authors would like to thank Aakash Bhagat for his help in developing the testbed used in this study.

Conflicts of Interest

There are no conflicts of interest to report.

Funding

This study was supported in part by the National Science Foundation (NSF grant: #1566346, Program Manager: Dr. Ephraim Glinert; NSF grant: #1750850, Program Manager: Dr. Andrew Kerne).

ORCID

Shannon Patricia Devlin http://orcid.org/0000-0003-2652-012X
Nadine Marie Moacdieh http://orcid.org/0000-0002-7677-7946
Christopher D. Wickens http://orcid.org/0000-0002-6789-4446
Sara Lu Riggs http://orcid.org/0000-0002-0112-9469

References

Amin, S. G. (2011). *A study to determine the influence of workload on nursing personnel.* (Western Michigan University). https://doi.org/10.1155/2014/167674

Arrabito, G. R., Ho, G., Lambert, A., Rutley, M., Keillor, J., Chiu, A., … Hou, M. (2010). *Human factors issues for controlling uninhabited aerial vehicles: Preliminary findings in support.* www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA543186

Bowers, M. A., Christensen, J. C., & Eggemeier, F. T. (2014). The effects of workload transitions in a multitasking environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 220–224. https://doi.org/10.1177/1541931214581046

Clare, A. S., Hart, C. S., & Cummings, M. L. (2010). *Assessing operator workload and performance in expeditionary multiple unmanned vehicle control[Paper presentation]*.48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, (January), 1–10. http://www.scopus.com/inward/record.url?eid=2-s2.0-78649888643&partnerID=tZOtx3y1 https://doi.org/10.2514/6.2010-763

Cohen, J. (1988). Analysis of variance. In *Statistical power analysis for the behavioral sciences* (pp. 273–406). Lawrence Erlbaum Associates.

Cox-Fuenzalida, L.-E. (2007). Effect of workload history on task performance. *Human Factors*, *49*(2), 277–291. https://doi.org/10.1518/001872007X312496

Cox-Fuenzalida, L.-E., & Angie, A. D. (2005). The effects of workload history on dual task performance. *Current Psychology*, *24*(3), 171–179. https://doi.org/10.1007/s12144-005-1020-y

Cox-Fuenzalida, L.-E., Angie, A., Holloway, S., & Sohl, L. (2006). Extraversion and task performance: A fresh look through the workload history lens. *Journal of Research in Personality*, *40*(4), 432–439. https://doi.org/10.1016/j.jrp.2005.02.003

Cox-Fuenzalida, L.-E., Beeler, C., & Sohl, L. (2006). Workload history effects: A comparison of sudden increases and decreases on performance. *Current Psychology*, *25*(1), 8–14. https://doi.org/10.1007/s12144-006-1012-6

Cox-Fuenzalida, L.-E., Swickert, R., & Hittner, J. B. (2004). Effects of neuroticism and workload history on performance. *Personality and Individual Differences*, *36*(2), 447–456. https://doi.org/10.1016/S0191-8869(03)00108-9

Cumming, R. W., & Croft, P. G. (1973). Human information processing under varying task demand. *Ergonomics*, *16*(5), 581–586. https://doi.org/10.1080/00140137308924548

Cummings, M. L., Huang, L., Zhu, H., Finkelstein, D., & Wei, R. (2019). The impact of increasing autonomy on training requirements in a UAV supervisory control task. *Journal of Cognitive Engineering and Decision Making*, *13*(4), 295–309. https://doi.org/10.1177/1555343419868917

Cummings, M. L., Nehme, C. E., Crandall, J., & Mitchell, P. (2007). Predicting operator capacity for supervisory control of multiple UAVs. In J. S. Chahl, L. C. Jain, A. Mizutani, & M. Sato-Ilic (Eds.), *Innovations for Intelligent Machines - 1. Studies in Computational Intelligence* (Vol. 70, pp. 11–37). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72696-8_2

Edwards, T., Sharples, S., Wilson, J. R., & Kirwan, B. (2012). Factor interaction influences on human performance in air traffic control: The need for a multifactorial model. *Work*, *41*, 159–166. https://doi.org/10.3233/WOR-2012-0151-159

Feitshans, G., Rowe, A., Davis, J., Holland, M., & Berger, L. (2008). *Vigilant spirit control station (VSCS) the face of counter[Paper presentation].* AIAA Guidance, Navigation and Control Conference and Exhibit. https://doi.org/10.2514/6.2008-6309

Gateau, T., Chanel, C. P. C., Le, M. H., & Dehais, F. (2016). *Considering human's non-deterministic behavior and his availability state when designing a collaborative human-robots system [Paper presentation].* IEEE International Conference on Intelligent Robots and Systems, 4391–4397. https://doi.org/10.1109/IROS.2016.7759646

Gluckman, J. P., Warm, J. S., Dember, W. N., & Rosa, R. R. (1993). Demand transitions and sustained attention. *The Journal of General Psychology*, *120*(3), 323–337. https://doi.org/10.1080/00221309.1993.9711151

Goldberg, R. A., & Stewart, M. R. (1980). Memory overload or expectancy effect? 'Hysteresis' revisited*. *Ergonomics*, *23*(12), 1173–1178. https://doi.org/10.1080/00140138008924824

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The vigilance decrement reflects limitations in effortful attention, not mindlessness. *Human Factors*, *45*(3), 349–359. https://doi.org/10.1518/hfes.45.3.349.27253

Grier, R. A., Wickens, C. D., Kaber, D., Strayer, D., Boehm-Davis, D., Trafton, J. G., … Wickens, C. D. (2008). *The red-line of workload: Theory, research, and design[Paper presentation].* Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting, New York, NY., 1204–1208 https://doi.org/10.1177/154193120805201811

Hancock, P. A., Williams, G., Manning, C. M., & Miyake, S. (1995). Influence of task demand characteristics on workload and performance. *The International Journal of Aviation Psychology*, *5*(1), 63–86.

Hancock, P. A., & Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, *31*(5), 519–537. https://doi.org/10.7771/2327-2937.1024 https://doi.org/10.1177/001872088903100503

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Helton, W. S., & Russell, P. N. (2012). Brief mental breaks and content-free cues may not keep you focused. *Experimental Brain Research*, *219*(1), 37–46. https://doi.org/10.1007/s00221-012-3065-0

Helton, W. S., Shaw, T., Warm, J. S., Matthews, G., & Hancock, P. (2008). Effects of warned and unwarned demand transitions on vigilance performance and stress. *Anxiety, Stress, and Coping*, *21*(2), 173–184. https://doi.org/10.1080/10615800801911305

Hockey, G. R. J. (1986). A state control theory of adaptation to stress and individual difference is stress management. In G. R. J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (Eds.), *Energetics and human information processing* (pp. 285–298). Martinus Nijhoff Publishers.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive energetical framework. *Biological Psychology*, *45*, 73–93. https://doi.org/10.1016/S0301-0511(96)05223-4

Hooey, B. L., Kaber, D. B., Adams, J. A., Fong, T. W., & Gore, B. F. (2017). The underpinnings of workload in unmanned vehicle systems. *IEEE Transactions of Human-Machine Systems*, *48*(5), 1–16. https://doi.org/10.1109/TCST.2007.894653

Huey, B. M., & Wickens, C. D. (Eds.). (1993). *Workload transition: Implications for individual and team performance*. National Academy Press.

Jansen, R. J., Sawyer, B. D., van Egmond, R., de Ridder, H., & Hancock, P. A. (2016). Hysteresis in mental workload and task performance: The influence of demand transitions and task prioritization. *Human Factors*, *58*(8), 1143–1157. https://doi.org/10.1177/0018720816669271

Jasper, P., Sibley, C., & Coyne, J. (2016). Using heart rate variability to assess operator mental workload in a command and control simulation of multiple unmanned aerial vehicles. *Proceedings of the Human Factors and Ergonomics Society*, 1124–1128. https://doi.org/10.1177/1541931213601264

Kim, N. Y., House, R., Yun, M. H., Nam, C. S., & Harrivel, A. R. (2019). Neural correlates of workload transition in Multitasking: An ACT-R model of hysteresis effect. *Frontiers in Human Neuroscience*, *12*(January), 535–514. https://doi.org/10.3389/fnhum.2018.00535

Krulewitz, J. E., Warm, J. S., & Wohl, T. H. (1975). Effects of shifts in the rate of repetitive stimulation on sustained attention. *Perception & Psychophysics*, *18*(4), 245–249. https://doi.org/10.3758/BF03199369

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863–812. https://doi.org/10.3389/fpsyg.2013.00863

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simultated driving performance. *The Quarterly Journal of Experimental Psychology Section A*, *55*(2), 659–686. https://doi.org/10.1080/02724980143000505

Matthews, M. L. (1986). The influence of visual workload history on visual performance. *Human Factors*, *28*(6), 623–632. https://doi.org/10.1177/001872088602800601

Moacdieh, N. M., Devlin, S. P., Jundi, H., & Riggs, S. L. (2020). Effects of workload and workload transitions on attention allocation in a dual-task environment: Evidence from eye tracking metrics. *Journal of Cognitive Engineering and Decision Making*, *14*(2), 132–151. https://doi.org/10.1177/1555343419892184

Morgan, J. F., & Hancock, P. A. (2011). The effect of prior task loading on mental workload: An example of hysteresis in driving. *Human Factors*, *53*(1), 75–86. https://doi.org/10.1177/0018720810393505

Moroney, B. W., Warm, J. S., & Dember, W. N. (1995). Effects of demand transitions on vigilance performance and perceived workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *39*(21), 1375–1379. https://doi.org/10.1177/154193129503902104

Moroney, W. F., Reising, J., Biers, D. W., & Eggemeier, F. T. (1993). *The effect of previous level of workload on the NASA-Task Load Index (TLX) in a simulated flight task [Paper presentation]*. International Symposium on Aviation Psychology, Columbus, OH, 882–885

Murray, C. A., & Cox, C. B. (1989). *Apollo: The race to the moon*. Simon & Schuster.

Parush, A., & Rustanjaja, N. (2013). The impact of sudden events and spatial configuration on the benefits of prior information to situation awareness and performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1395–1399. https://doi.org/10.1177/1541931213571311

Prytz, E. G., & Scerbo, M. W. (2015). Changes in stress and subjective workload over time following a workload transition. *Theoretical Issues in Ergonomics Science*, *16*(6), 586–605. https://doi.org/10.1080/1463922X.2015.1084397

Ramchurn, S. D., Fischer, J. E., Ikuno, Y., Wu, F., Flann, J., & Waldock, A. (2015). *A study of human-agent collaboration for multi-UAV task allocation in dynamic environments [Paper presentation]*. IJCAI International Joint Conference on Artificial Intelligence, Buenos Aires, Argentina., 1184–1192

Sibley, C., Coyne, J., & Morrison, J. (2015). *Research considerations for managing future unmanned systems [Paper presentation]*. AAAI Spring Symposium, 54–59.

Ungar, N. R., Matthews, G., Warm, J. S., Dember, W. N., Thomas, J. K., Finomore, V. S., & Shaw, T. H. (2005). Demand transitions and tracking performance efficiency: Structural and strategic models. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(17), 1523–1526. https://doi.org/10.1177/154193120504901704

Van Acker, B. B., Parmentier, D. D., Vlerick, P., & Saldien, J. (2018). Understanding mental workload: From a clarifying concept analysis toward an implementable framework. *Cognition, Technology & Work*, *20*(3), 351–365. https://doi.org/10.1007/s10111-018-0481-3

Voorheis, C. M., Moroney, W. F., Biers, D. W., & Eggemeier, F. T. (2005). In search of workload context effects using a compensatory tracking task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(24), 2095–2098. https://doi.org/10.1177/154193120504902406

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Wickens, C. D., Gutzwiller, R. S., & Santamaria, A. (2015). Discrete task switching in overload: A meta-analyses and a model. *International Journal of Human-Computer Studies*, *79*, 79–84. https://doi.org/10.1016/j.ijhcs.2015.01.002

Wickens, C. D., Tsang, P., & Pierce, B. (1985). The dynamics of resource allocation. *Advances in Man-Machine Systems Research*, *2*, 1–49.

Williams, K. W. (2006). Human factors implications of unmanned aircraft accidents: Flight-control problems. https://doi.org/10.1016/S1479-3601(05)07008-6

Winnefeld, J. A., Jr., & Kendall, F. (2013). Unmanned systems integrated roadmap FY2013–2038.