



Task difficulty impact on multitasking in mixed reality environments

Safanah Abbas^a, Heejin Jeong^{b,*}

^a Department of Mechanical and Industrial Engineering, University of Illinois at Chicago, USA

^b Human Systems Engineering, The Polytechnic School, Ira A. Fulton Schools of Engineering, Arizona State University, USA

ARTICLE INFO

Keywords:

Task difficulty
Digital-physical hybrid
Hand use
Priority
Multitasking
Learning environment

ABSTRACT

Mixed reality applications allow users to experience both physical and digital worlds simultaneously. However, limited research has compared multitasking performance in these two worlds with different difficulty levels. This study investigated the effect of task difficulty on mixed-reality multitasking performance. A block-matching task was used for the physical task experiment, and the N-back test was used for the virtual task experiment. Thirty-six participants completed eight experimental conditions and were tested on four measures: NASA-TLX, accuracy, priority, and hand use. Physical demand and effort subscales were significant in all the experimental manipulations: real-world tasks, virtual-world tasks, and multitasking. The priority variable was not found to be substantial. Most participants prioritized working on either real-world tasks or multitasking, leaving the virtual-world tasks to the least preferred. Such a study can empower instructional designers to tailor tasks that align with learners' cognitive abilities for an optimal learning experience.

1. Introduction

Multitasking helps us manage the complex environments that we live in and is used to describe performing various tasks simultaneously (Dzubak, 2008; Lin, 2013). For instance, learners inherently exhibit multitasking, engaging in various tasks, whether related to educational activities or not, while listening to their instructors (Dönmez & Akbulut, 2021; Ettinger & Cohen, 2020; Jamet et al., 2020). Multitasking has been essential for brain and memory capacity research (Spink et al., 2008). It generally refers to the ability to coordinate multiple tasks at the same time to achieve an overall goal, and it is manifest in different ways, such as task switching (serial) and dual tasking (concurrent) (MacPherson, 2018; Salvucci et al., 2009). Dual tasking enables two tasks to be performed simultaneously, unlike task switching (Bridenbaugh & Kressig, 2015; Fraser & Li, 2012). In this research, we investigate multitasking performance in mixed reality and compare it with single-task scenarios in both the real (physical) and virtual (digital) worlds.

Successful multitasking becomes very challenging depending on the complexity of the activities being performed, not only to complete the tasks but also to effectively perceive the information associated with each task (Dzubak, 2008). Multitasking has received increasing attention with the development of new technologies (Lee et al., 2012). Because technologies today are not confined to a single location as

before, they allow for extensive multitasking (Abbas & Jeong, 2024). Thus, we often multitask and navigate through our phones or computers. Modern technology platforms broadly enable multitasking (Adler & Benbunan-Fich, 2012). This generates the importance of investigating multitasking behavior for human-computer interaction research.

Mixed reality (MR) or digital-physical hybrid worlds are great examples of modern technology. They allow users to use both environments, physical and digital worlds (Drascic & Milgram, 1996). Therefore, they deal with more user interaction than other technologies, such as virtual reality headsets (Rokhsaritalemi et al., 2020). The MR is increasingly gaining popularity and relevance (Speicher et al., 2019). Since it was predicted that smart glasses make workers more productive (Bourque, 2017), it is an excellent opportunity to research the effect of this technology on human performance. The field of mixed reality applications is expanding across diverse domains (Dai et al., 2021; Park et al., 2021; Gaol & Prasolova-Førland, 2021; Choi et al., 2022; Kim et al., 2023); hence, multitasking is considered one of the essential elements of this environment (North et al., 2021). By 2025, it was predicted that more than 14 million workers in the United States would be wearing smart glasses in work environments (Rogers, 2018). Furthermore, according to a report from Technavio (2020), the augmented and virtual reality market is projected to experience a compound annual growth rate exceeding 35% from 2020 to 2024, considering the influence of COVID-19. However, there are limitations to the widespread

* Corresponding author. Mesa, AZ 85212, USA.

E-mail addresses: sabbas28@uic.edu (S. Abbas), heejin.jeong@asu.edu (H. Jeong).

<https://doi.org/10.1016/j.cexr.2024.100065>

Received 13 December 2023; Received in revised form 28 February 2024; Accepted 6 April 2024

Available online 25 April 2024

2949-6780/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

adoption of immersive technologies in terms of comfort, usability, and mental effort, as highlighted in the report. For example, Apple Vision Pro, the most recent addition to the Apple product line, recommends stopping use if individuals experience physical discomfort, motion sickness, visual discomfort, or skin irritation (Santos & Castillo, 2024).

In the learning and training environment, on the other hand, as a dynamic domain, learners may need to navigate virtual environments, manipulate digital objects, and interact with the real world simultaneously. Augmented and virtual realities have increased students' and workers' learning and working efficiency by providing real-time information in education and training (Bednar & Welch, 2020; Lal et al., 2021; Lee, 2012). This plays a crucial role in understanding the impact of such technology on individuals for better educational experiences, considering factors such as task difficulty, cognitive load, and the interaction between real and virtual worlds. Therefore, multitasking has grown in both theoretical and practical research (Spink et al., 2008). Cognitive scientists and human factors engineers have performed considerable efforts to study multitasking performance (Meyer & Kieras, 1997). This paper investigates multitasking in mixed reality or hybrid worlds.

Task difficulty has been acknowledged as one of the most critical determinants of human performance (Cohen & Lefkowitz, 1977; Liu & Li, 2011). It is defined as the description of the characteristics of a task that represent the likelihood of completing the task successfully (Martin & Manning, 1995; Nawaz et al., 2020). According to the Yerkes-Dodson law, performance success is a function of both task difficulty and arousal (Yerkes & Dodson, 1908). This initiates the importance of taking task difficulty as a factor when studying human performance.

Completing high-complexity tasks typically demands more memory capacity, cognitive load, and other resources, impacting resource allocation (Gentili et al., 2018; Jaquess et al., 2018). Therefore, studying the relationship between task difficulty and performance is essential. This relationship has scientifically shown that a neural component increases intensity when making decisions with increasing task difficulty, which influences general cognitive processes (Philiastides, 2006). Due to the presence of large amounts of information, our cognitive systems can overload, which decreases task performance; hence, this implies the concept of mental workload when having limited information processing capabilities (Gentili et al., 2014; Guadagnoli & Lee, 2004; Sweller, 2010).

According to Sweller's cognitive load theory (CLT), a high mental workload would need extra resources to enter more information, creating a demand for these additional resources, leading to reduced performance (Sweller, 1988). Therefore, efficiently modifying task characteristics influences an individual's cognitive needs, as more challenging activities demand increased information, resulting in a positive correlation between heightened mental workload and informational demands (Sweller, 2010).

When individuals, such as learners, perform multiple tasks simultaneously, each requiring cognitive resources, the cognitive load theory provides insights into how the mental effort needed to complete a task influences learning and performance. It offers guidelines to assist in presenting information that encourages learner activities to optimize performance (Kirschner, 2002). The CLT addresses the limitations of working memory, as learners' memory is typically limited to about seven items or elements of information at any one time (Baddeley, 1992; Miller, 1956). Due to this limitation, instructions should be designed to ensure that working memory can effectively process the information, a principle that holds significant value for education. Achieving this involves navigating multiple factors, including casual factors like characteristics of the task (e.g., task complexity) and the type of environment, as well as assessment factors that involve evaluating the three measurable dimensions of cognitive load: mental load, mental effort, and performance (Paas & Van Merriënboer, 1993). Some studies investigating the use of multiple forms of media for learning, where information processing from different channels is involved, employ the

cognitive load theory to demonstrate that these processes can generate additional cognitive overload and challenge working memory capacity (Anmarkrud et al., 2019; Kalyuga, 2007).

Multiple resource theory states that people have fixed capacity resource pools separate to process information (Wickens, 2008). More resources are applied to the task, resulting in better performance (Norman & Bobrow, 1975). Hence, a decreased task performance indicates a shortage in these resources, which causes excess workload. In educational settings, multitasking has been shown to hinder meaningful learning by rapidly utilizing the limited capacity of learners' information processing channels, thereby resulting in insufficient attentional resources (May & Elder, 2018).

This study selected three main variables to investigate the effect of task difficulty on humans in MR multitasking: workload, priority, and hand use. It has been found in the literature that a link exists between workload and performance (Bruggen, 2015; Hancock & Matthews, 2018). Furthermore, workload and performance are sensitive to multiple characteristics of the presented task (Hancock et al., 1995) and other factors such as individual past experiences and abilities (Meshkati, 1988). The priority variable indicates which task is prioritized in each MR multitasking scenario we experimented with. Dual tasking consists of a primary and a secondary task (Kosson, 1996; Tran, 2020). Thus, taking the priority as a dependent measure helps to conclude which task difficulty level led our participants to prioritize which world task. In other words, it helps us know what task (real-world task, virtual-world task) was considered the primary one under each presented difficulty level. Using dominant and non-dominant hands affected task performance (Cramond et al., 1989; Schweiger et al., 2021). Moreover, there has been evidence of an interference effect between the simultaneous performance of both cognitive and motor tasks. (Baddeley & Della, 1996; Lindenberger et al., 2000; Theill et al., 2011). Nevertheless, research has shown that performance on a task with the non-dominant hand tends to be generally slower (Annett et al., 1979), and handedness as a variable has received much less attention in research (Prichard et al., 2013). Therefore, the hand-use variable is considered to aid in better designs for MR applications.

Given the complexity of the relationship between workload and task performance, it is essential to note that increasing workload doesn't always lead to decreased performance (Nachreiner, 1995). This study examines the connection between task difficulty and performance in a mixed-reality multitasking environment. By understanding how the introduction of task difficulty influences cognitive demands and efficiency in multitasking, this research sheds light on the intricate relationship between multitasking, workload, and task difficulty. As task difficulty varies, it can impact the mental workload experienced by individuals engaged in simultaneous tasks. Essentially, the study explores how task difficulty directly affects performance in multitasking within mixed reality environments compared to monotasking. The current paper reports the detailed answers to the following research questions (RQs).

- **RQ1-(a):** Does task difficulty (complexity) affect human workload in performing physical (real) tasks, digital (virtual) tasks, and mixed reality tasks?
- **RQ1-(b):** Is there a correlation between the different workload dimensions in performing physical (real) tasks, digital (virtual) tasks, and mixed reality tasks?
- **RQ2:** Does task difficulty affect the priority reported by participants in mixed reality multitasking?
- **RQ3:** Is there a correlation between the collected data of the hand use variable (self-reported hand used and observed hand used) in all our experimental manipulations, the physical (real) world, digital (virtual) world, and mixed reality multitasking?

In general, while increasing task difficulty may initially enhance performance, it could lead to a decline at higher levels, as Wood (1986)

suggested. Additionally, the difficulty level at which a task is performed has been found to influence mental workload (Colle & Reid, 1998; Jaquess et al., 2018). Therefore, we hypothesize that task difficulty in mixed reality significantly affects human workload and performance. The detailed hypotheses derived for each RQ are as follows.

- **H1-(a):** An increase in task difficulty (complexity) will affect the perceived human workload for physical (real) tasks, digital (virtual) tasks, and mixed reality tasks.
- **H1-(b):** Significant correlations will be observed among the different workload dimensions across physical (real) tasks, digital (virtual) tasks, and mixed reality tasks.
- **H2:** Task difficulty will impact the reported priority by participants between the physical and virtual tasks in multitasking.
- **H3:** Significant correlations will be found between self-reported hand use and observed hand use across all the experimental conditions.

The current research has two main contributions. First, it provides empirical science into the effect of task difficulty on performance in MR applications. Second, it helps organizations using MR technology better understand and expect its returns. The study design involves conducting multiple experimental conditions for both monotasking and multitasking. These conditions include performing only physical tasks, only virtual tasks, and multitasking between the two—referred to as mixed reality. The study also examines combinations of two difficulty levels: easy vs. hard. Participants experience all conditions, indicating a within-subject design. This approach directly compares performance within the same individuals across various task difficulty levels and conditions.

2. Related work

2.1. Task difficulty effect in digital (virtual) reality environments

We witness increased augmented, virtual, and mixed-reality display use (Kaplan, Cruit, Endsley, Beers, Sawyer, & Hancock, 2021). With this increase in high-speed communication, virtual and augmented reality are becoming the next-generation display platforms; they provide deeper interactions between humans and the digital world (Xiong et al., 2021). Hence, they become more accessible (Willemssen et al., 2018). Unlike existing methods, such technologies are beneficial because they provide a controlled stimulus environment that can present cognitive challenges in an explicit control distraction (Rizzo et al., 2000). One of the main advantages of VR is that it provides humans with the training needed to acquire new skills cost-efficiently. Simulation-based training has gained wide acceptance among organizations and industries (Hancock, 2009). VR applications have been used to test various cognitive processes such as spatial abilities (Larson et al., 1999; McComas et al., 1998; Rizzo et al., 1998; Stanton et al., 1998), memory (Dinh et al., 1999; Grealy et al., 1999), and attention processes (Rizzo et al., 2009; Wann et al., 1997).

Since virtual reality applications provide tasks required to be fulfilled by their users, task difficulty has been assigned as one of the most critical task aspects (Sheridan, 1992). Several studies were conducted to investigate the effect of task difficulty on performance in VR settings. Poeschl (2017) conducted a study to measure public speaking performance in a VR application with two difficulty levels: low and high. Low task difficulty was for giving a speech with preparation, and high task difficulty was otherwise. Task difficulty was found to have an insignificant effect on VR public speaking training.

Shuggi et al. (2017) investigated the impact of mental workload under different levels of task difficulty in a reaching movement task using a virtual robotic arm. The results of NASA-TLX in this study indicated that the mental demand for the group assigned a higher difficulty was significantly higher. The other dimensions of the NASA-TLX did not disclose any significance. In another example, a study was

conducted to assess the performance of 60 participants in navigating through a virtual office building while performing various spatial orientation activities (Neale, 1997). Two levels of difficulty were tested: easy vs. hard. The results showed that as task difficulty increased, more errors occurred.

Kaber et al. (2012) examined human performance in a VR surgical simulation under different task difficulty settings. It was found that user performance degraded with increasing task difficulty. On the contrary, Orvis et al. (2007) conducted a study to measure participants' performance in instructional video games with four levels of task difficulty: static, increasing, adaptive-low, and adaptive-high. It was found that task difficulty had a significant effect on improving participants' performance and motivation.

Rao et al. (2020) developed a VR search-and-shoot simulation to test user performance with two scenarios of task difficulty, novice and professional. Participants were divided into two groups: the first group completed the novice scenario first, and the second group performed the professional scenario first. Various cognitive and behavioral measures were taken, including NASA-TLX. The result showed that the group that started with the professional scenario performed better than the others. Jaquess et al. (2018) conducted a study to monitor mental workload through multiple practice sessions with two different difficulty levels. The task was to control an airplane in a high-fidelity flight-simulated environment. Results were analyzed via NASA-TLX and electroencephalography (EEG). It was found that self-reported mental demand was reduced while working memory engagement was increased.

In the previous literature, there has been evidence that handedness as a variable can provide beneficial insights into individual differences in behavior (Kaploun & Abeare, 2010; Wright et al., 2004). However, these individual differences in the hand-use variable can also be seen in VR applications. For example, Pedowitz et al. (2016) examined the performance of the dominant versus the non-dominant hand in a virtual reality surgical setting. Two groups of surgeons were assigned based on difficulty level (expert vs. residents). It has been shown that a statistical difference between the two hands was found only among resident surgeons. For this reason, virtual reality that embodies cognitive loads can be used to measure motor skills.

In another study, Nataraj et al. (2022) found a positive relationship between performance and binding when the dominant hand was observed in a virtual reach control application. The reaching path length represented performance, while binding was the time interval estimation of the beep sound after the grasping contact happened. These authors conclude that considering the dominant hand helps optimize the performance and engagement of VR applications. Challenging goals/objectives have led to better performance than easy goals (Locke et al., 1981). The importance of individual differences in abilities and behavior might also depend on the task's difficulty (Fleishman, 1957; Liu et al., 2010).

2.2. Task difficulty effect in physical (real) environments

The real-world environment can also be referred to as the actual environment in everyday life. Several studies examined the effect of task difficulty on performance in the physical (real) world. For example, Li et al. (2019) conducted a study to investigate whether the performance of the P300-speller can be improved when mental effort is increased. This P300-speller provides a means of communication to patients with amyotrophic lateral sclerosis. In this study, twenty-seven subjects were recruited, and two difficulty levels were determined: (DC-task) and (DC + CN task). DC-task required the subjects to distinguish the color of a target. In contrast, the DC + CN task required them to determine the color of a target and simultaneously count the number of times the target flashed. It was found that the largest amplitudes were induced in the DC + CN task. This study concluded that mental tasks with various difficulty levels could significantly optimize the performance of the P300-speller and improve its accuracy. Salehi and Orak (2020) investigated the

task difficulty factor on a physical basketball chest pass performance. It was found that task difficulty (nominal and functional) significantly affected performance. The higher effect was for the functional difficulty level.

The hand used to perform a task could impact individual performance; the literature showed contradictory findings on task complexity and its influence on hand selection. Several studies found a significant effect of task difficulty on hand selection (Gabbard et al., 2003; Provins & Magliaro, 1989), while other studies found the opposite (Bryden et al., 2007; Bryden & Huszczyński, 2011; Bryden & Roy, 2006). Lee et al. (2014) conducted a study investigating the effect of cognitive workload between robotic and laparoscopic surgeries using both dominant and non-dominant hands. The mental workload was assessed using NASA-TLX. Participants were asked to remove rings from a wire using a grasper by their nondominant hand and place them in a bin. When all the rings were collected, they were asked to return them to the wire using their dominant hand. The authors showed that robotic surgery's task performance was better than laparoscopy. The muscular workload of the non-dominant hand was significantly higher than that of the dominant hand.

2.3. Task difficulty effect in multitasking

Having explored the effects of task difficulty in both the real and digital worlds, multitasking becomes a dynamic interplay of cognitive demands and efficiency in our evolving environment. In the face of rapid changes and the expansive reach of new technologies, the responsibility of managing multiple tasks is entrusted to us. For example, workers simultaneously delivered numerous projects to different stakeholders (Lindbeck & Snower, 2000; Schieman & Young, 2015) and neither did students work on a single homework submission. Thus, it became an explicit expectation in most organizations that employees would simultaneously handle their time by implementing different work activities (Kaufman-Scarborough & Lindquist, 1999). Multitasking is beneficial because it satisfies complex work demands promptly; it is fair to say that such activity would result in an increased workload and mental fatigue (Maume & Purcell, 2007; Perlow, 1999).

Liu et al. (2016) studied the relationship between task difficulty and performance in a multitasking scenario. Three levels of task difficulty were assigned (low, medium, and high), and the NASA-TLX questionnaire was used to assess workload. Results have shown that task difficulty significantly impacted performance. On the other hand, Bowers and Christensen (2014) found that task difficulty did not reveal any significance on workload performance via the NASA-TLX measure. Bowers et al. conducted a study that employed the Air Force multi-attribute task battery with two difficulty levels, low and high. Participants performed conditions that transitioned from easy to hard levels and vice versa. In another example, Haga et al. (2002) tested twelve subjects on dual tasking (a tracking task and a memory search task) with various difficulty levels on a computer screen. NASA-TLX has been used as the subjective measure of workload, and the results have shown that workload was sensitive to the change in task difficulty.

Shin et al. (2020) conducted a study investigating task performance and mind wandering of light, intermediate, and heavy media multitaskers as a function of task difficulty in an N-back task. The findings showed no significant difference in performance across difficulty levels between light and heavy media multitaskers. On the other hand, intermediate media multitaskers performed better at the more difficult levels than heavy multitaskers. Sanjram and Gupta (2013) examined the effect of task difficulty and time constraints on performance and cognitive workload in multitasking within a programming context. The results showed increased task difficulty and time constraints affected the prospective memory performance. It was indicated that different difficulty levels of the programming task did not necessarily significantly impact cognitive workload, but time constraints did. Liu and Nam (2018) investigated the relationship between working memory capacity, task

difficulty, and multitasking performance. Their study aimed to quantitatively evaluate multitasking systems better to understand the interaction between systems and human operators. Their experimental results have shown that task performance decreased while subjective workload increased when task difficulty increased.

In education, it is expected to observe learners simultaneously engaging in multiple tasks in libraries and classrooms while studying or preparing for upcoming exams. However, most previous studies have been conducted in familiar environments, such as classrooms or libraries, rather than attempting immersion in virtual reality learning environments. Tugtekin and Odabasi (2023) explored the effects of multitasking and task characteristics like difficulty on learning outcomes and subjective cognitive load in virtual environments. The study's participants were divided into different groups, involving multitasking with a regular task, a hard task, and monotasking. The results, tested by MANCOVA with prior knowledge as the covariate, indicated a significant effect for both multitasking and task characteristics in the VR learning environment.

Although the complexity of work in multitasking learning environments has received relatively little attention (Hahnel et al., 2019), task characteristics are expected to influence learners' interaction with learning tasks (Lee & List, 2021). Previous literature studies have examined multitasking in learning environments. However, there is still a gap in the research field regarding an interactive learning environment rather than a traditional one, multitasking, task difficulty, and their impact on learning outcomes and cognitive load (Tugtekin & Odabasi, 2023).

North et al. (2021) studied multitasking between the physical and digital worlds to determine whether immersion in MR significantly impacts performance and perception. Although it was not clearly stated whether task difficulty was taken as a factor in this study, a negative impact was found on the overall effectiveness when multitasking in a mixed reality environment. Vallejo et al. (2017) developed a virtual serious game-based multitasking assessment tool for cognition and activities in daily life in comparison with the real world. Participants were asked to execute a cooking activity in the virtual and real-life game. The study aimed to measure the cognitive processes engaged in a multitasking activity. The results revealed that the virtual task performance was highly similar to what participants did in the real-life task. Hence, it was concluded that performances in the virtual world task were good predictors of real-world performance.

Since mixed reality involves an engaging interface, efforts to optimize human-computer interactions have increased (Nataraj et al., 2022). The functioning of the human hand, including its physiology and motor control, has been a critical topic for scientific research in the 19th century (Jones & Lederman, 2006). Consequently, the behavior of the hand dominance of different individuals could indicate a difference in multitasking performance. According to Steingrueber (1975), task performance changes with changes in task complexity when the dominant hand is observed.

2.4. Lack of multitasking research of task difficulty effect in digital-physical hybrid worlds

Most research indicated that task performance was negatively related to task difficulty levels, such as Bonner (1994), Xu et al. (2008), Ziefle and Bay (2005), Topi et al. (2005), and Cho (2018). On the other hand, other studies showed a positive relationship between task complexity and performance (Mascha, 2001; Wu et al., 2012). As we are interested in studying the performance in MR, the research field is desperate to determine which factors specifically contribute to better training in augmented reality and virtual reality (AR/VR) (Kaplan et al., 2021). Hence, we aim to study the impact of task difficulty, which is a pivotal factor influencing human effectiveness, on performance in comparison between the physical and digital worlds.

Most studies have investigated the impact of task difficulty on

multitasking performance in physical (real-world) settings. However, many studies compared the physical and digital worlds (Brade et al., 2017; Coutrot et al., 2019; Joyner et al., 2021), but we found almost no research related to the investigation of task difficulty effect on multitasking performance between the two worlds. Recently, mixed reality applications received greater attention and have been implemented in health care (Kim et al., 2021; Lungu et al., 2020), education (Gaol & Forland, 2021), construction (Dai et al., 2021; Ogunseju et al., 2021), and so on. Hence, our research in mixed reality addresses the gap in understanding the impact of task difficulty on human performance in dynamic environments. This study builds upon previous relevant research. Additionally, it was found that virtual reality tasks conducted in previous studies had an increased difficulty, requiring additional cognitive resources because of the poorer performance obtained in VR than in traditional or computerized assessment tools (Negut et al., 2016). Therefore, it is necessary to provide a better virtual reality cognitive test.

3. Method

3.1. Participants

Thirty-six participants were recruited in this study, ranging from 19 to 43 ($M = 23.9$, $SD = 4.22$) years old. Multiple effect sizes were analyzed using a power analysis (Cohen, 1992) with 80% and 90% power, respectively. With an effect size of 0.5 and an alpha of 0.05, it was determined that a minimum of 28 participants was needed to achieve 80% power. However, given budget and time constraints, other trials indicated larger sample sizes beyond our capacity to collect. Participants reported varying levels of expertise in using AR/VR devices. No hearing impairment was present, and all participants had normal or correct to normal vision. A total of 34 participants reported being right-handed, while only two participants were left-handed. The authors obtained Institutional Review Board approval (IRB# 2020-0466) for this study.

3.2. Experimental tasks

Two different worlds were examined in this study: (1) the digital (virtual) world (VW) and (2) the physical (real) world (RW), as modern technologies would be taking over organizations and industries (Susskind & Susskind, 2015), comparing each world environment separately

and altogether was taken into consideration for this study to cover the most of existing work contexts.

3.2.1. The digital (virtual) world task

We developed a virtual N-back task application augmented in a Microsoft HoloLens 2 device to assess the digital world performance while performing a physical world task simultaneously if necessary (See Fig. 1). The N-back task we used in this study was matching the colors of a virtual cube. We employed two difficulty levels for the N-back task, $N = 1$ and $N = 2$. The N-back task is a working memory measure to assess individual differences. (Jaeggi et al., 2010), The N parameter represents the number of steps requested to return to memory to return the required response (Chen et al., 2008). The main advantage of this N-back task is that when changing the size of the N parameters, demands on executive working memory can easily be adjusted (Monk et al., 2011). Fig. 2 shows an illustration of the virtual N-back task.

We employed several features in this N-back task application, including timer duration where it stops at a customized time (e.g., 90 s), a stop button to end the current task, double-tap/touches to register the answers of the participants and lock object's position where participants can place the virtual cube in the space anywhere they feel comfortable with allowing their hands to reach it without any extra exertion. The application plays an alerting sound once the timer is up. Then, the display will show the final results with the four specific pieces of information: (1) user correct answers, (2) missing correct answers, (3) user incorrect answers, and (4) total possible stimuli. In this study, we selected a constant total of 34 possible stimuli for all participants. Fig. 1 shows a snapshot of the virtual screen for the N-back application. This application was designed to measure working memory capacity and used as an assessment tool for working memory training with two different modes. Mode 1 displays only one object's shape (i.e., a cube). Mode 2 displays two object forms (i.e., a cube and a cylinder). The object's position will be locked once the participant clicks on either Mode 1 or Mode 2. Thus, the participant cannot move the object to a different location unless the task is stopped. In this study, only Mode 1 was used.

This application allows participants to immerse themselves in augmented reality and interact with the virtual 3D cube, which can test their cognitive abilities. Such applications provide an assessment tool to safely measure and train people's mental skills using this augmented reality N-back task. Hence, it ultimately helps understand how people's working memory capacity operates when exposed to a mixed reality cognitive test.

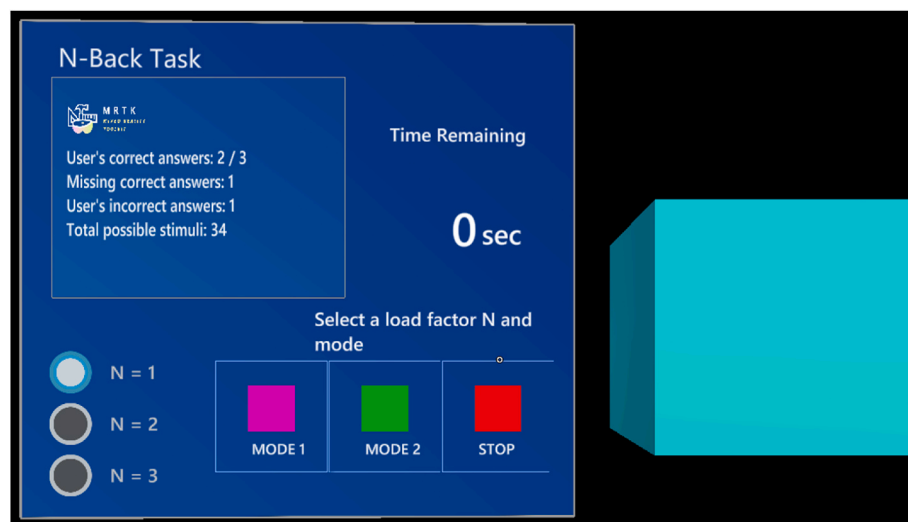


Fig. 1. The virtual screen of the N-back application.

Fig. 1 Alt Text: A screen-captured image of the N-back task interface, including the result of the tasks on the upper-left, the time remaining on the middle, the selection buttons for each mode on the bottom, and the colored cube on the right.

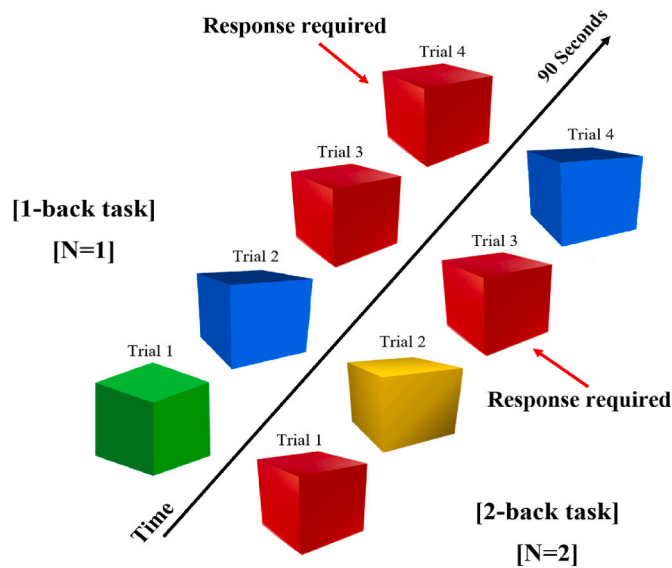


Fig. 2. Illustration of the virtual 1-back and 2-back tasks.

Fig. 2 Alt Text: An example of four trials for the virtual N-back test for each task, showing different colored virtual cubes by time frame, with required responses for each task indicated by red arrows. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

MR technology is human-computer interaction (Dubois et al., 2009; Pan et al., 2006). Thus, interacting with such systems impacts the user's cognitive skills. It has been found that the knowledge of human cognitive behavior is sufficiently advanced to enable applications in several functional domains to design efficient and easy-to-use interactive computer systems. Therefore, bringing cognitive behavior into the practical problems of the design represents an attempt to bridge the gap between science and application. (Card et al., 2018). Targeting the cognitive processes during the development of VR systems is needful as it helps those having attention difficulties such as deficit/hyperactivity disorder (ADHD), traumatic brain injury (TBI), and any other neurodegenerative disorders (i.e., Alzheimer's disease) (Martinussen et al., 2005; Rizzo et al., 2000; Willcutt et al., 2005).

3.2.2. The physical (real) world task

A block-matching task was used to measure participants' performance in the physical world using an educational material (Getianlai Toys), a wooden blocks puzzle board. The board has twenty-six uppercase English letters, and participants were asked to pair pre-assigned blocks on the board using their hands within a given time limit (90 s). Participants would visually search for the correct English letter to pair on the wooden board with its matching letter. Thus, their mental

workload would be affected (May et al., 1990; Greef et al., 2009), and their reliance on their working memory would be increased (Luria & Vogel, 2011). When participants were not given any order to pair the English letters into the wooden board, this was set as the easy level of the physical task. For the hard level, three different orders of pairing the letters were given to the participants: (1) forward-bottom-top (FBT), (2) backward-top-bottom (BTB), and (3) backward-bottom-top (BBT). The wooden board had two rows of English letters; the first row had the letters from A to M, and the second row had the rest. FBT order asked participants to pair the letters starting from N to Z, then from A to M. While the BTB order was about pairing letters starting from M to A, then from Z to N. Lastly, the BBT order was from Z to N, then M to A. Fig. 3 shows a screenshot from the recorded videos of one of the participants performing the physical world task.

3.3. Experimental design

A total of 36 participants were recruited in this experiment. Each participant performed eight experimental conditions. Two conditions for the physical task, two for the digital task, and four for multitasking between both. Task difficulty was set as the independent variable. Two task difficulty levels were assigned (easy vs. hard). The two physical world task conditions were described as easy and hard RW tasks. The digital world tasks were described as an easy VW task and a hard VW task. The four multitasking conditions were easy RW & easy VW, easy RW & hard VW, hard RW & easy VW, and hard RW & hard VW. The dependent variables were set as human workload, total completion time, accuracy, priority, and dominant hand. We hypothesize that task difficulty significantly affects human workload and performance. The order of the experimental conditions was randomly given to each participant to avoid biases from the experiment side.

3.4. Procedure

At first, participants' recruitment processes started with filling out an eligibility survey to ensure the absence of any impairment that would affect their performance. Eligible participants signed a consent form and completed a demographic survey (gender, age, etc.). We also asked the demographic survey participants about their expertise in working with AR/VR devices and their dominant hands. Each participant performed a few practice experimental conditions to get familiar with the nature of the tasks; they also watched a demonstration video illustrated for the digital task. After each practice experimental condition, they completed a pairwise comparison survey for the NASA-TLX dimensions. Each participant randomly performed eight experimental conditions when familiarizing themselves with the tasks. They were asked to use their hands to pair the English letters with the wooden board for the physical task and to interact with the virtual cube when they needed to respond to the stimulus during the digital task. After each actual experimental

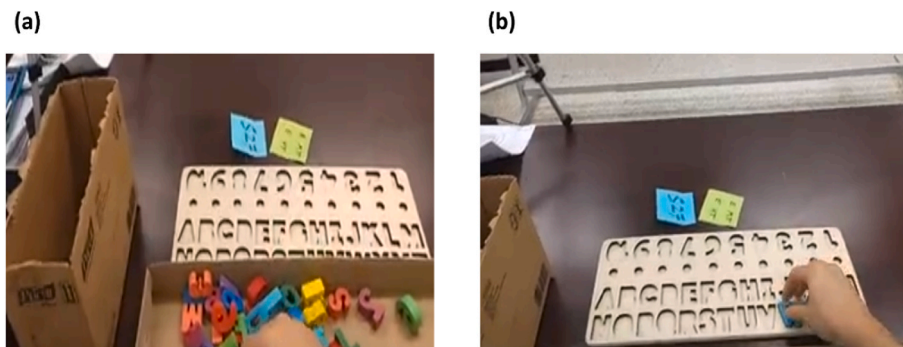


Fig. 3. The physical (real) world task from the participant's point of view.

Fig. 3 Alt Text: the participant's view (a) looking for the correct English letter to pair on the wooden board, and (b) placing the letter on the board after finding it.

condition, they completed a post-task survey, NASA-TLX. At the end of the survey, they were also asked which hand they used (right-hand, left-hand, or both hands) and RW-VW multitasking priority (prioritizing only the RW task, only the VW task, or both tasks). Each experimental condition had a timer of 90 s; therefore, they were asked to stop if the timer hit this limited time. The digital world task (virtual) had a built-in timer in the n-back test application, which stopped the test immediately when the timer hit 90 s. However, a video camera timer set the 90-s time limit for the physical world task. After completing all the experimental conditions, they filled out a payment form. The total duration of the experiment was approximately 90 min for each participant. Fig. 4 shows the experiment flow.

4. Measurements

4.1. NASA-TLX

NASA-TLX is a multi-dimensional scale designed to assess operators' workload while performing a task or afterward (Hart, 2006). It measures performance according to six subscales: (1) mental demand (MD), (2) physical demand (PD), (3) temporal demand (TD), (4) performance (P), (5) effort (E), and (6) frustration (F) (McKendrick & Cherry, 2018). Participants rated their scores for each dimension on a percentage scale from 0 to 100. These data were taken to calculate the weighted rating. As mentioned in the procedure section, each participant completed a pairwise comparison survey for the NASA-TLX dimensions. A total of 15 comparisons were made, and a count of each dimension was calculated. Thus, the final weighted rating resulted from the sum product of each participant's answers to NASA-TLX dimensions ratings (0–100%) in the post-task survey with the results of the counts from the pairwise comparison survey divided by 15.

4.2. Accuracy

We assigned accuracy as a dependent measure to determine the degree of task correctness. Different formulas were used to calculate the accuracy of each world task. Equation (i) for the physical world task and equation (ii) for the digital world task. Both equations (i) and (ii) were combined into an average to calculate the accuracy for the four multitasking conditions. A separate accuracy was calculated for each multitasking condition for the RW and VW tasks. Then, these calculated accuracies were combined into an average. This ensured that the differences in the complexity level between the tasks in the four multitasking conditions were accounted for.

$$\text{RW accuracy} = \text{total number of paired English letters} \div 26 \quad (\text{i})$$

$$\text{VW accuracy} = \text{the number of correct answers} \div (\text{total number of answers})(\text{ii})$$

4.3. Total completion time (TCT)

TCT is collected to measure how long it took each participant to complete each experimental condition. We took the TCT as a dependent variable for only the physical task conditions (i.e., easy RW task, hard RW task) because the rest included the VW task with a built-in timer of 90 s. A video camera was used to measure the total completion time, and the experimenter recorded the TCT of each participant.

4.4. Priority

This measure was taken for only the multitasking conditions, where participants needed to prioritize one of the tasks. After each multitasking condition, participants were asked to self-report their priorities in the post-task survey of NASA-TLX, including a multiple-choice question at the end. They were prompted to indicate whether they prioritized tasks from the physical world, the digital world, or both equally.

4.5. Hand use

Two data types were collected for this variable: observed hand use recorded by the experimenter's visual inspection and self-reported hand use recorded by each participant. Three answers were collected: right hand, left hand, or both hands. Studying the discrepancy between these data increases the perceived behavioral control, contributing to better habitual behavior (Zomer et al., 2013). The hands of our participants were involved in both experimental tasks, the physical block-matching task and the virtual n-back task. Hence, studying this variable is essential if the workload is high, which could lead the participants to incorrectly report the hand they used, resulting in their memory impairment. They used their hands to look for the correct letter to pair for the physical task, and then they had to place it on the wooden board. Thus, they repeated this process until they paired all the English letters as requested within a 90-s limit. For the virtual world task, participants had to place the virtual object (cube) in the space in front of their eyes, where they felt comfortable reaching with their hands. However, participants were not consistent with each other regarding the virtual object's position placement. Every time a correct stimulus appeared, they needed to respond by touching the virtual cube with their hands. This means they had to constantly move their hands back and forth while performing this task. The movement of participants' hands for physical and the digital tasks might affect their performance.

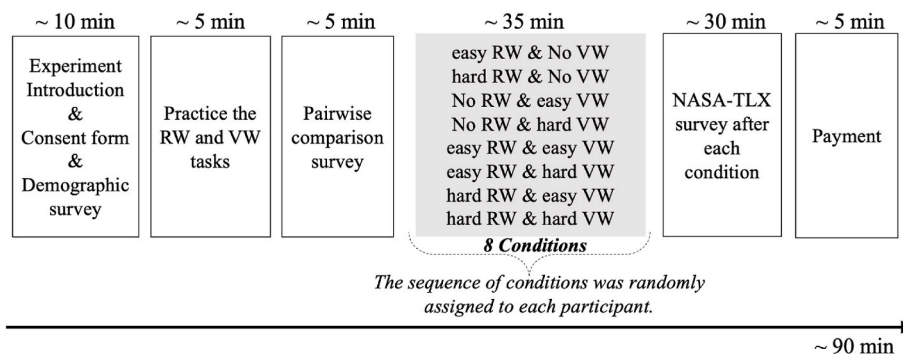


Fig. 4. The flow of the experiment.

Fig. 4 Alt Text: The procedure of the experiment is depicted in a flow chart. Initially, participants signed a consent form, received an introduction to the experiment, and completed a demographic survey. Next, they engaged in a brief practice session to familiarize themselves with the tasks, followed by the completion of a pairwise comparison survey after each practice session. Subsequently, they performed the actual eight experimental conditions in a random order and filled out a NASA-TLX survey to evaluate each dimension. Finally, participants were compensated for their time in cash.

5. Data analysis

A total of 288 data were collected (8 conditions $\times$ 36 participants). Table 1 summarizes each task difficulty level data for the single and multitasking conditions. Six dependent variables were collected for the data: (1) NASA-TLX, (2) weighted rating, (3) total completion time (TCT), (4) accuracy, (5) multitasking priority, and (6) hand use. Only (5) and (6) are categorical, whereas the rest DVs are numerical. Normality tests and visual inspection of normality plots and histograms were conducted to check the data. All numerical dependent variables were not normally distributed even though trials of data transformation were implemented. Two significant tests were used to analyze the numerical measures: Mann-Whitney and Kruskal-Wallis. The Mann-Whitney test is the non-parametric alternative to the *t*-test (MacFarland & Yates, 2016). The Kruskal-Wallis test is the non-parametric approach to the one-way ANOVA (McKight & Najab, 2010).

The analysis of the data was divided into three groups: (1) single physical (real) world task, (2) digital (virtual) world task, and (3) physical-digital multitasking. For the analysis of (1) and (2), we performed the Mann-Whitney test to determine the significance of the task difficulty level since there are only two groups of each division (easy RW, hard RW, easy VW, and hard VW). We used the Kruskal-Wallis test for the four multitasking conditions because we have more than two groups (easy RW & easy VW, easy RW & hard VW, hard RW & easy VW, hard RW & hard VW). We obtained the *p*-values for each case to determine its significance. Effect sizes were also brought to specify how task difficulty accounts for variance in each dependent variable. The importance of securing effect sizes relies on accompanying it with an interval estimate that acknowledges the uncertainty with the population value of the effect size (Kelley & Preacher, 2012). Two different formulas were used for calculating the effect sizes: Pearson correlation coefficient *r* (Equation iii) for the Mann-Whitney test (Tomczak & Tomczak, 2014) and epsilon squared estimate (Equation iv) for the Kruskal-Wallis test (Tomczak & Tomczak, 2014). Post hoc pairwise comparison tests were performed for the multitasking conditions analysis.

$$r = |Z| \div \sqrt{n} \quad (\text{iii})$$

$$E_R^2 = H \div ((n^2 - 1) \div (n + 1)) \quad (\text{iv})$$

Where *n* is the total number of observations, and *H* is the value obtained from the Kruskal-Wallis test. A Chi-squared test was conducted to test the significance of the priority measure. Correlation analysis was conducted on the hand use variable to determine the relationship between observed and self-reported hand use by each task difficulty level. Data analysis was performed using R programming language (Gardener, 2012), with an assigned significance level of $\alpha = 0.05$.

6. Results

6.1. NASA-TLX

To answer RQ1-(a), a Mann-Whitney test was implemented to test the significance of each NASA-TLX subscale on the first two divisions: (1) physical (real) world task and (2) digital (virtual) world task. For single RW task conditions (easy RW, hard RW), a significant difference was found on four subscales: PD, TD, E, and F, with *p*-values of .01, .03, .04, and .01, respectively. Mental demand and performance were not significant. 9.4% of the PD scores variance was accounted for by task difficulty, while 6.62% for TD, 5.82% for E, and 10.56% for F. Comparing all the four significant subscales, we found the higher significance was for frustration. The weighted rating was also significant, with a *p*-value of .03 and an effect size of 7.1%, indicating that task difficulty is associated with a 7.1% difference in the weighted rating results. Fig. 5 shows the boxplots of the significant NASA-TLX dimensions for single physical (real) world tasks.

For single VW conditions (easy VW, hard VW), we also found a significant difference in the same four subscales with *p*-values < .05 and effect sizes of 16.6%, 5.83%, 21.90%, and 13.21%, respectively. The highest significance was for PD and E. The weighted rating was also significant, with a *p* < .01. Fig. 6 shows the boxplots of the NASA-TLX

Table 1
Descriptive statistics by task difficulty for all the experimental conditions.

Dependent measures (Numerical)	Single task (physical)		Single task (digital)		Multitasking (physical-digital)			
	Task difficulty		Task difficulty		Task difficulty			
	Easy	Hard	Easy	Hard	Easy RW – Easy VW	Easy RW – Hard VW	Hard RW – Easy VW	Hard RW-Hard VW
Mental demand (MD)%								
Mean	30.44	29.77	20	24.31	41.11	44.75	41.06	44.11
(SD)	(24.69)	(26.53)	(16.15)	(23.05)	(26.57)	(26.48)	(27.48)	(27.64)
Physical demand (PD)%								
Mean	25.94	37.34	36.11	57.4	63.14	75.39	72.42	82.19
(SD)	(24.19)	(23.71)	(24.62)	(24.51)	(23.35)	(19.29)	(26.25)	(23.03)
Temporal demand (TD)%								
Mean	40.14	55.17	27.91	38.83	55.17	64.56	67.67	66.47
(SD)	(26.87)	(28.27)	(22.07)	(23.00)	(23.33)	(25.41)	(27.66)	(28.56)
Performance (P)%								
Mean	48.47	43.91	48.23	47.71	51.67	50.83	47.53	52.86
(SD)	(48.71)	(23.58)	(40.36)	(28.39)	(24.14)	(22.85)	(22.60)	(27.55)
Effort (E)%								
Mean	31.83	51.94	34.43	56.54	66.42	72.89	75.25	82.11
(SD)	(21.67)	(39.46)	(22.31)	(22.27)	(22.50)	(19.79)	(20.34)	(17.80)
Frustration (F)%								
Mean	11.11	23.31	17.34	33.2	36.53	44.28	49	53.14
(SD)	(14.33)	(21.82)	(16.71)	(24.04)	(28.14)	(27.44)	(32.78)	(33.44)
Weighted rating%								
Mean	32.65	43.96	33.22	49.77	56.97	63.41	64.59	70.10
(SD)	(19.61)	(21.79)	(16.69)	(16.37)	(14.62)	(14.59)	(16.50)	(16.20)
TCT in seconds								
Mean	70.42	86.58	–	–	–	–	–	–
(SD)	(12.22)	(6.09)	(–)	(–)	(–)	(–)	(–)	(–)
Accuracy%								
Mean	98.81	84.72	74.05	41.25	53.83	36.84	40.18	27.86
(SD)	(4.43)	(20.77)	(34.07)	(33.73)	(25.61)	(16.14)	(22.82)	(21.08)

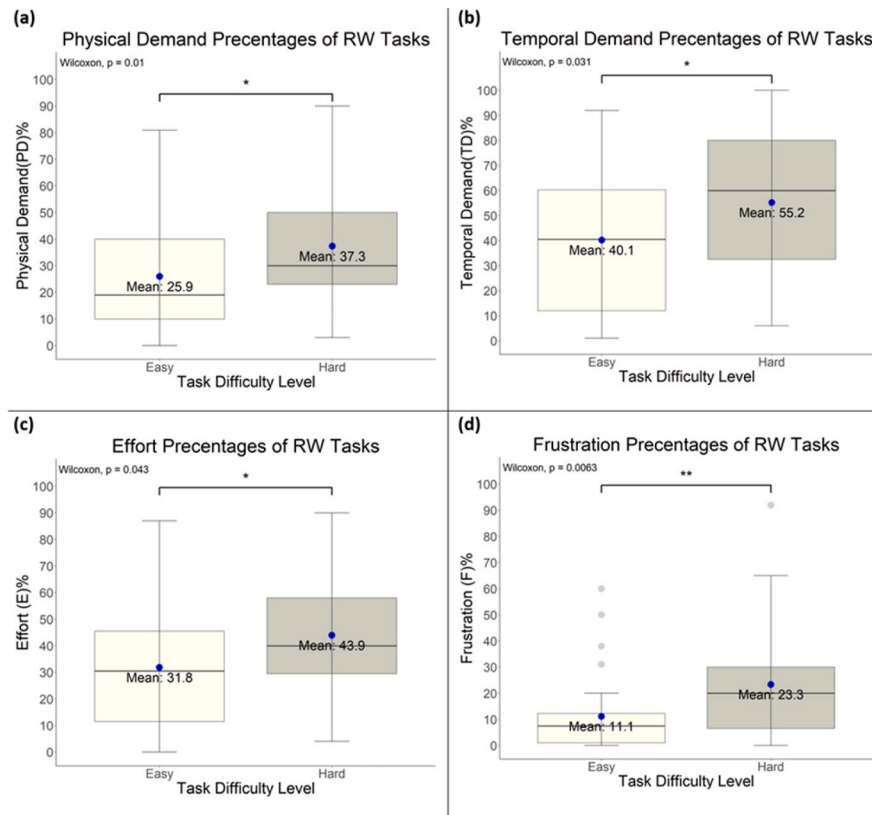


Fig. 5. Significant dependent measure of NASA-TLX dimensions for single physical (real) world tasks.

Fig. 5 Alt Text: Four boxplots showing task difficulty level (easy and hard) on the x-axis and the percentages of the significant NASA-TLX subscales on the y-axis: (a) PD% data, (b) TD% data, (c) E% data, and (d) F% data (ns: $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

dimensions that showed significant results for single digital (virtual) world tasks. The results in Figs. 5 and 6 show the same pattern of the four significance measures, where the hard level was higher than the easy level.

For RW-VW multitasking, a significant difference was found in the results of PD and E, with a p -value of $<.01$ and $.02$, respectively. The rest of the NASA-TLX subscales were insignificant. We found that about 10% of the variance in PD scores was accounted for by task difficulty, while the same factor accounted for 7.25% of the variance in the E scores. Post hoc tests were implemented for both significant subscales to locate those specific differences. Six pairwise comparisons resulted: (1) easy RW & easy VW-hard RW & easy VW, (2) easy RW & easy VW-hard RW & easy VW, (3) easy RW & hard VW-hard RW & easy VW, (4) easy RW & easy VW-hard RW and hard VW, (5) easy RW & hard VW-hard RW & hard VW, and (6) hard RW & easy VW-hard RW & hard VW. Table 2 provides a summary of the significant NASA-TLX subscales.

We found the exact significant difference in task difficulty level for both physical demand (PD) and effort (E) in comparing easy RW & easy VW-hard RW & hard VW with a p -value of $<.001$ and $<.01$, respectively. Fig. 7 shows the boxplots of the significant NASA-TLX dimensions for the multitasking conditions, where the higher significance was for the physical demand subscale. The weighted rating of multitasking was also significant, with a p -value of $.01$ and an effect size of 9.02% . A post hoc test was also implemented on the results of the weighted rating measure, and we found the exact significance was for the same comparison group as in physical demand and effort. Fig. 8 shows the significant weighted rating measures for all three experimental manipulations.

Additionally, to answer RQ1-(b), we performed the correlation analyses among each NASA-TLX subscale to see whether one is associated with another. We analyzed the results for each experimental manipulation by each task difficulty level: (1) single physical tasks (easy RW,

hard RW), (2) single digital tasks (easy VW, hard VW), and (3) MR multitasking (easy RW & easy VW, easy RW & hard VW, hard RW & easy VW, hard RW & hard VW). For (1), we found the strongest correlation for the easy level was between physical demand and mental demand, as well as between physical demand and effort, with positive correlation coefficients of 0.81 and 0.80 , respectively. The same subscales were also highly correlated for the hard level with positive correlation coefficients of 0.84 and 0.71 , respectively. For (2), the highest correlation for the easy level was found between temporal demand and effort, with a positive correlation coefficient of 0.70 , followed by a coefficient of 0.69 for the correlation between physical demand and effort. The highest correlation was between physical demand and effort for the hard level, with $r = 0.61$. For (3), we found the highest correlation was between physical demand and effort for all four multitasking conditions, with positive correlation coefficients of 0.70 , 0.61 , 0.64 , and 0.70 , respectively. Fig. 9 shows the correlation plot for MR multitasks.

6.2. Accuracy

The accuracy for single RW and VW tasks was found to be significant from the Mann-Whitney test results, with p -values $<.001$ and effect sizes of 21.4% and 20.67% , respectively. The accuracy measure was also significant for RW-VW multitasking from the results of the Kruskal-Wallis test, with a $p < .0001$ and an effect size of 19.06% . A post hoc test was performed on the accuracy, where three pairwise combinations were found significant: (1) easy RW & easy VW-hard RW & hard VW, (2) easy RW & hard VW-hard RW & hard VW, and (3) hard RW & easy VW-hard RW & hard VW, with p -values $<.001$, $.03$, and $.04$ respectively. Fig. 10 shows the boxplots of the significant accuracies for all three experimental manipulations, where the hard level is higher than the easy level.

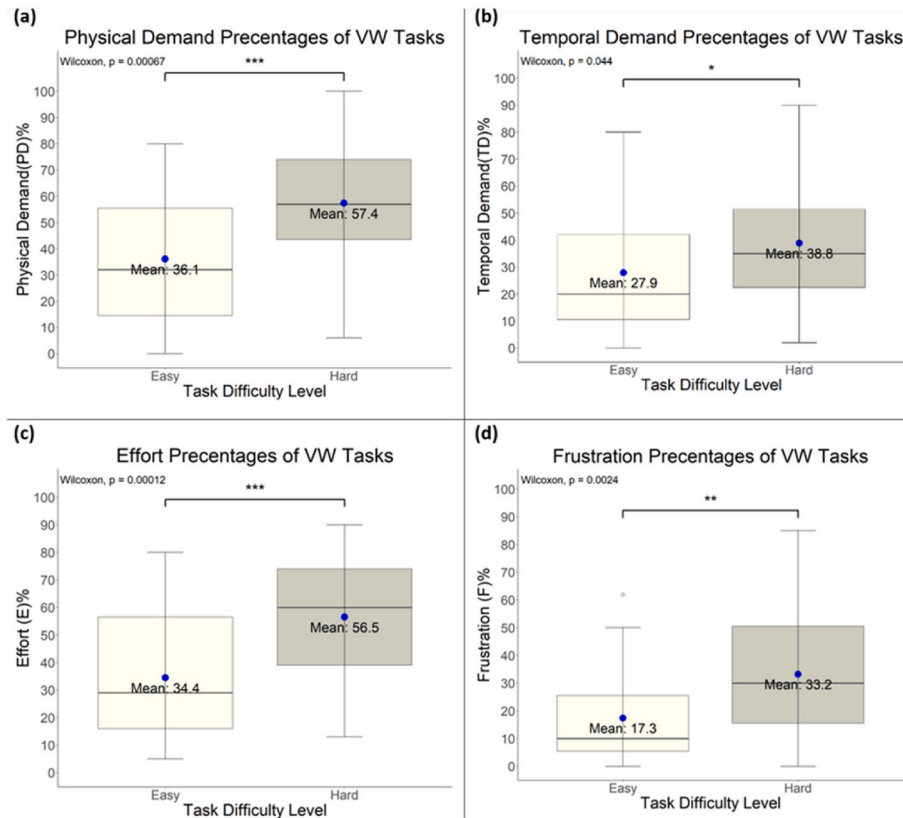


Fig. 6. Significant dependent measure of NASA-TLX dimensions for single digital (virtual) world tasks.

Fig. 6 Alt Text: Four boxplots showing task difficulty level (easy and hard) on the x-axis and the percentages of the significant NASA-TLX subscales on the y-axis: (a) PD% data, (b) TD% data, (c) E% data, and (d) F% data (ns: $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

Table 2
Significant NASA-TLX subscales.

Experimental manipulation	Significant NASA-TLX subscales, p -value <.05	Effect size %
Single RW tasks	Physical demand (PD)	9.4
	temporal demand (TD)	6.62
	Effort (E)	5.82
	Frustration (F)	10.56
Single VW tasks	Physical demand (PD)	16.6
	Temporal demand (TD)	5.83
	Effort (E)	21.90
	Frustration (F)	13.21
RW-VW multitasking	Physical demand (PD)	10
	Effort (E)	7.25

6.3. Total completion time

We found a significant difference in TCT with a p -value of $<.001$ from the results of the Mann-Whitney test. The task difficulty level accounted for 40.3% of the variance in the TCT scores, indicating a substantial effect size. Fig. 11 shows the boxplots of the TCT scores for the single physical (real) world tasks.

6.4. Priority

A total of 144 data were collected for priority (4 multitasking conditions $\times$ 36 participants). Fig. 12 represents the answers to the additional question we asked each participant in the post-task survey about the priority variable. Each participant performed four multitasking conditions in a random order of task difficulty level. After each multitasking condition, participants were asked which task they prioritized from the following options: prioritizing only the real-world task,

prioritizing only the virtual-world task, or prioritizing both tasks equally. These answers were represented in three different sections in this figure. We recorded the number of subjects who selected each task prioritization answer for each multitasking condition. At maximum, for easy RW & easy VW tasks condition, 17 participants equally prioritized both real-world and virtual-world tasks. For the easy RW & hard VW tasks condition, 16 participants prioritized working on the real-world task, and the remaining 20 participants did not prioritize this task. For the hard RW & easy VW tasks condition, 17 participants prioritized working on both tasks. For the hard RW & hard VW tasks condition, most participants (=19) prioritized working on the real-world task. We can note from the results of Fig. 12 that participants placed their highest priority on either the real-world task alone or both tasks. This leaves the virtual-world task by itself to be the least prioritized. Such results could be beneficial in indicating the level of importance of each task when designing mixed reality applications with different difficulty levels. However, to answer RQ2, we conclude that the RW-VW task priority perception of mixed reality multitasking is not affected by task difficulty level because of its insignificance in the Chi-squared test.

6.5. Hand use

A correlation analysis was performed between self-reported and observed hand use to answer RQ3. At first, three correlation tests were performed based on the division of the experiment: a single physical (real) world task, a single digital (virtual) world task, and multitasking between both. Both p -values of the correlation analysis of single real and virtual tasks were less than .001, with correlation coefficients of 0.59 and 0.36, respectively. This represents a moderate positive correlation between self-reported and observed hand use when performing only a real-world task and a weak correlation when performing only a virtual-

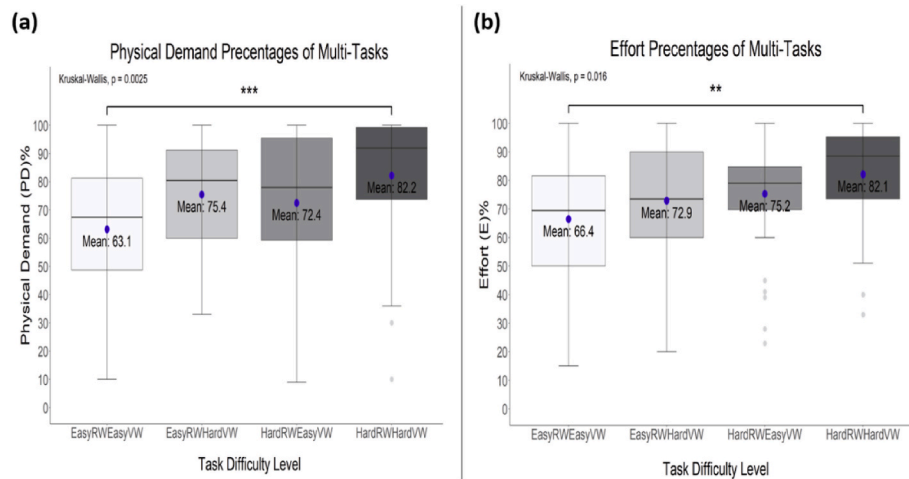


Fig. 7. Significant dependent measure of NASA-TLX dimensions for RW-VW multitasking conditions.

Fig. 7 Alt Text: Two boxplots showing task difficulty level on the x-axis and the percentages of the significant NASA-TLX subscales on the y-axis: (a) PD% data and (b) E% data (ns: $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

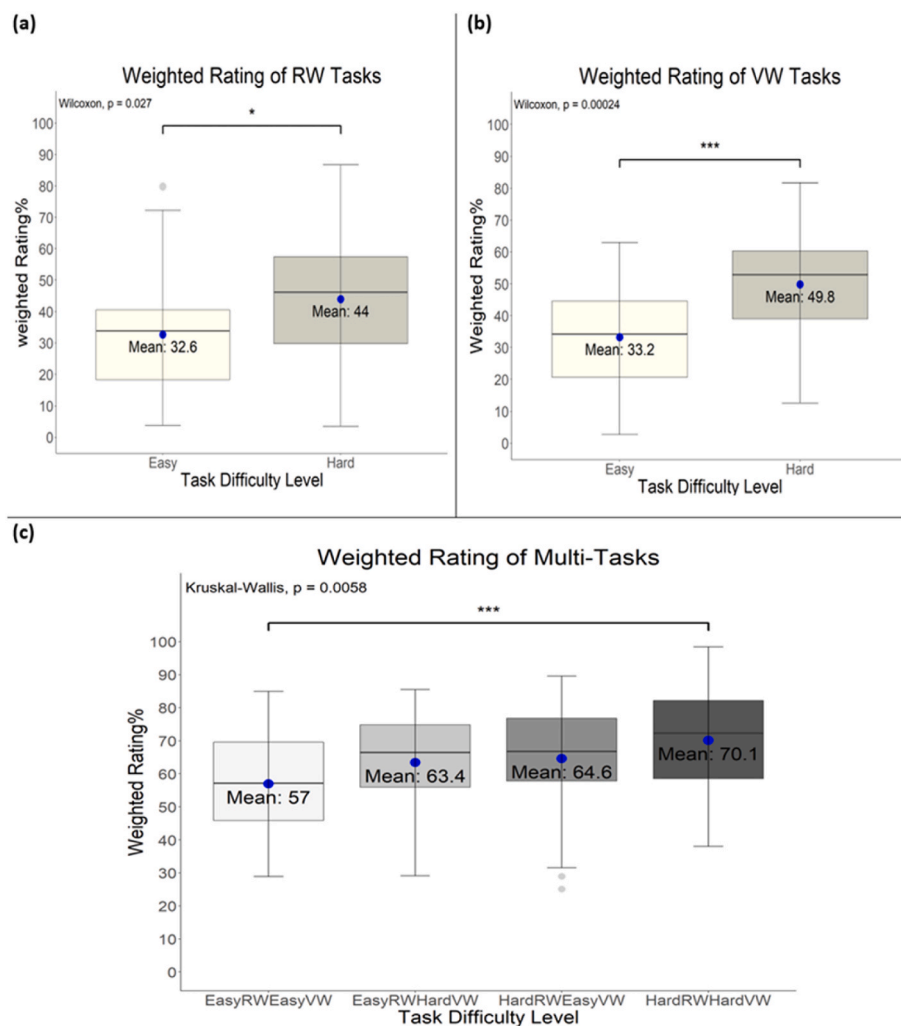


Fig. 8. Significant weighted rating measure of RW-VW multitasking conditions.

Fig. 8 Alt Text: Three boxplots showing task difficulty level on the x-axis and the percentages of the significant weighted rating on the y-axis: (a) for RW tasks, (b) for VW tasks, and (c) for RW-VW multitasking (ns: $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

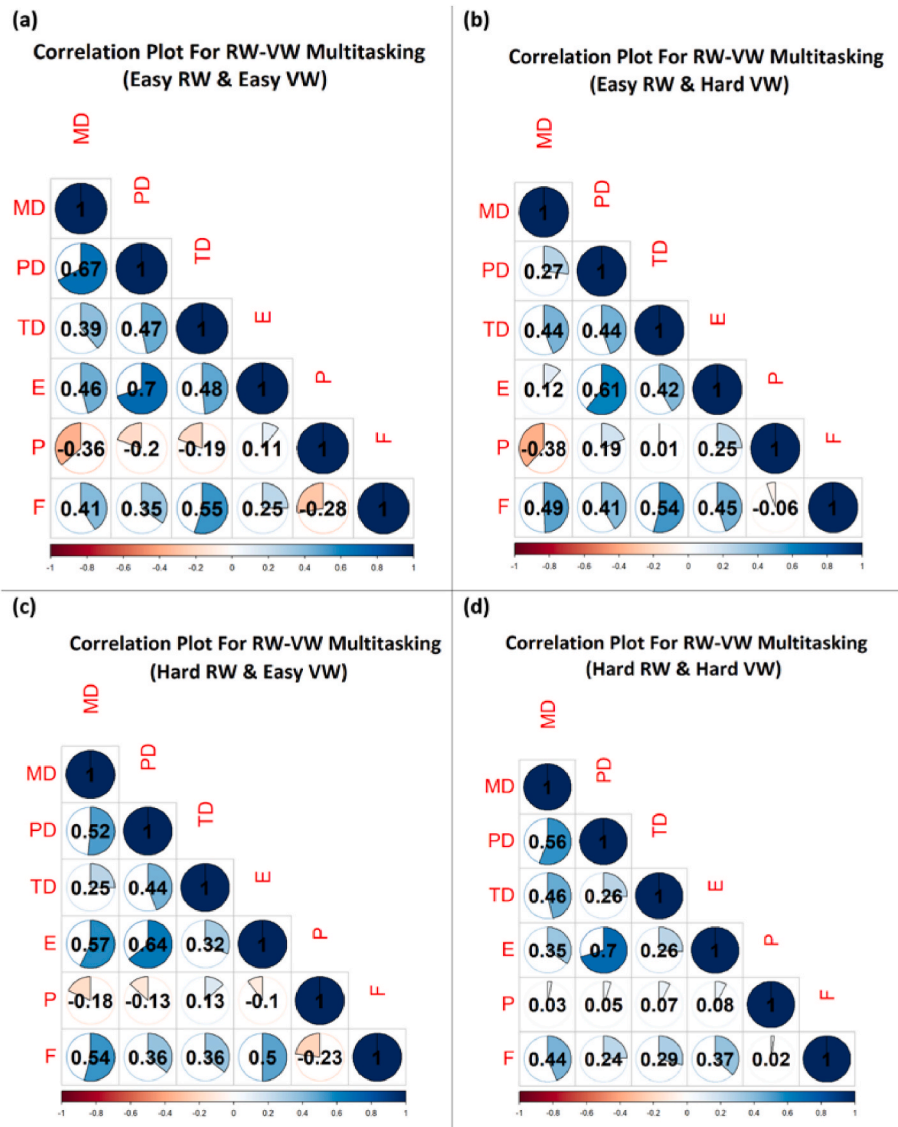


Fig. 9. Correlation plots of NASA-TLX subscales for RW-VW multitasking

Fig. 9 Alt Text: Four plots showing correlation coefficient values between NASA-TLX subscales in each circle: for (a) easy RW & easy VW, (b) easy RW & hard VW, (c) hard RW & easy VW, and (d) hard RW & hard VW.

world task. Comparing these results with multitasking conditions, we found a strong positive relationship between self-reported and observed hand use with $r = 0.7$ and $p < .001$. Considering each task's difficulty level for the correlation analysis, the results did not indicate significant differences with each experimental manipulation's general correlation analysis. Single real-world tasks resulted in a moderate positive relationship for both levels, easy and hard, with correlation coefficients of 0.6 and 0.58, respectively. Single virtual-world tasks resulted in weak positive relationships for both levels, with correlation coefficients of 0.4 and 0.3. The correlation relationship of the hand use variable for the multitasking conditions was strong, except for the highest multitasking difficulty level, namely hard RW & hard VW, which was moderate with a correlation coefficient of 0.6. Participants perceived their hand usage better during MR multitasking conditions than when performing a single task, either physically or digitally. Fig. 13 shows the bar plots of this variable in MR multitasking, separated by each task difficulty level. We can see that almost all participants used both hands for working on physical tasks, while the right hand was the most dominant when performing the virtual task. Such information would be helpful for MR application designers to account for the dominant hand of their users.

However, we have shown that the hand-use variable in our experiment has affected the physical demand of our participants. This is the cut-off point where MR applications involving their users interacting with their hands could be designed to reduce their users' physical demand.

7. Discussion

The present study aimed to examine the effect of task difficulty on human performance in multitasking between physical and digital tasks. NASA-TLX and accuracy measured performance. Subjective workload and accuracy are used as a measure of workload under specific time pressures (Wang et al., 2016). Thus, workload depends upon each individual and is related to task complexity (De Waard & Brookhuis, 1996). Other variables included total completion time, multitasking priority, and hand use. We were interested in studying task difficulty as a factor as it has been shown evidence to affect task performance for several decades generally (Campbell & Ilgen, 1976; Dunn & Williamson, 2012; Huber, 1985; Radeborg et al., 1999) and in multitasking particularly (Adler & Benbunan-Fich, 2015; Adler et al., 2015; Lansdown et al., 2004; Liu et al., 2016). Since increasing the task complexity would

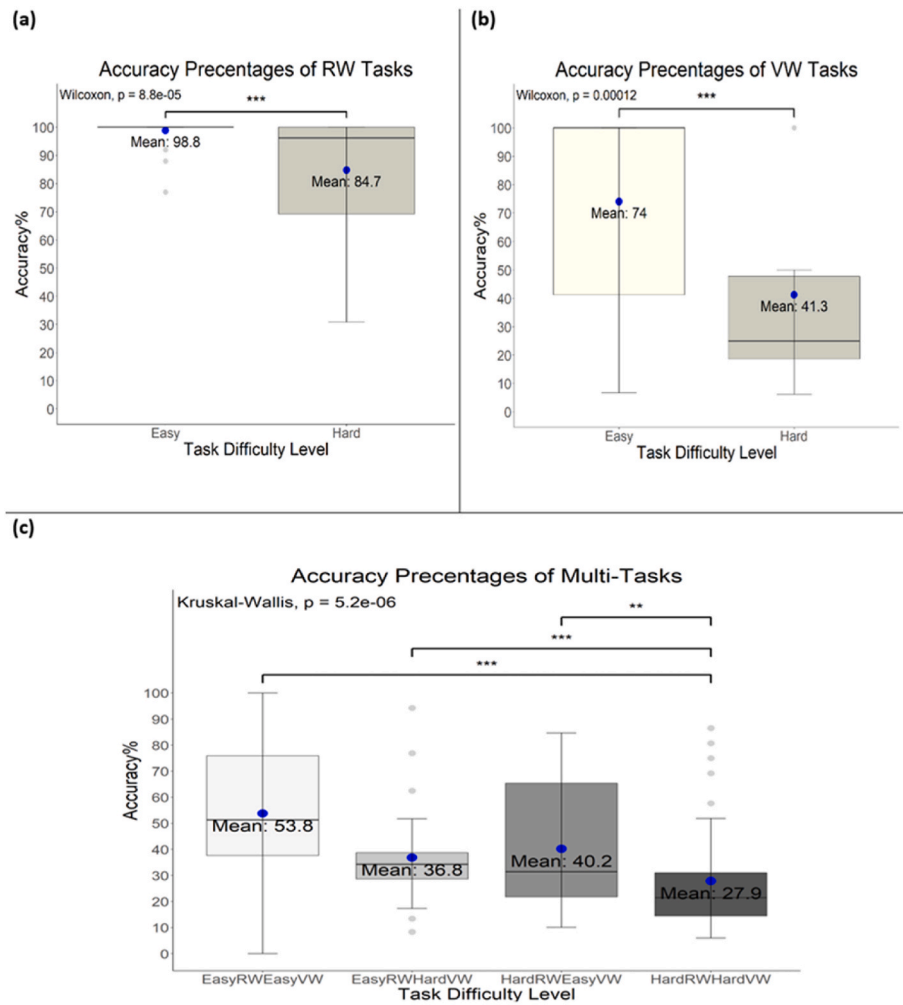


Fig. 10. Significant accuracy measures:

Fig. 10 Alt Text: Three boxplots showing task difficulty level on the x-axis and the percentages of task performance accuracy on the y-axis: (a) accuracy% for RW tasks, (b) accuracy% for VW tasks, and (c) accuracy% for MR multitasking (ns: $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

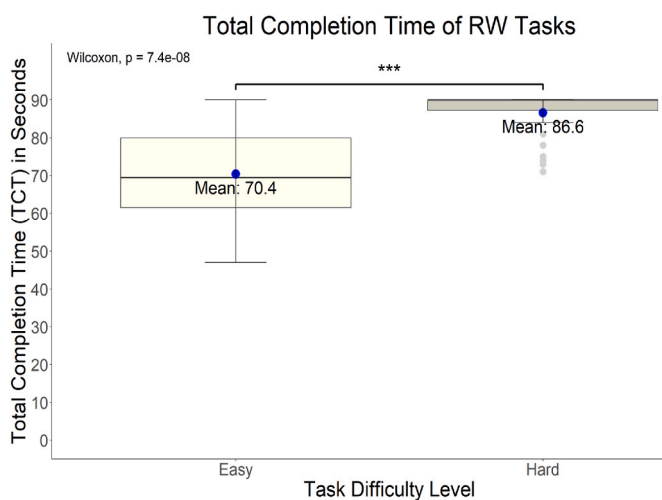


Fig. 11. TCT scores of single physical (real) world task.

Fig. 11 Alt Text: A boxplot showing task difficulty level (easy and hard) on the x-axis and the total completion time on the y-axis: (a) PD% data, (b) TD% data (ns: $p > 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

make performance maintenance challenging (Liu et al., 2016), positive relationships were found between task difficulty and mental workload measures (Pagnotta et al., 2022).

In our study, concerning RQ-1 (a) and (b), physical demand and effort were significant in all the experimental conditions in this study. Temporal demand and frustration were significant across only RW and VW tasks. The exact significance between the two conditions was found: easy RW & easy VW, hard RW & hard VW. Physical demand and effort were highly correlated in all four multitasking conditions and single virtual tasks. Meanwhile, temporal demand was highly correlated with effort in single physical tasks. Accuracy measure was found significant for all the experimental manipulations, (1) single RW task, (2) single VW task, and (3) MR multitasking. Participants performed their best at single RW tasks, with an average accuracy of 91.76%, and their worst at multitasking, with an average accuracy of 39.68%.

Participants' performance was measured separately in monotasking physical and digital tasks and multitasking in mixed reality. The comparison between monotasking and multitasking has been made in this study to conclude whether the results of performing a monotask physically or virtually would be consistent with multitasking the same tasks. However, out of the six NASA-TLX dimensions, only the physical demand and effort were constant in their significance across all experimental manipulations. Such results are helpful for future research to conclude if performance could be degraded or not if users, for example, should monotask over multitask or vice versa. In a similarly conducted

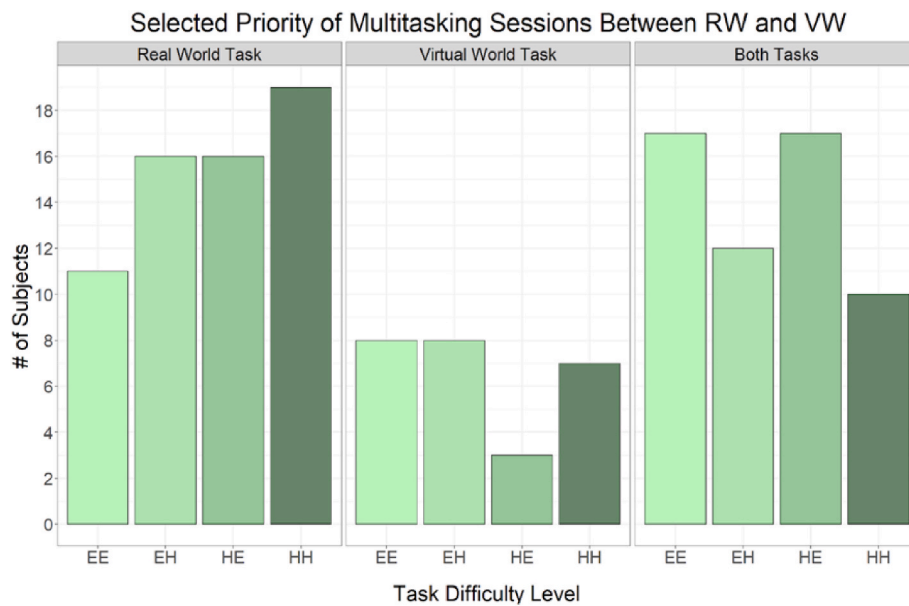


Fig. 12. Selected priority of multitasking between RW-VW

Fig. 12 Alt Text: A bar graph showing task difficulty level on the x-axis (EE = easy RW & easy VW; EH = easy RW & hard VW; HE = hard RW & easy VW; and HH = hard RW & hard VW) and the number of users' choices in their priority between the levels on the y-axis.

study by Xi et al. (2022), participants performed a shopping-related task with and without using augmented and virtual realities involving navigation, hand interaction, and information processing. The study yielded consistent results, aligning with our findings, where only the physical demand and effort, measured by NASA-TLX, were significant. At the same time, the other dimensions showed no significant effect in the combination of augmented and virtual realities. This occurred because users had to expend considerable physical resources, such as head and hand movements, to interact with augmented information and separate content in the real environment. A similar explanation could account for the results found in this study.

Total completion time measured only one experimental condition, single RW tasks. As explained earlier, this is because of the built-in automatic timer in the VW task, which continued until reaching the 90-s limit. However, although participants had the same time limit in single RW tasks, some could have finished the task in less time. Therefore, the TCT could be accounted for only in this case. We found that the task difficulty of single RW tasks (easy vs. hard) significantly differed in participants' completion time. The difference in completion time with varying difficulty levels (easy vs. hard) can be attributed to the distinct challenges posed by each difficulty level. Participants likely encountered greater complexity, cognitive load, or obstacles in the hard RW tasks than the easy ones, leading to variations in completion times. In a related study, Alessa et al. (2023) investigated the impact of task difficulty (simple vs. complex) on task performance in terms of completion time, focusing on interactions between instructions provided via augmented reality and paper-based guidance. Their findings reveal that task difficulty significantly influenced the total time, similar to our findings. Completion times showed a notable increase from simple to complex tasks across all instructional types.

When individuals experience time pressure, as suggested by van der Kleij et al. (2009) and Kelly and Loving (2004), it often leads to faster performance rates, lower performance quality, and increased heuristic information processing. Thus, the perception of time pressure to complete a task affects performance significantly. It has been demonstrated that working under time constraints adversely affects performance across various task types, including skill learning (Ferrari, 2001), decision-making (Kocher & Sutter, 2004; Payne et al., 1993), and information processing (Maule et al., 2000). The efficiency theory

(Eysenck & Calvo, 1992) posits that as task difficulty increases, such as under time pressure, more cognitive resources are needed to complete the task effectively. This explains the adverse effects of time pressure on task performance. In digital environments, the time spent on a task is considered an important feature (De Boeck & Jeon, 2019; Naumann, 2019).

To answer RQ-2, the multitasking priority measure was found to be insignificant. Participants performed both physical and digital tasks simultaneously to conclude the results of this variable. For all four multitasking conditions, we found that participants mostly prioritized either the real-world or both tasks, leaving the virtual-world task to be the least important. The priority variable answers were collected from the survey and distributed to the participants after each experimental condition. The correlation between self-reported and observed hand use in multitasking conditions was positively strong, with a correlation coefficient of $r > 0.07$ (answering RQ-3). Participants' perception of their dominant hand was consistent with the results of performing a VW task in multitasking, where almost all participants (=34) reported they were righthanded. On the other hand, it was inconsistent with performing RW tasks in multitasking. Most participants used both hands when working on the RW task in multitasking.

According to McGrath and Kantak (2016), the handedness variable could influence performance. Moreover, cognitive abilities significantly differed between right and left-handed individuals (Al-Hashel et al., 2016). Previous studies have supported that performance can vary between right and left-handed individuals. For example, Peters and Ivanoff (1999) demonstrated that left-handed women react faster to auditory stimuli with their left hand than right-handed women with their right hand. Similarly, Dane and Erzurumluoglu (2003) found that left-handed handball players were faster than their right-handed counterparts.

Güntürkün et al. (2020) explained that the difference in hand dominance in individuals is connected to both the right and left hemispheres of the brain, influencing behavior that affects sensory, cognitive, and motor efficiency. The primary motor cortex of the brain controls voluntary movements in the body, including those of the hands, through contralateral control, an organization of the nervous system in which one side of the brain controls and communicates with the opposite side of the body. As a result, right-handed individuals often exhibit left-hemisphere dominance for motor control, as indicated by the

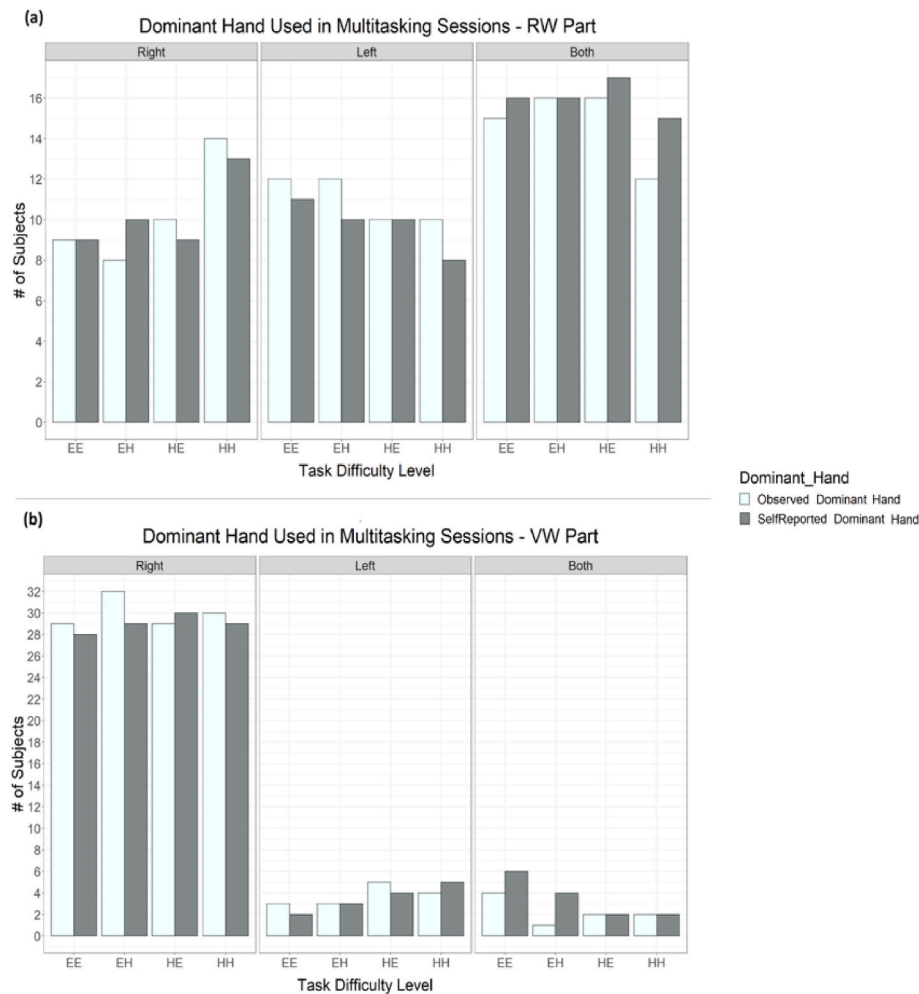


Fig. 13. Dominant hand used in multitasking conditions

Fig. 13 Alt Text: Two bar graphs showing task difficulty level on the x-axis (EE = easy RW & easy VW; EH = easy RW & hard VW; HE = hard RW & easy VW; and HH = hard RW & hard VW) and the number of users' choices in their dominant hand used in multitasking conditions on the y-axis: (a) RW part of the multitasking, and (b) VW part of the multitasking.

findings of Güntürkün et al. (2020). Güntürkün et al. (2020) added that brain lateralization, which refers to the functional differences between the left and right hemispheres of the brain, where cognitive functions are more strongly associated with one hemisphere than the other, affects the movements of the dominant hand depending on the task demands. This is because the hemispheric functional differences influence it in various individuals. However, brain asymmetry is shaped by the complex interaction of genetic, environmental, and epigenetic factors. Johnstone et al. (2021) supported the same understanding that brain lateralization asymmetry influences handedness among different individuals.

In educational environments, task difficulty plays a pivotal role. For example, Lynch et al. (2019) explored the relationship between homework task difficulty and student engagement, revealing significant differences between student groups assigned tasks of lesser and greater difficulty. The study suggested that task difficulty could impact student engagement, as their involvement with a task depended on their perceived ability to succeed. In particular, engagement decreased with an increase in the level of difficulty. Pavlov et al. (2021) matched this finding, demonstrating that task difficulty significantly affected students' engagement and learning rates. In this study, fourth-grade students performed better academically on math assignments when the task conditions closely matched their instructional skill levels. Differentiated instruction creates better learning opportunities for students by providing an appropriate match to their individual needs, improving

their academic outcomes and engagement (Haydon, 2012; Noltemeyer et al., 2019). This highlights the importance of considering task difficulty when designing learning tasks to provide learners with appropriate opportunities to experience success.

Several limitations were present in this study. Currently, our results focused on young university students; thus, the first limitation is having similar groups of participants, which might have been a barrier to generalizing the results. Second, the TCT variable could not be considered for all the experimental conditions because the built-in virtual timer in the n-back test application allows participants to complete the task forcibly for 90 s. Third, short interviews/questions could be performed on participants to understand their decisions toward prioritizing the presented tasks. Fourth, only two task difficulty levels were assigned: easy vs. hard. Increasing the number of difficulty levels might provide more detailed results for our experiment. Lastly, handedness was measured using a single self-report rather than a comprehensive survey instrument. In a future study, it might be interesting to collect more demographic data to see links between this data and performance in MR multitasking. In addition, developing AR medical applications can be useful to help patients in need of a better cognitive assessment.

We conclude that the results of this study provide a basis for future research to explore the effect of mixed-reality multitasking on performance. The goal often involves ergonomists and designers utilizing information technology to create technology systems that reduce

workload or at least keep it within an acceptable range. Workload management is crucial for user acceptance and performance (Dang et al., 2020). However, with increasing task complexity or difficulty, the perceived workload escalates, and if the acceptable level is exceeded, performance could suffer. Given this reasoning, theories like task-technology fit (Goodhue & Thompson, 1995) have gained significant prominence in human-computer interaction over the past decades. However, further research is needed to expand the effect of difficulty level as a factor in mixed reality environments for a larger and more diverse sample size.

8. Conclusion

This study explored the performance effectiveness of MR multitasking for different task difficulty levels. Physical demand, effort, and accuracy were significant compared to the other performance metrics. The findings of this research carry substantial implications for designers working in mixed reality. As revealed by this study, understanding the correlation between task difficulty and user performance provides a foundation for the development of predictive models. Such models can prove valuable for designers, offering insights that aid in anticipating user navigation and task performance within mixed-reality environments, especially when faced with varying levels of task difficulty.

Some pedagogical implications of this study involve informing instructional designers about the critical role of task difficulty in developing mixed reality educational content. This knowledge empowers designers to tailor tasks that align with the cognitive abilities of learners, ensuring an optimal learning experience by designing tasks that are challenging yet manageable. Researchers have highlighted the importance of addressing the varied needs of learners and creating personalized learning environments to promote active learning (Jung & Latchem, 2011; O'Donnell et al., 2015).

Determining the generalizability of immersive reality's effectiveness as a pedagogical method is challenging, given significant differences in age, task difficulty, and applications (Hamilton et al., 2020). Therefore, emphasis can be placed on creating user-friendly interfaces aligned with learners' cognitive abilities or implementing systems that dynamically adjust task difficulty based on individual or group performance. Considering learner characteristics such as age, this approach not only enhances engagement but could also improve learning outcomes in mixed-reality settings.

Funding

This research was supported by the College of Engineering at the University of Illinois at Chicago and the first author's Saudi Arabian Cultural Mission (SACM) fellowship.

CRedit authorship contribution statement

Safanah Abbas: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Funding acquisition, Formal analysis. **Heejin Jeong:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors thank Edison Larco for developing the experiment interface and Bryan Havens and Yalda Ghasemi for providing initial insights to collect the data.

References

- Abbas, S., & Jeong, H. (2024). Unveiling gender differences: A mixed reality multitasking exploration. *Frontiers in Virtual Reality*, 4, Article 1308133.
- Adler, R. F., & Benbunan-Fich, R. (2012). Juggling on a high wire: Multitasking effects on performance. *International Journal of Human-Computer Studies*, 70(2), 156–168. <https://doi.org/10.1016/j.ijhcs.2011.10.003>
- Adler, R. F., & Benbunan-Fich, R. (2015). The effects of task difficulty and multitasking on performance. *Interacting with Computers*, 27(4), 430–439.
- Adler, R. F., Rubin, D., Mohammad, A. R., Irfan, A., Senadeera, H., & Nguyen, T. (2015). That's easy! the effects of objective and subjective task difficulty when multitasking. *Procedia Manufacturing*, 3, 5467–5474. <https://doi.org/10.1016/j.promfg.2015.07.684>
- Al-Hashel, J. Y., Ahmed, S. F., Al-Mutairi, H., Hassan, S., Al-Awadhi, N., & Al-Saraji, M. (2016). Association of cognitive abilities and brain lateralization among primary school children in Kuwait. *Neuroscience Journal*, 2016, 1–5. <https://doi.org/10.1155/2016/6740267>
- Alessa, F. M., Alhaag, M. H., Al-harkan, I. M., Nasr, M. M., Kaid, H., & Hammami, N. (2023). Evaluating physical stress across task difficulty levels in augmented reality-assisted Industrial Maintenance. *Applied Sciences*, 14(1), 363. <https://doi.org/10.3390/app14010363>
- Anmarkrud, Ø., Andresen, A., & Bråten, I. (2019). Cognitive load and working memory in Multimedia learning: Conceptual and measurement issues. *Educational Psychologist*, 54(2), 61–83. <https://doi.org/10.1080/00461520.2018.1554484>
- Annett, J., Annett, M., Hudson, P. T., & Turner, A. (1979). The control of movement in the preferred and non-preferred hands. *Quarterly Journal of Experimental Psychology*, 31(4), 641–652. <https://doi.org/10.1080/14640747908400755>
- Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Baddeley, A. D., & Della Sala, S. (1996). Working memory and executive control. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1346), 1397–1404.
- Bednar, P. M., & Welch, C. (2020). Socio-technical perspectives on smart working: Creating meaningful and Sustainable Systems. *Information Systems Frontiers*, 22(2), 281–298. <https://doi.org/10.1007/s10796-019-09921-1>
- Bonner, S. E. (1994). A model of the effects of audit task complexity. *Accounting, Organizations and Society*, 19(3), 213–234. [https://doi.org/10.1016/0361-3682\(94\)90033-7](https://doi.org/10.1016/0361-3682(94)90033-7)
- Bourque, A. (2017). Smart glasses are making workers more productive. *CIO*. Retrieved February 11, 2022, from <https://www.cio.com/article/235049/smart-glasses-are-making-workers-more-productive.html>
- Bowers, M. A., & Christensen, J. C. (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>
- Brade, J., Lorenz, M., Busch, M., Hammer, N., Tscheligi, M., & Klimant, P. (2017). Being there again – presence in real and virtual environments and its relation to usability and user experience using a mobile navigation task. *International Journal of Human-Computer Studies*, 101, 76–87. <https://doi.org/10.1016/j.ijhcs.2017.01.004>
- Bridenbaugh, S. A., & Kressig, R. W. (2015). Motor cognitive dual tasking. *Zeitschrift für Gerontologie und Geriatrie*, 48(1), 15–21. <https://doi.org/10.1007/s00391-014-0845-0>
- Bruggen, A. (2015). An empirical investigation of the relationship between workload and performance. *Management Decision*, 53(10), 2377–2389. <https://doi.org/10.1108/md-02-2015-0063>
- Bryden, P. J., & Huszczyński, J. (2011). Under what conditions will right-handers use their left hand? The effects of object orientation, object location, arm position, and task complexity in preferential reaching. *Laterality: Asymmetries of Body, Brain and Cognition*, 16(6), 722–736. <https://doi.org/10.1080/1357650x.2010.514344>
- Bryden, P. J., & Roy, E. A. (2006). Preferential reaching across regions of hemispace in adults and children. *Developmental Psychobiology*, 48(2), 121–132. <https://doi.org/10.1002/dev.20120>
- Bryden, P. J., Roy, E. A., Rohr, L. E., & Egilo, S. (2007). Task demands affect manual asymmetries in Pegboard performance. *Laterality: Asymmetries of Body, Brain and Cognition*, 12(4), 364–377. <https://doi.org/10.1080/13576500701356244>
- Campbell, D. J., & Ilgen, D. R. (1976). Additive effects of task difficulty and goal setting on subsequent task performance. *Journal of Applied Psychology*, 61(3), 319–324. <https://doi.org/10.1037/0021-9010.61.3.319>
- Card, S. K., Moran, T. P., & Newell, A. (2018). Applying psychology to design. *The Psychology of Human-Computer Interaction*, 403–424. <https://doi.org/10.1201/9780203736166-12>
- Chen, Y.-N., Mitra, S., & Schlaghecken, F. (2008). Sub-processes of working memory in the N-back task: An investigation using ERPs. *Clinical Neurophysiology*, 119(7), 1546–1559. <https://doi.org/10.1016/j.clinph.2008.03.003>
- Cho, M. (2018). Task complexity, modality, and working memory in L2 task performance. *System*, 72, 85–98. <https://doi.org/10.1016/j.system.2017.10.010>
- Choi, S. H., Park, K. B., Roh, D. H., Lee, J. Y., Mohammed, M., Ghasemi, Y., & Jeong, H. (2022). An integrated mixed reality system for safety-aware human-robot collaboration using deep learning and digital twin generation. *Robotics and Computer-Integrated Manufacturing*, 73, Article 102258. <https://doi.org/10.1016/j.rcim.2021.102258>
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3), 98–101. <https://doi.org/10.1111/1467-8721.ep10768783>
- Cohen, R. S., & Lefkowitz, J. (1977). Self-esteem, locus of control, and task difficulty as determinants of Task Performance. *Journal of Vocational Behavior*, 11(3), 314–321. [https://doi.org/10.1016/0001-8791\(77\)90027-6](https://doi.org/10.1016/0001-8791(77)90027-6)

- Colle, H. A., & Reid, G. B. (1998). Context effects in subjective mental workload ratings. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(4), 591–600. <https://doi.org/10.1518/001872098779649283>
- Coutrot, A., Schmidt, S., Coutrot, L., Pittman, J., Hong, L., Wiener, J. M., Hölscher, C., Dalton, R. C., Hornberger, M., & Spiers, H. J. (2019). Virtual navigation tested on a mobile app is predictive of real-world wayfinding navigation performance. *PLoS One*, 14(3), 1–15. <https://doi.org/10.1371/journal.pone.0213272>
- Cramond, H. J., Clark, M. S., & Smith, D. S. (1989). The effect of using the dominant or nondominant hand on performance of the Rivermead Perceptual Assessment Battery. *Clinical Rehabilitation*, 3(3), 215–221. <https://doi.org/10.1177/026921558900300306>
- Dönmez, O., & Akbulut, Y. (2021). Timing and relevance of secondary tasks impact multitasking performance. *Computers & Education*, 161, Article 104078. <https://doi.org/10.1016/j.compedu.2020.104078>
- Dai, F., Olorunfemi, A., Peng, W., Cao, D., & Luo, X. (2021). Can mixed reality enhance safety communication on construction sites? An industry perspective. *Safety Science*, 133, 1–12. <https://doi.org/10.1016/j.ssci.2020.105009>
- Dane, S., & Erzurumluoglu, A. (2003). Sex and handedness differences in eye-hand visual reaction times in handball players. *International Journal of Neuroscience*, 113(7), 923–929. <https://doi.org/10.1080/00207450390220367>
- Dang, Y. M., Zhang, Y. G., Brown, S. A., & Chen, H. (2020). Examining the impacts of mental workload and task-technology fit on user acceptance of the Social Media Search System. *Information Systems Frontiers*, 22(3), 697–718. <https://doi.org/10.1007/s10796-018-9879-y>
- De Boeck, P., & Jeon, M. (2019). An overview of models for response times and processes in cognitive tests. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00102>
- De Waard, D., & Brookhuis, K. A. (1996). *The measurement of drivers' mental workload*. Dinth, H. Q., Walker, N., Hodges, L. F., Song, C., & Kobayashi, A. (1999). Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE virtual reality (Cat. No. 99CB36316)* (pp. 222–228). <https://doi.org/10.1109/vr.1999.756955>
- Drascic, D., & Milgram, P. (1996). Perceptual issues in augmented reality. *Stereoscopic Displays and Virtual Reality Systems*, III(2653), 123–135.
- Dubois, E., Gray, P., & Nigay, L. (Eds.). (2009). *The engineering of mixed reality systems*. Springer Science & Business Media.
- Dunn, N., & Williamson, A. (2012). Driving monotonous routes in a train simulator: The effect of task demand on driving performance and subjective experience. *Ergonomics*, 55(9), 997–1008. <https://doi.org/10.1080/00140139.2012.691994>
- Dzubak, C. M. (2008). Multitasking: The good, the bad, and the unknown. *The Journal of the Association for the Tutoring Profession*, 1(2), 1–12.
- Ettinger, K., & Cohen, A. (2020). Patterns of multitasking behaviours of adolescents in digital environments. *Education and Information Technologies*, 25(1), 623–645. <https://doi.org/10.1007/s10639-019-09982-4>
- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, 6(6), 409–434. <https://doi.org/10.1080/02699939208409696>
- Ferrari, J. R. (2001). Procrastination as self-regulation failure of performance: Effects of Cognitive Load, self-awareness, and time limits on 'working best under pressure'. *European Journal of Personality*, 15(5), 391–406. <https://doi.org/10.1002/per.413>
- Fleishman, E. A. (1957). Factor structure in relation to task difficulty in psychomotor performance. *Educational and Psychological Measurement*, 17(4), 522–532. <https://doi.org/10.1177/001316445701700408>
- Fraser, S. A., & Li, K. Z. (2012). Dual-task performance in motor learning. In *Encyclopedia of the sciences of learning* (pp. 1048–1050). https://doi.org/10.1007/978-1-4419-1428-6_1703
- Güntürkün, O., Ströckens, F., & Ocklenburg, S. (2020). Brain lateralization: A comparative perspective. *Physiological Reviews*, 100(3), 1019–1063. <https://doi.org/10.1152/physrev.00006.2019>
- Gabbard, C., Tapiá, M., & Helbig, C. R. (2003). Task complexity and limb selection in reaching. *International Journal of Neuroscience*, 113(2), 143–152.
- Gaol, F. L., & Prasolova-Förland, E. (2021). Special section editorial: The Frontiers of Augmented and mixed reality in all levels of Education. *Education and Information Technologies*, 27(1), 611–623. <https://doi.org/10.1007/s10639-021-10746-2>
- Gardner, M. (2012). *Beginning R: The statistical programming language*. John Wiley & Sons.
- Gentili, R. J., Jaquess, K. J., Shuggi, I. M., Shaw, E. P., Oh, H., Lo, L., Tan, Y. Y., Domingues, C. A., Blanco, J. A., Rietschel, J. C., Miller, M. W., & Hatfield, B. D. (2018). Combined assessment of attentional reserve and cognitive-motor effort under various levels of challenge with a dry EEG system. *Psychophysiology*, 55(6). <https://doi.org/10.1111/psyp.13059>
- Gentili, R. J., Rietschel, J. C., Jaquess, K. J., Li-Chuan, L., Prevost, C. M., Miller, M. W., Mohler, J. M., Oh, H., Ying, Y. T., & Hatfield, B. D. (2014). Brain biomarkers based assessment of cognitive workload in pilots under various task demands. In *2014 36th annual international conference of the IEEE engineering in medicine and biology society*. <https://doi.org/10.1109/embc.2014.6944961>
- Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 19(2), 213. <https://doi.org/10.2307/249689>
- Grealy, M. A., Johnson, D. A., & Rushton, S. K. (1999). Improving cognitive function after brain injury: The use of exercise and virtual reality. *Archives of Physical Medicine and Rehabilitation*, 80(6), 661–667. [https://doi.org/10.1016/s0003-9993\(99\)90169-7](https://doi.org/10.1016/s0003-9993(99)90169-7)
- Greef, T. D., Lafeber, H., van Oostendorp, H., & Lindenberg, J. (2009). Eye movement as indicators of mental workload to trigger adaptive automation. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, 219–228. https://doi.org/10.1007/978-3-642-02812-0_26
- Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36(2), 212–224. <https://doi.org/10.3200/jmbr.36.2.212-224>
- Haga, S., Shinoda, H., & Kokubun, M. (2002). Effects of task difficulty and time-on-task on mental workload. *Japanese Psychological Research*, 44(3), 134–143. <https://doi.org/10.1111/1468-5884.00016>
- Hahnel, C., Schöor, C., Kroehne, U., Goldhammer, F., Mahlow, N., & Artelt, C. (2019). The role of cognitive load in university students' comprehension of multiple documents. *Zeitschrift für Pädagogische Psychologie*, 33(2), 105–118. <https://doi.org/10.1024/1010-0652/a000238>
- Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2020). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, 8(1), 1–32. <https://doi.org/10.1007/s40692-020-00169-2>
- Hancock, P. (2009). The future of simulation. *Human Factors in Simulation and Training*, 169–186. <https://doi.org/10.1201/9781420072846.ch9>
- Hancock, P. A., & Matthews, G. (2018). Workload and performance: Associations, insensitivities, and dissociations. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 61(3), 374–392. <https://doi.org/10.1177/0018720818809590>
- Hancock, P. A., Williams, G., & Manning, C. M. (1995). Influence of task demand characteristics on workload and performance. *The International Journal of Aviation Psychology*, 5(1), 63–86. https://doi.org/10.1207/s15327108ijap0501_5
- Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. *PsychEXTRA Dataset*, 904–908. <https://doi.org/10.1037/e577632012-009>
- Haydon, T. (2012). Using functional behavior assessment to match task difficulty for a 5th grade student: A case study. *Education & Treatment of Children*, 35(3), 459–476. <https://doi.org/10.1135/etc.2012.0019>
- Huber, V. L. (1985). Effects of task difficulty, goal setting, and strategy on performance of a heuristic task. *Journal of Applied Psychology*, 70(3), 492–504. <https://doi.org/10.1037/0021-9010.70.3.492>
- Jaeggi, S. M., Buschkuhl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of then-back task as a working memory measure. *Memory*, 18(4), 394–412. <https://doi.org/10.1080/09658211003702171>
- Jamet, E., Gonther, C., Cojean, S., Colliot, T., & Erhel, S. (2020). Does multitasking in the classroom affect learning outcomes? A naturalistic study. *Computers in Human Behavior*, 106, Article 106264. <https://doi.org/10.1016/j.chb.2020.106264>
- Jaquess, K. J., Lo, L.-C., Oh, H., Lu, C., Ginsberg, A., Tan, Y. Y., Lohse, K. R., Miller, M. W., Hatfield, B. D., & Gentili, R. J. (2018). Changes in mental workload and motor performance throughout multiple practice sessions under various levels of task difficulty. *Neuroscience*, 393, 305–318. <https://doi.org/10.1016/j.neuroscience.2018.09.019>
- Johnstone, L. T., Karlsson, E. M., & Carey, D. P. (2021). Left-handers are less lateralized than right-handers for both left and right hemispheric functions. *Cerebral Cortex*. <https://doi.org/10.1093/cercor/bhab048>
- Jones, L. A., & Lederman, S. J. (2006). *Human hand function*. Oxford university press.
- Joyner, J. S., Vaughn-Cooke, M., & Benz, H. L. (2021). Comparison of dexterous task performance in virtual reality and real-world environments. *Frontiers in Virtual Reality*, 2, 1–18. <https://doi.org/10.3389/fvrv.2021.599274>
- Jung, I., & Latchem, C. (2011). A model for e-education: Extended teaching spaces and extended learning spaces. *British Journal of Educational Technology*, 42(1), 6–18.
- Kaber, D. B., Li, Y., Clamann, M., & Lee, Y.-S. (2012). Investigating human performance in a virtual reality haptic simulator as influenced by fidelity and system latency. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 42(6), 1562–1566. <https://doi.org/10.1109/tsmca.2012.2201466>
- Kalyuga, S. (2007). Enhancing instructional efficiency of interactive e-learning environments: A cognitive load perspective. *Educational Psychology Review*, 19(3), 387–399. <https://doi.org/10.1007/s10648-007-9051-6>
- Kaplan, A. D., Cruik, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. A. (2021). The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 63(4), 706–726. <https://doi.org/10.1177/0018720820904229>
- Kaploun, K. A., & Abeare, C. A. (2010). Degree versus direction: A comparison of four handedness classification schemes through the investigation of lateralised semantic priming. *Laterality: Asymmetries of Body, Brain and Cognition*, 15(5), 481–500. <https://doi.org/10.1080/13576500902958871>
- Kaufman-Scarborough, C., & Lindquist, J. D. (1999). Time management and polychronicity. *Journal of Managerial Psychology*, 14(3/4), 288–312. <https://doi.org/10.1108/02683949910263819>
- Kelley, K., & Preacher, K. J. (2012). On effect size. *Psychological Methods*, 17(2), 137–152. <https://doi.org/10.1037/a0028086>
- Kelly, J. R., & Loving, T. J. (2004). Time pressure and group performance: Exploring underlying processes in the attentional focus model. *Journal of Experimental Social Psychology*, 40(2), 185–198. [https://doi.org/10.1016/s0022-1031\(03\)00094-5](https://doi.org/10.1016/s0022-1031(03)00094-5)
- Kim, J., Bayro, A., Lee, J., Soltis, I., Kim, M., Jeong, H., & Yeo, W. H. (2023). Mixed reality-integrated soft wearable biosensing glove for manipulating objects. *Biosensors and Bioelectronics X*, 14, Article 100343. <https://doi.org/10.1016/j.biosx.2023.100343>
- Kim, K.-J., Choi, M.-J., & Kim, K.-J. (2021). Effects of nursing simulation using mixed reality: A scoping review. *Healthcare*, 9(8), 1–11. <https://doi.org/10.3390/healthcare9080947>
- Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, 12(1), 1–10. [https://doi.org/10.1016/s0959-4752\(01\)00014-7](https://doi.org/10.1016/s0959-4752(01)00014-7)

- Kocher, M. G., & Sutter, M. (2004). Time is money - time pressure, incentives, and the quality of decision-making. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.505242>
- Kosson, D. S. (1996). Psychopathy and dual-task performance under focusing conditions. *Journal of Abnormal Psychology*, 105(3), 391–400. <https://doi.org/10.1037/0021-843x.105.3.391>
- Lal, B., Dwivedi, Y. K., & Haag, M. (2021). Working from home during covid-19: Doing and managing technology-enabled social interaction with colleagues at a distance. *Information Systems Frontiers*, 25(4), 1333–1350. <https://doi.org/10.1007/s10796-021-10182-0>
- Lansdown, T. C., Brook-Carter, N., & Kersloot, T. (2004). Distraction from multiple in-vehicle secondary tasks: Vehicle performance and mental workload implications. *Ergonomics*, 47(1), 91–104. <https://doi.org/10.1080/00140130310001629775>
- Larson, P., Rizzo, A. A., Buckwalter, J. G., van Rooyen, A., Kratz, K., Neumann, U., ... van der Zaag, C. (1999). Gender issues in the use of virtual environments. *CyberPsychology and Behavior*, 2(2), 113–123.
- Lee, K. (2012). Augmented reality in education and training. *TechTrends*, 56(2), 13–21. <https://doi.org/10.1007/s11528-012-0559-3>
- Lee, G. I., Lee, M. R., Clanton, T., Sutton, E., Park, A. E., & Marohn, M. R. (2014). Comparative assessment of physical and cognitive ergonomics associated with robotic and traditional laparoscopic surgeries. *Surgical Endoscopy*, 28(2), 456–465. <https://doi.org/10.1007/s00464-013-3213-z>
- Lee, J., Lin, L., & Robertson, T. (2012). The impact of media multitasking on learning. *Learning, Media and Technology*, 37(1), 94–104.
- Lee, H. Y., & List, A. (2021). Examining students' self-efficacy and perceptions of task difficulty in learning from multiple texts. *Learning and Individual Differences*, 90, Article 102052. <https://doi.org/10.1016/j.lindif.2021.102052>
- Li, Q., Lu, Z., Gao, N., & Yang, J. (2019). Optimizing the performance of the visual p300-speller through active mental tasks based on color distinction and modulation of Task Difficulty. *Frontiers in Human Neuroscience*, 13, 1–14. <https://doi.org/10.3389/fnhum.2019.00130>
- Lin, L. (2013). Multiple dimensions of multitasking phenomenon. *International Journal of Technology and Human Interaction*, 9(1), 37–49. <https://doi.org/10.4018/jthi.2013010103>
- Lindbeck, A., & Snower, D. J. (2000). Multitask learning and the reorganization of work: From Tayloristic to holistic organization. *Journal of Labor Economics*, 18(3), 353–376. <https://doi.org/10.1086/209962>
- Lindenberger, U., Marsiske, M., & Baltes, P. B. (2000). Memorizing while walking: Increase in dual-task costs from young adulthood to old age. *Psychology and Aging*, 15(3), 417–436. <https://doi.org/10.1037/0882-7974.15.3.417>
- Liu, J., Gwizdzka, J., Liu, C., & Belkin, N. J. (2010). Predicting task difficulty for different task types. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1–10. <https://doi.org/10.1002/meet.14504701173>
- Liu, P., & Li, Z. (2011). Toward understanding the relationship between task complexity and task performance. *Lecture Notes in Computer Science*, 192–200. https://doi.org/10.1007/978-3-642-21660-2_22
- Liu, S., & Nam, C. S. (2018). Quantitative modeling of user performance in multitasking environments. *Computers in Human Behavior*, 84, 130–140. <https://doi.org/10.1016/j.chb.2018.02.035>
- Liu, S., Wadeson, A., Kim, N. Y., & Nam, C. S. (2016). Effects of working memory capacity, task switching, and task difficulty on multitasking performance. *Proceedings of the Human Factors and Ergonomics Society - Annual Meeting*, 60(1), 502–506. <https://doi.org/10.1177/1541931213601114>
- Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. <https://doi.org/10.21236/ada086584>
- Lungu, A. J., Swinkels, W., Claesen, L., Tu, P., Egger, J., & Chen, X. (2020). A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: An extension to different kinds of surgery. *Expert Review of Medical Devices*, 18(1), 47–62. <https://doi.org/10.1080/17434440.2021.1860750>
- Luria, R., & Vogel, E. K. (2011). Visual search demands dictate reliance on working memory storage. *Journal of Neuroscience*, 31(16), 6199–6207. <https://doi.org/10.1523/jneurosci.6453-10.2011>
- Lynch, R., Hurley, A., Cumiskey, O., Nolan, B., & McGlynn, B. (2019). Exploring the relationship between homework task difficulty, student engagement and performance. *Irish Educational Studies*, 38(1), 89–103. <https://doi.org/10.1080/03323315.2018.1512889>
- MacFarland, T. W., & Yates, J. M. (2016). Nonparametric statistics for the biological sciences. *Introduction to Nonparametric Statistics for the Biological Sciences Using R*, 1–50. https://doi.org/10.1007/978-3-319-30634-6_1
- MacPherson, S. E. (2018). Definition: Dual-tasking and multitasking. *Cortex*, 106, 313–314. <https://doi.org/10.1016/j.cortex.2018.06.009>
- Martin, B. A., & Manning, D. J. (1995). Combined effects of normative information and task difficulty on the goal commitment-performance relationship. *Journal of Management*, 21(1), 65–80. <https://doi.org/10.1177/014920639502100104>
- Martiniussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44, 377–384.
- Mascha, M. F. (2001). The effect of task complexity and expert system type on the acquisition of Procedural Knowledge. *International Journal of Accounting Information Systems*, 2(2), 103–124. [https://doi.org/10.1016/s1467-0895\(01\)00016-1](https://doi.org/10.1016/s1467-0895(01)00016-1)
- Maule, A. J., Hockey, G. R. J., & Bdzola, L. (2000). Effects of time-pressure on decision-making under uncertainty: Changes in affective state and information processing strategy. *Acta Psychologica*, 104(3), 283–301. [https://doi.org/10.1016/s0001-6918\(00\)00033-0](https://doi.org/10.1016/s0001-6918(00)00033-0)
- Maume, D. J., & Purcell, D. A. (2007). The 'over-paced' American: Recent trends in the intensification of work. *Workplace Temporalities*, 251–283. [https://doi.org/10.1016/s0277-2833\(07\)17008-4](https://doi.org/10.1016/s0277-2833(07)17008-4)
- May, K. E., & Elder, A. D. (2018). Efficient, helpful, or distracting? A literature review of media multitasking in relation to academic performance. *International Journal of Educational Technology in Higher Education*, 15(1). <https://doi.org/10.1186/s41239-018-0096-z>
- May, J. G., Kennedy, R. S., Williams, M. C., Dunlap, W. P., & Brannan, J. R. (1990). Eye movement indices of mental workload. *Acta Psychologica*, 75(1), 75–89. [https://doi.org/10.1016/0001-6918\(90\)90067](https://doi.org/10.1016/0001-6918(90)90067)
- McComas, J., Pivik, J., & Laflamme, M. (1998). Children's transfer of spatial learning from virtual reality to real environments. *CyberPsychology and Behavior*, 1(2), 121–128.
- McGrath, R. L., & Kantak, S. S. (2016). Reduced asymmetry in motor skill learning in left-handed compared to right-handed individuals. *Human Movement Science*, 45, 130–141. <https://doi.org/10.1016/j.humov.2015.11.012>
- McKendrick, R. D., & Cherry, E. (2018). A deeper look at the NASA TLX and where it falls short. *Proceedings of the Human Factors and Ergonomics Society - Annual Meeting*, 62(1), 44–48. <https://doi.org/10.1177/1541931218621010>
- McKnight, P. E., & Najab, J. (2010). Kruskal-wallis test. *The Corsini Encyclopedia of Psychology*. <https://doi.org/10.1002/9780470479216.corpsy0491>
- Meshkati, N. (1988). Toward development of a cohesive model of workload. *Advances in Psychology*, 305–314. [https://doi.org/10.1016/s0166-4115\(08\)62394-8](https://doi.org/10.1016/s0166-4115(08)62394-8)
- Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. *Psychological Review*, 104(1), 3.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97. <https://doi.org/10.1037/h0043158>
- Monk, A. F., Jackson, D., Nielsen, D., Jefferies, E., & Olivier, P. (2011). N-Backer: An auditory N-back task with automatic scoring of spoken responses. *Behavior Research Methods*, 43(3), 888–896. <https://doi.org/10.3758/s13428-011-0074-z>
- Nachreiner, F. (1995). Standards for ergonomics principles relating to the design of work systems and to mental workload. *Applied Ergonomics*, 26(4), 259–263. [https://doi.org/10.1016/0003-6870\(95\)00029-c](https://doi.org/10.1016/0003-6870(95)00029-c)
- Nataraj, R., Sanford, S., Liu, M., & Harel, N. Y. (2022). Hand dominance in the performance and perceptions of virtual reach control. *Acta Psychologica*, 223, Article 103494. <https://doi.org/10.1016/j.actpsy.2022.103494>
- Naumann, J. (2019). The skilled, the knowledgeable, and the motivated: Investigating the strategic allocation of time on task in a computer-based assessment. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01429>
- Nawaz, S., Srivastava, N., Yu, J. H., Baker, R. S., Kennedy, G., & Bailey, J. (2020). Analysis of task difficulty sequences in a simulation-based Poe Environment. *Lecture Notes in Computer Science*, 423–436. https://doi.org/10.1007/978-3-030-52237-7_34
- Neale, D. C. (1997). Factors influencing spatial awareness and orientation in desktop virtual environments. *Proceedings of the Human Factors and Ergonomics Society - Annual Meeting*, 41(2), 1278–1282. <https://doi.org/10.1177/1071181397041002123>
- Negut, A., Matu, S.-A., Sava, F. A., & David, D. (2016). Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: A meta-analytic approach. *Computers in Human Behavior*, 54, 414–424. <https://doi.org/10.1016/j.chb.2015.08.029>
- Noltemeyer, A., Palmer, K., James, A. G., & Petrasek, M. (2019). Disciplinary and achievement outcomes associated with school-wide positive behavioral interventions and supports implementation level. *School Psychology Review*, 48(1), 81–87. <https://doi.org/10.17105/spr-2017-0131.v48-1>
- Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, 7(1), 44–64. [https://doi.org/10.1016/0010-0285\(75\)90004-3](https://doi.org/10.1016/0010-0285(75)90004-3)
- North, S., North, M., Garofalo, D., & Parajapati, D. (2021). The effects of mixed reality immersion on users performance and perception of multitasking while performing concurrent real world tasks. *The Journal of Computing Sciences in Colleges*, 36(8), 73–88.
- O'Donnell, E., Lawless, S., Sharp, M., & Wade, V. P. (2015). A review of personalised e-learning. *International Journal of Distance Education Technologies*, 13(1), 22–47. <https://doi.org/10.4018/ijdet.2015010102>
- Ogunseiju, O. O., Akanmu, A. A., & Bairaktarova, D. (2021). Mixed reality based environment for Learning Sensing Technology applications in construction. *Journal of Information Technology in Construction*, 26, 863–885. <https://doi.org/10.36680/j.itcon.2021.046>
- Orvis, K., Horn, D., & Belanich, J. (2007). Task difficulty and prior videogame experience: Their role in performance and motivation in instructional videogames. *Computers in Human Behavior*. <https://doi.org/10.1037/e500442012-001>
- Paas, F. G., & Van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 35(4), 737–743. <https://doi.org/10.1177/001872089303500412>
- Pagnotta, M., Jacobs, D. M., de Frutos, P. L., Rodríguez, R., Ibáñez-Gijón, J., & Travieso, D. (2022). Task difficulty and physiological measures of mental workload in Air traffic control: A scoping review. *Ergonomics*, 65(8), 1095–1118. <https://doi.org/10.1080/00140139.2021.2016998>
- Pan, Z., Cheok, A. D., Yang, H., Zhu, J., & Shi, J. (2006). Virtual reality and mixed reality for virtual learning environments. *Computers & Graphics*, 30(1), 20–28. <https://doi.org/10.1016/j.cag.2005.10.004>
- Park, K. B., Choi, S. H., Lee, J. Y., Ghasemi, Y., Mohammed, M., & Jeong, H. (2021). Hands-free human-robot interaction using multimodal gestures and deep learning in

- wearable mixed reality. *IEEE Access*, 9, 55448–55464. <https://doi.org/10.1109/ACCESS.2021.3071364>
- Pavlov, A., Duhon, G., & Dawes, J. (2021). Examining the impact of task difficulty on student engagement and learning rates. *Journal of Behavioral Education*, 32(3), 527–542. <https://doi.org/10.1007/s10864-021-09465-y>
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). The adaptive decision maker. <https://doi.org/10.1017/cbo9781139173933>
- Pedowitz, R., Nicandri, G., & Tuchscheid, S. (2016). Asymmetry in dominant/non-dominant hand performance differentiates novices from experts on an arthroscopy virtual reality serious game. *Medicine Meets Virtual Reality 22: NextMed/MMVR22*, 220, 289.
- Perlow, L. A. (1999). The time famine: Toward a sociology of work time. *Administrative Science Quarterly*, 44(1), 57–81. <https://doi.org/10.2307/2667031>
- Peters, M., & Ivanoff, J. (1999). Performance asymmetries in computer mouse control of right-handers, and left-handers with left- and right-handed mouse experience. *Journal of Motor Behavior*, 31(1), 86–94. <https://doi.org/10.1080/00222899909601894>
- Philiastides, M. G. (2006). Neural representation of task difficulty and decision making during perceptual categorization: A timing diagram. *Journal of Neuroscience*, 26(35), 8965–8975. <https://doi.org/10.1523/jneurosci.1655-06.2006>
- Poeschl, S. (2017). Virtual reality training for public speaking—a quest-VR framework validation. *Frontiers in ICT*, 4, 1–13. <https://doi.org/10.3389/fict.2017.00013>
- Prichard, E., Propper, R. E., & Christman, S. D. (2013). Degree of handedness, but not direction, is a systematic predictor of cognitive performance. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00009>
- Provins, K. A., & Magliaro, J. (1989). Skill, strength, handedness, and fatigue. *Journal of Motor Behavior*, 21(2), 113–121. <https://doi.org/10.1080/00222895.1989.10735469>
- Radeborg, K., Briem, V., & Hedman, L. R. (1999). The effect of concurrent task difficulty on working memory during simulated driving. *Ergonomics*, 42(5), 767–777.
- Rao, A. K., Chandra, S., & Dutt, V. (2020). Desktop and virtual-reality training under varying degrees of task difficulty in a complex search-and-shoot scenario. In *HCI international 2020 – late breaking papers: Virtual and augmented reality* (pp. 421–439). https://doi.org/10.1007/978-3-030-59990-4_31
- Rizzo, A. A., Bowerly, T., Buckwalter, J. G., Klimchuk, D., Mitura, R., & Parsons, T. D. (2009). A virtual reality scenario for all seasons: the virtual classroom. *CNS Spectrums*, 11(1), 35–44. <https://doi.org/10.1017/s1092852900024196>
- Rizzo, A. A., Buckwalter, J. G., Bowerly, T., Van Der Zaag, C., Humphrey, L., Neumann, U., Chua, C., Kyriakakis, C., Van Rooyen, A., & Sisemore, D. (2000). The virtual classroom: A virtual reality environment for the assessment and rehabilitation of attention deficits. *CyberPsychology and Behavior*, 3(3), 483–499. <https://doi.org/10.1089/10949310050078940>
- Rizzo, A., Buckwalter, J. G., Larson, P., Van Rooyen, A., Kratz, K., Neumann, U., ... Thiebaut, M. (1998). Preliminary findings on a virtual environment targeting human mental rotation/spatial abilities. In *Proceedings of the 2nd european conference on disability, virtual reality and associated techniques* (pp. 213–220). Sweden: Skövde.
- Rogers, S. (2018). What is mixed reality and what does it mean for enterprise? *Forbes*. Retrieved September 8, 2022, from <https://www.forbes.com/sites/solrogers/2018/12/04/what-is-mixed-reality-and-what-does-it-mean-for-enterprise/>
- Rokhsaritalemi, S., Sadeghi-Niaraki, A., & Choi, S.-M. (2020). A review on Mixed reality: Current trends, challenges and prospects. *Applied Sciences*, 10(2), 1–26. <https://doi.org/10.3390/app10020636>
- Salehi, S. K., & Orak, L. L. (2020). The effect of knowledge of performance feedback and task difficulty on performance and learning of basketball chest pass: Challenge based learning approach. *European Journal of Physical Education and Sport Science*, 6(20), 63–76. <https://doi.org/10.5281/zenodo.3700095>
- Salvucci, D. D., Taatgen, N. A., & Borst, J. P. (2009). Toward a unified theory of the multitasking continuum. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/1518701.1518981>
- Sanjram, P. K., & Gupta, M. (2013). Task difficulty and time constraint in programmer multitasking. *International Journal of Green Computing*, 4(1), 35–57. <https://doi.org/10.4018/jgc.2013010103>
- Santos, E., & Castillo, V. (2024). *Apple vision Pro: Comments in healthcare*. *arXiv preprint arXiv:2401.08685*.
- Schieman, S., & Young, M. (2015). Who engages in work–family multitasking? A study of Canadian and American workers. *Social Indicators Research*, 120(3), 741–767. <https://doi.org/10.1007/s11205-014-0609-7>
- Schweiger, D., Stone, R., & Genschel, U. (2021). Nondominant Hand Computer Mouse training and the bilateral transfer effect to the dominant hand. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-83770-4>
- Sheridan, T. B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, 1(1), 120–126. <https://doi.org/10.1162/pres.1992.1.1.120>
- Shin, M., Linke, A., & Kemps, E. (2020). Moderate amounts of media multitasking are associated with optimal task performance and minimal mind wandering. *Computers in Human Behavior*, 111, 1–6. <https://doi.org/10.1016/j.chb.2020.106422>
- Shuggi, I. M., Oh, H., Shewokis, P. A., & Gentili, R. J. (2017). Mental workload and motor performance dynamics during practice of reaching movements under various levels of task difficulty. *Neuroscience*, 360, 166–179. <https://doi.org/10.1016/j.neuroscience.2017.07.048>
- Speicher, M., Hall, B. D., & Nebeling, M. (2019). What is mixed reality?. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–15). <https://doi.org/10.1145/3290605.3300767>
- Spink, A., Cole, C., & Waller, M. (2008). Multitasking behavior. *Annual Review of Information Science & Technology*, 42(1), 93–118. <https://doi.org/10.1002/aris.2008.1440420110>
- Stanton, D., Foreman, N., & Wilson, P. N. (1998). Uses of virtual reality in clinical training: Developing the spatial skills of children with mobility impairments. In *Virtual environments in clinical psychology and neuroscience* (pp. 219–232).
- Steingrueber, H. J. (1975). Handedness as a function of test complexity. *Perceptual and Motor Skills*, 40(1), 263–266. <https://doi.org/10.2466/pms.1975.40.1.263>
- Susskind, R. E., & Susskind, D. (2015). *The future of the professions: How technology will transform the work of human experts*. USA: Oxford University Press.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Technavio. Covid-19 impacts: Augmented reality (AR) and virtual reality (VR) market will accelerate at a CAGR of over 35% through 2020-2024 : The increasing demand for AR and VR technology to boost growth: Technavio. Business Wire. <https://www.businesswire.com/news/home/20200903005356/en/COVID-19-Impacts-Augmented-Reality-AR-Virtual-Reality>
- Theill, N., Martin, M., Schumacher, V., Bridenbaugh, S. A., & Kressig, R. W. (2011). Simultaneously measuring gait and cognitive performance in cognitively healthy and cognitively impaired older adults: The basel motor-cognition dual-task paradigm. *Journal of the American Geriatrics Society*, 59(6), 1012–1018. <https://doi.org/10.1111/j.1532-5415.2011.03429.x>
- Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences*, 1(21), 19–25.
- Topi, H., Valacich, J. S., & Hoffer, J. A. (2005). The effects of task complexity and time availability limitations on human performance in database query tasks. *International Journal of Human-Computer Studies*, 62(3), 349–379. <https://doi.org/10.1016/j.ijhcs.2004.10.003>
- Tran, N. (2020). *Exploring mixed reality robot communication under different types of mental workload*. Colorado School of Mines.
- Tugtekin, U., & Odabasi, H. F. (2023). Effect of multitasking and task characteristics interaction on cognitive load and learning outcomes in virtual reality learning environments. *Education and Information Technologies*, 28(11), 14915–14942. <https://doi.org/10.1007/s10639-023-11813-6>
- Vallejo, V., Wyss, P., Chesham, A., Mitache, A. V., Müri, R. M., Mosimann, U. P., & Nef, T. (2017). Evaluation of a new serious game based multitasking assessment tool for cognition and activities of daily living: Comparison with a real cooking task. *Computers in Human Behavior*, 70, 500–506. <https://doi.org/10.1016/j.chb.2017.01.021>
- van der Kleij, R., Lijkwan, J. T. E., Rasker, P. C., & De Dreu, C. K. W. (2009). Effects of time pressure and communication environment on team processes and outcomes in dyadic planning. *International Journal of Human-Computer Studies*, 67(5), 411–423. <https://doi.org/10.1016/j.ijhcs.2008.11.005>
- Wang, L., He, X., & Chen, Y. (2016). Quantitative relationship model between workload and time pressure under different flight Operation Tasks. *International Journal of Industrial Ergonomics*, 54, 93–102. <https://doi.org/10.1016/j.ergon.2016.05.008>
- Wann, J. P., Rushton, S. K., Smyth, M., & Jones, D. (1997). Virtual environments for the rehabilitation of disorders of attention and movement. *Virtual Reality in Neuro-psychology-physiology*, 157–164.
- Wickens, C. D. (2008). Multiple Resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3), 449–455. <https://doi.org/10.1518/001872008x288394>
- Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, 57, 1336–1346.
- Willemsen, P., Jaros, W., McGregor, C., Downs, E., Berndt, M., & Passafaro, A. (2018). Memory task performance across augmented and virtual reality. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*. <https://doi.org/10.1109/vr.2018.8446457>
- Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, 37(1), 60–82. [https://doi.org/10.1016/0749-5978\(86\)90044-0](https://doi.org/10.1016/0749-5978(86)90044-0)
- Wright, L., Hardie, S., & Rodway, P. (2004). Pause before you respond: Handedness influences response style on the Tower of Hanoi Task. *Laterality: Asymmetries of Body, Brain and Cognition*, 9(2), 133–147. <https://doi.org/10.1080/13576500244000265>
- Wu, X., Lowyck, J., Sercu, L., & Elen, J. (2012). Task Complexity, student perceptions of vocabulary learning in EFL, and task performance. *British Journal of Educational Psychology*, 83(1), 160–181. <https://doi.org/10.1111/j.2044-8279.2011.02061.x>
- Xi, N., Chen, J., Gama, F., Riar, M., & Hamari, J. (2022). The challenges of entering the metaverse: An experiment on the effect of extended reality on workload. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-022-10244-x>
- Xiong, J., Hsiang, E.-L., He, Z., Zhan, T., & Wu, S.-T. (2021). Augmented reality and virtual reality displays: Emerging technologies and future perspectives. *Light: Science & Applications*, 10(1), 1–30. <https://doi.org/10.1038/s41377-021-00658-8>
- Xu, S., He, X., & Salvendy, G. (2008). An ergonomics study of computerized emergency operating procedures: Presentation style, task complexity, and training level. *Reliability Engineering & System Safety*, 93(10), 1500–1511. <https://doi.org/10.1016/j.res.2007.09.006>

- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18(5), 459–482. <https://doi.org/10.1002/cne.920180503>
- Ziefle, M., & Bay, S. (2005). How older adults meet complexity: Aging effects on the usability of different mobile phones. *Behaviour & Information Technology*, 24(5), 375–389. <https://doi.org/10.1080/0144929042000320009>
- Zomer, T. P., Erasmus, V., van Empelen, P., Looman, C., van Beeck, E. F., Tjon-A-Tsien, A., Richardus, J. H., & Voeten, H. A. C. M. (2013). Sociocognitive determinants of observed and self-reported compliance to hand hygiene guidelines in Child Day Care Centers. *American Journal of Infections Control*, 41(10), 862–867. <https://doi.org/10.1016/j.ajic.2012.11.023>