IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 11 March 2024, accepted 26 March 2024, date of publication 29 March 2024, date of current version 22 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3383142

TOPICAL REVIEW

Effect of Interruptions and Cognitive Demand on Mental Workload: A Critical Review

NITIN KOUNDAL[1], ABDUALRHMAN ABDALHADI[1], MAGED S. AL-QURAISHI[2], (Member, IEEE), IRRAIVAN ELAMVAZUTHI[3], (Senior Member, IEEE), MAHDIYEH SADAT MOOSAVI[4], CHRISTOPHE GUILLET[5], FRÉDÉRIC MERIENNE[4], (Member, IEEE), AND NAUFAL M. SAAD[1], (Member, IEEE)

[1]Centre for Intelligent Signal and Image Research (CISIR), Department of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS (UTP), Seri Iskandar 32610, Malaysia
[2]Interdisciplinary Research Center for Smart Mobility and Logistics, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
[3]Smart Assistive and Rehabilitative Technology (SMART), Department of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS (UTP), Seri Iskandar 32610, Malaysia
[4]Arts et Métiers Institute of Technology, LISPEN, HESAM Université, 71100 Chalon-sur-Saône, France
[5]University of Burgundy, LISPEN, 71100 Chalon-sur-Saône, France

Corresponding author: Naufal M. Saad (naufal_saad@utp.edu.my)

This work was supported in part by Universiti Teknologi PETRONAS (UTP), Bandar Seri Iskandar, Perak, Malaysia, for supporting this research through the Graduate Assistantship (GA) Scheme; in part by the Higher Institution Centre of Excellence (HICoE) Scheme awarded by the Center for Intelligent Signal and Imaging Research (CISIR) under Grant 015MAO-050; and in part by Yayasan Universiti Teknologi PETRONAS under Grant 015LC0-354.

ABSTRACT Worker safety and productivity are crucial for effective job management. Interruptions to an individual's work environment and their impact on mental health can have adverse effects. One prospective instrument for assessing and calculating an individual's mental state in an interrupted scenario and cognitive demand levels is the use of physiological computing devices in conjunction with behavioral and subjective measurements. This study sought to address how to gather and compute data on individuals' cognitive states in interrupted work settings through critical analysis. Thirty-three papers were considered after the literature search and selection procedure. This descriptive study is conducted from three perspectives: parameter measurement, research design, and data analysis. The variables evaluated were working memory, stress, emotional state, performance, and resumption lag. The subject recruitment, experimental task design, and measurement techniques were examined from the standpoint of the experimental design. Data analysis included computing and cognitive pre-processing. Four future research directions are suggested to address the shortcomings of the present studies. This study offers suggestions for researchers on experiment planning and using computing to analyze individuals' cognitive states during interrupted work scenarios. Additionally, it offers helpful recommendations for organizing and conducting future research.

INDEX TERMS Interruptions, cognitive task, mental workload, performance, emotion.

I. INTRODUCTION

An interruption is defined as any event that hinders productivity and is not directly related to the main task [1], provided that there is an intention to resume and complete the initial workstream [2]. According to Coraggio [3], an interruption refers to an external event that hinders an operator from carrying out their current task and diverts their attention to another task. Today's workplace is rife with interruptions. In contemporary work environments, there exists an abundance of interruptions that disrupt productivity. The advent of technological advancements, specifically in the realm of advanced information and communication technology, has enabled individuals to engage in multiple tasks concurrently [4]. Knowledge workers change their working spheres due to interruptions every 11.5 minutes, and they work in 10 different spheres daily, as noted by Gonzalez and Mark [5]. According to Leroy and Glomb, for today's workers,

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Boubchir.

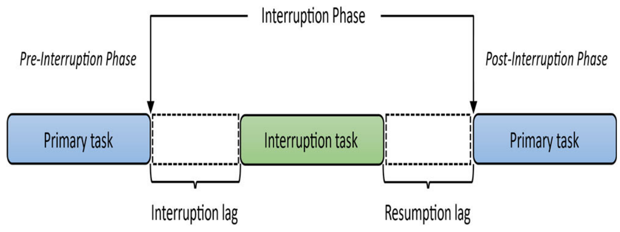


FIGURE 1. **Single interruption anatomy.**

"complete duties without interruptions have become a luxury." [6] Studies conducted by Wajcman and Rose [7] and Spira and Feintuch [8] revealed that knowledge workers encounter an average of 86 interruptions during their workday, resulting in a substantial annual cost of $588 billion to the US economy. Interruptions that are sudden and inevitable can cause frustration and stress, resulting in a negative impact on work performance, particularly in terms of accuracy and time to complete tasks [9], [10]. Interruptions worsen psychiatric symptoms, increase physical and mental pressure, and reduce productivity [11], [12], [13]. Interruptions reduce performance quality and negatively affect task performance. Moreover, they can give rise to feelings of unease, irritation, and a tendency to commit errors, hold-ups, and fluctuations [14], [15], arising mainly from limited cognitive capacities that are not appropriately allocated among various tasks [16], [17], [18], [19]. The majority of research has indicated that interruptions have a negative impact on the execution of the main task, primarily due to the presence of a delay in resuming the primary activity following the completion of a secondary task [20] (as depicted in Figure 1). However, alternative research argues that interruptions are advantageous for finishing tasks as they promote simple activities while having a detrimental impact on more intricate ones [21]. Recently, a substantial and increasing body of research has focused on the effects of interruptions on individual performance, error handling, and cognitive workload.

Mental workload (MWL) is an important subject in the field of work systems [22]. MWL refers to the cognitive energy required to complete a task within a limited timeframe [23]. When an individual finishes a particular assignment, the burden can either be cognitive or physical, although they are closely connected and cannot be entirely distinguished [24]. Operators exert increased effort to accomplish challenging and arduous tasks [25]. Subjects experience boredom and errors when their MWL decreases to an inadequate level [26]. An operator overload can occur due to a rise in the requirement for available resources [27]. The operator's daily MWL and lack of rest can lead to health problems such as chronic stress, burnout, and depression [28], which may affect their overall well-being and performance [29]. Moreover, improving the MWL of a system operator can enhance operator contentment, reduce the likelihood of errors and training costs, and enhance system efficiency and security [30], [31], [32], [33]. The most widely used methods for assessing MWL are evaluations of secondary task performance, measures of primary task performance, physiological techniques, and subjective measures [26], [34], [35]. The majority of experimenters created tools and techniques based on laboratory research, and then applied their findings to analyze performance reactions in real-world settings; thus, field research employing scientific evaluation techniques may yield more advantageous results than laboratory-based studies, particularly in the case of MWL, which significantly affects job performance [36].

One notable characteristic that is exhibited is the concept of restricted capacity, which refers to the inherent limitation in human beings' ability to manage a finite amount of knowledge at any given time. In the realm of cognitive psychology, Wickens [37], [38] made groundbreaking contributions to the development of the MWL concept, which elucidates the notion that the cognitive resources required for job accomplishment are inherently constrained. Advancing this line of inquiry, Sharples and Megaw have shed light on the intricate relationship between the availability of cognitive resources and task performance [39]. It is posited that when an individual possesses surplus cognitive resources, they may be capable of undertaking additional tasks (can be referred as cognitive demand levels) simultaneously; however, as the demands of these tasks start to surpass the available cognitive resources, the levels of MWL become excessively burdensome, leading to a precipitous decline in performance. Importantly, this model also recognizes that the performance of tasks may deteriorate due to underload, wherein individuals fail to allocate sufficient attention to the task at hand.

There is a vast and expanding body of research in the field since interruptions and their impact on job implementation, mistake management, and mental burden have gained attention in recent years. Two reviews have shown the effects of distraction and interruption, cognitive load, and workplace stress. Engström et al. [40] reviewed 84 articles and reports investigating cognitive load's effects on driving performance. This review focuses on various aspects of driving performance, including object and event detection, lateral control, longitudinal control, and decision-making. This review mainly focused on the effects of cognitive load on driving performance in controlled settings, such as simulators or test tracks, and may not fully capture the complexities of real-world driving conditions. This study did not include naturalistic driving studies, which could provide insights into the effects of cognitive load on driving performance in real-world scenarios. In the review paper, Jean-François Stich [41] adopted the "overview of reviews" method to integrate findings from distinct disciplines and themes, leveraging existing reviews to compare workplace stress in virtual and traditional offices. This study uses several key reviews to compare workplace stress in virtual and traditional offices by selecting review papers. The selected reviews covered various sources of stress, such as interruptions, workload, communication, and work-life conflicts. However, the consequences of interruptions and mental job levels on cognitive capacity in practical settings have not been thoroughly examined. More significantly, specific concerns in previous research continue to lack systematic answers.

1) How can an experiment be planned to produce high-quality data for a field study?
2) How can mental burdens be calculated in an interrupted real-world setting while processing data accurately and efficiently?
3) What research should be conducted in the future, and what are the existing research constraints for calculating the variation in mental capacity and cognitive states caused by interruptions and cognitive task levels?

This study aimed to provide a systematic asset for determining the effects of interruptions and mental task levels on mental strain in a real-world scenario by performing a comprehensive study and assessment of significant peer-reviewed academic journal articles. Interruptions and mental or cognitive workload, research strategy, and data management in the current research paper were analyzed and summarized. This study's deficiencies were recognized to create a guide for future investigation and immediate actionable engineering practice.

The remainder of the paper is organized as follows: Section II briefly describes interruptions, MWL, and cognitive task levels. The methodological structure of this study is described in Section III. Section IV summarizes these studies from the perspectives of interruption and MWL selection, study design, and data management. Section V encapsulates the current limitations of this study and offers suggestions for future research. The conclusions are presented in Section VI.

II. STUDY BACKGROUND

A. INTERRUPTIONS

An unexpected delay in behavioral performance, attentional focus, or both from an ongoing task is referred to as a job interruption [20]. There are five interruptions based on Jett and George's typologies [42] and those of Leroy, Schmidt, and Madjar [43]: *intrusions, distractions, multitasking, breaks*, and *surprises*. Mark, Gonzalez, and Harris divided interruption into internal and external categories, defining *internal interrupts* as "situations where a person purposefully pauses a task" and *external interrupts* as "those that arise from the events in the environment" [44]. Sasangohar, Donmez, Easty, and Trbovich refer to interruption as a *nested interruption* when the tertiary task disrupts the secondary task (which interrupts the primary work). For instance, a doctor asks an Intensive Care Unit nurse to order medicines via a computer system (a secondary task) while preparing the medications (a first or primary task). The nurse was disturbed by an urgent pump alert (a third or tertiary task) while completing the drug order [45]. The four types of interruption studies were (a) *objective*, (b) *subjective*, (c) *episodic*, and (d) *frequency* [20].

Objective approach: In this method, the researcher determines whether an event is interrupted, not the participant. Observers may directly or indirectly observe disruptions through video recordings of the task. This method is appropriate for examining work interruptions that are externally visible to the researcher because it does not capture the participants' subjective responses. The objective approach generally helps research how observed task pauses impact objectively quantifiable outcomes, including task performance and resumption, mistakes, and completion times [46].

Subjective approach: This method arrests the individuals' subjective appraisal of interruptions. This method is suitable for researching emotional, attitudinal, and stress reactions to interruptions because it focuses on participants' assessments of work interruptions. However, the act of recording the participants' subjective experiences might turn into an interruption. Changes in research design (such as administering surveys outside of working hours) can solve this problem but may incur additional expenses (e.g., retrospective bias) [46].

Episodic approach: This strategy assumes that the content of each interruption varies. Researchers can examine individual interruption episodes in laboratory settings or field investigations using participants' memories of specific interruptions. This method enables researchers to concentrate on the components of a particular task disruption and the characteristics of the jobs, the participants, and the circumstances of that disruption period influence the results. However, because of its event-level emphasis, the episodic method is less suited for examining the cumulative impact of frequent interruptions experienced in modern workplaces.

Frequency approach: Instead of focusing on the experience of a specific interruption, this strategy emphasizes the overall implications of several interruptions. The presumption is that interruptions have identical substances and that their aggregate effect determines their impact. This method aids in analyzing the overall impact of interruptions spread across days, weeks, or months. The characteristics of the interrupting or interrupted activities or the interrupter engaged in each work disruption are challenging to identify and examine because of the focus on aggregate impacts [20].

B. MENTAL WORKLOAD

The total extent of cognitive or memory work essential for executing a job is called MWL, often called the cognitive load [47]. The quantity of mental energy required to perform a job in a finite amount of time is known as the MWL [23]. Multiple methods were employed to estimate cognitive workload. Various studies have employed a *subjective approach*. Difficulty in executing a particular job determines workload. The National Aeronautics and Space Administration (NASA) task load index (TLX) technique [48] is a good example of this strategy. Integrating apparent ratings on the six subscales and weighting evaluates and quantifies the workload level. This subjective self-reporting measurement assumes respondents can identify tasks with varying workloads [49].

The *behavioral task performance method* considers poor behavioral performance to be a reliable indicator of workload and a quantifiable result for a specific cognitive activity [37], [50]. Quantifiable or qualitative outputs show how well an individual performed the work. Although MWL and task performance are related, they do not always match. When the MWL is optimal, task performance is at its maximum; however, it may decline when it is too low or too high. The crucial components of this approach include choosing appropriate tasks and considering individual variations [49].

The *physiological technique* incorporates statistics from the heart rate (HR) [51], HR variability (HRV) [52], galvanic skin response [53], skin temperature, brain measurements, and breathing rate to determine the upper and lower limits of cognitive capacity. However, this information is collected using intrusive and invasive techniques, such as wearing or directly attaching sensors, and there is a risk that personal health data will be exposed [49].

C. COGNITIVE DEMAND LEVELS

The Multiple Resource Model (MRM) [38] expands upon the concept of limited resources by taking into account various types of resources and different stages of processing. This theoretical framework encompasses three dimensions, which propose that encoded information can be either spatial (visual) or verbal (auditory) in nature (left dimension). Additionally, the perception and cognition of this information can be either spatial or verbal (middle dimension), and the selected responses can be either spatial (manual) or verbal (vocal) in nature (right dimension). The successful completion of multiple tasks (cognitive demand levels) simultaneously relies on the absence of competition or overlap in the required resources across these three dimensions. For instance, engaging in texting while driving can have disastrous consequences because both activities involve visual encoding, spatial processing, and require a manual response. On the other hand, driving while having a conversation is theoretically feasible since it necessitates different resources across the three dimensions. Therefore, based on the MRM, such a dual-task scenario is considered viable.

While managing numerous concurrent activities is a common task in the workplace, varying task complexities often necessitate different amounts of data processing resources, resulting in varying degrees of MWL [54]. Rasmussen [55] created a prominent taxonomy of various information-handling tasks to discern the different levels of individual mental behavior: *rule*-, *skill*-, and *knowledge-centered behavior*. Performing *skill-centered* tasks often entails robust, automated, and seamless signal-response processes [56]. Propositions stored in long-term memory are used to predict rule-based behaviors [55]. *Rule-based* behavioral cognitive function is slower and less automated than *skill-centered behavior*. *Knowledge-centered behavior* includes new conditions that lack existing solutions. Under these circumstances, dealing with the problem requires considerable mental effort and delayed responses. Knowledge tasks often require higher cognitive skills, including situational awareness in planning, decision-making, and problem-solving [57]. As a result, the mental load becomes overwhelming, and performance may decline quickly when the task needs to surpass the available cognitive resources [39]. Diverse behavioral demands can also affect the cognitive load, and varied cognitive behaviors may be related to various physiological responses [54].

D. FIELD VS. CONTROLLED STUDIES

Field research has been conducted in both natural and real environments. Instead of manipulating the factor under study, observing, studying, and clarifying what already exists is preferable [58]. The naturalness of the environment was maintained as the study settings mimicked real-world scenarios. The subjects in the field study may or may not have been aware they were under observation. *Controlled research* is an analysis performed in a situation created specifically for the study. *Laboratory research* is a closely regulated experiment in which the investigator controls the specific element under investigation to determine whether it manipulates changes in individuals [59]. In laboratory research, individuals may be picked more carefully, placed in more controlled environments, and are often aware that they are part of a study. The main benefit of *field research* is that it can be employed in numerous real-world scenarios because it represents a broader range of circumstances and locations than *laboratory studies* [60]. Be aware that this benefit may be deceptive; it may be challenging to determine the generalizability of the study because of the absence of control and the inability to describe the field setting precisely.

A code of ethics may occasionally influence the choice of where to conduct a study [61]. *Laboratory investigations* provide better control over irrelevant factors that may otherwise affect the results and provide clearer cues about the behavior studied [62]. Any change in the participants was due to the manipulated variables when all uncontrolled factors were successfully eliminated. This method successfully established a cause-and-effect connection. However, prudence is advised when considering such a connection. Like other study methodologies, there are drawbacks to laboratory research, possibly signifying an artificial setting that affects the participants' actions and outcomes [63].

III. RESEARCH METHODOLOGY

The research methodology of this study consists of three steps (see Figure 2). Two critical academic records were examined to identify relevant articles. Subsequently, a two-round assessment procedure eliminates unacceptable results and establishes the essential evaluation parameters. Finally, by reading the article's complete text, the study topics' characteristics, research designs, and data handling presented in the analyses were extracted and examined.

A. SEARCH STRATEGY

Two renowned academic research databases, *Scopus* and the *Web of Science Core Collection* (WSC), were thoroughly searched for papers published in this literature review using their titles, abstracts, and keywords. According to the

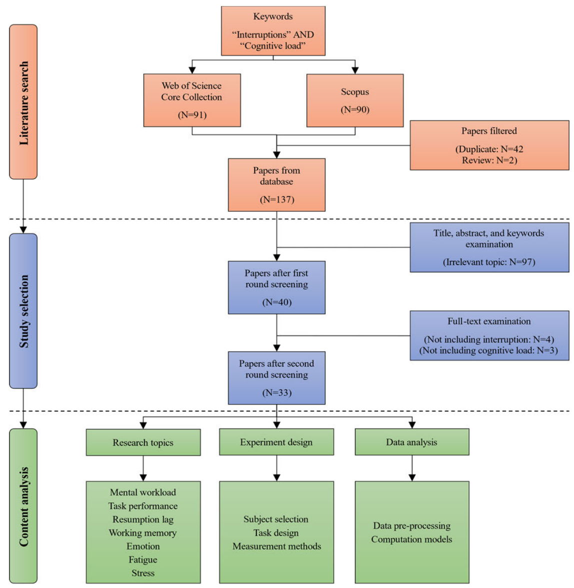


FIGURE 2. Research methodology.

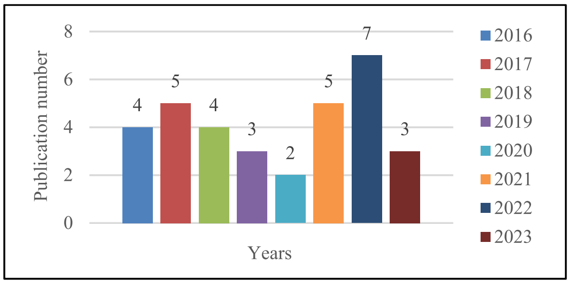


FIGURE 3. Number of publications selected for review annually on Interruption and MWL.

preliminary search results these two databases are among the most prominent academic resources available online regarding journals and current articles. The search scope considers workload levels, cognitive load, and interruptions because the present review focuses on how interruptions and cognitive task levels affect MWL. Therefore, the following keywords were used for the literature search: ("*workload*" OR "*cognitive load*" OR "*workload levels*") AND ("*interruptions*" OR "*work interruptions*" OR "*interruption*") in terms of topic for WSC and terms of article title, abstract, and keywords for Scopus. The study was refined to account for only English-language articles available in peer-reviewed journals between 2016 and 2023 in the engineering and neurosciences neurology research areas. A total of 181 publications were found in the initial search; 90 were from Scopus, and the remaining 91 were from the WSC. The 181 articles from the original inquiry were reduced to 137 after removing 44 articles (42 duplicates and 2 reviews).

B. INCLUSION AND EXCLUSION

This systematic review followed the PRISMA guidelines [64] and study was based on the effects of interruptions and cognitive task levels on the MWL, meaning that the experiment must include tasks for measuring mental or cognitive load, and the included studies must consist of interruptions. Studies that were not entirely related were excluded. Two separate researchers assessed publications, and the acceptability of the study was decided by their agreement. A two-round screening strategy was used to narrow the pool of included studies. Titles, keywords, and abstracts were searched in the first round to remove irrelevant papers. The feasibility of the remaining documents was determined in the second round by performing a full-text analysis.

In the initial review (based on title, abstract, and keywords), 97 irrelevant publications were eliminated from the 137 selected articles (does not include mental or cognitive load, and interruptions). After doing full-text analysis, 3 publications that included interruptions study but did not involve mental or cognitive workload and 4 papers that did not concentrate on interruptions (including mental or cognitive workload) were eliminated during the second round of full-text review. Finaly, 33 papers were selected for the systematic review after screening.

C. CONTENT ANALYSIS

The essential elements of the identified studies, such as publication information (year, journal), research topics (investigated interruptions and mental or cognitive workload), research design (subject selection, task design, and measurement process), and statistical handling (data preprocessing and computing techniques), were obtained and examined by analyzing the entire manuscript of the articles.

IV. RESULTS

A. DESCRIPTIVE ANALYSIS

Zeigarnik attempted to describe selective memory processes while executing duties, presenting one of the earliest documented interruption experiments [65]. The study of interruptions and their impacts gained prominence during the second half of the 20th century due to catastrophes in safety-critical fields (e.g., Edwards and Gronlund 1998 [66]; National Transportation Safety Board 1988 [67]). The study of interruptions has attracted interest from numerous scientific fields over the past few decades. According to the literature, mental psychology and human–computer interfaces are key fields for studying interruptions. The number of publications selected for review each year on how interruptions and the intensity of mental tasks affect cognitive state or workload is depicted in Figure 3. The 33 papers were selected to span the following application areas: medical (seven papers), office (six papers), aviation (four papers), naval (two papers), industry (one paper), and other (13 papers). A summary of information on interruptions and their impact on cognitive state taken from the 33 papers that were evaluated is discussed below. Apart from the widespread use of technology at work, today's companies emphasize cooperation and open workspaces, making interruptions more common [68]. It is hardly surprising that there have been many studies on interruptions over the last two decades. However, as researchers concentrate on ideas

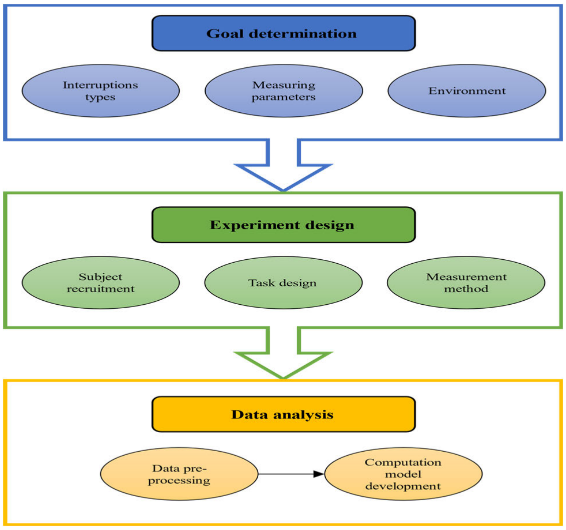


FIGURE 4. Research framework.

and results most consistent with their specialties, this study is dispersed across fields with minimal integration. The number of relevant studies will increase from 2016 to 2023 due to the rapid advancement of noninvasive biosignal (physiological) measuring methods and algorithms. Seven articles published in 2022 specifically demonstrated the existence of a sizable and expanding body of research on interruptions and their impact on the cognitive state in recent years.

The 33 highlighted studies can also be used to develop a generic research framework to investigate how interruptions and task levels affect cognitive states in practical environments (see Figure 4). Therefore, researchers must select their study topics before planning and conducting experiments. The authors must select sufficiently representative subjects, choose and use appropriate equipment, execute the experiment, and capture raw data. Before data processing, outliers were eliminated after acquiring raw data. Subsequently, appropriate computing models with abundant data are provided for additional analysis. The following section provides more details on each stage.

B. RESEARCH TOPICS

The three primary areas of interest in the publications under examination were interruption types, measurement parameters, and environment (see Table 1).

1) INTERRUPTION TYPES

Several disruptions have been examined in the publications under review. Therefore, they can be classified as internal or external. *Internal (intrinsic) interruptions* are caused by thought, discomfort, weariness, and other factors form within the individual [69]. Only two articles experienced internal interruptions during the review process. Gontar et al. investigated the effect of internal interruptions on cabin crew personnel performing turnaround duties. Internal interruptions occur less frequently and are more challenging to detect than external interruptions [70]. The stress effects of four different forms of observed interruptions were examined by Fletcher et al. in 2018. Rumination and breaks are internal disruptions in this study [46]. Other sources (stimuli) outside the individual such as phone calls, emails, environment, and other distractions lead to *external interruptions* [69]. Mental interruptions (such as reading, math problems, and sentence and noun repetition) [54], [71], [72], [73], [74], [75] social interruptions (such as phone calls and discussions) [54], [71], [76], [77], [78], [79], pop-ups and messages (such as chat-answering and suggestion messages) [49], [80], [81], [82], [83], [84], [85], [86], and auditory and verbal disruption (such as "Excuse me, could you please help me," giving feedback and alarm) [46], [72], [87], [88], [89], [90], [91], [92], [93] were the most frequent types of external interruptions that were observed. Falkland et al. identified an interruption in their study by verifying a prescription for a 68-year-old female of an abnormally low dosage of intravenous Panadol (180 mg) [94]. Falkland et al. regarded participants completing the NASA-TLX report as an interruption [95].

In the mental rotation task, part of the interruption task in Lodinger et al., participants indicated whether the alphanumeric stimuli were positioned in a regular or mirror-image direction [96]. Gontar et al. investigated cabin interruptions and provided a list of the people [70]. The four interconnected interruptions identified by Andreasson et al. are process-driven, social, nested, and notification interruptions. The primary responsibility of the maintenance staff is to execute planned and scheduled maintenance activities for the equipment, which constitutes their main task. However, they frequently face interruptions and are assigned to address more critical issues, known as interruption tasks. Additionally, they are sometimes called upon to handle urgent warnings, which are nested interruptions, thereby preventing them from completing their report and the original maintenance job described in an emergency work order report [97]. Spatial and non-spatial visual interruptions are the two visual interruptions introduced by Borowsky et al.. The spatial task presented a black screen containing ten white asterisks. The non-spatial task presented a black screen slightly tinted either red or blue [98]. The study done by Koundal et al., the impact of cognitive task levels and nested interruptions on mental states was assessed [99].

2) MEASURING PARAMETERS

In the reviewed articles, researchers examined mental or cognitive workloads, and with that, they also looked at additional factors such as task performance, resumption lag, working memory, emotion, fatigue, and stress. Performance and mental or cognitive workloads have often been investigated across all parameters. Cognitive states are interdependent [101]. The total amount of cognitive effort or memory humans require to execute a task is known as *mental* or *cognitive workload* [47]. The subjects experienced an increased workload when the condition was interrupted instead of continuous [54]. The number of interruptions increases the workload [76].

TABLE 1. Research topics in the publications under examination.

Articles	Research Topics		
	Interruptions	Measuring Parameters	Environment
[54]	❖Fun video (20 s) ❖Message from a friend ❖Short passage	Cognitive load, Performance	Laboratory
[71]	❖Mathematical questions ❖Communication	Cognitive load, Performance, Fatigue, Resumption lag	
[76]	❖Air traffic controller ❖Cabin crew ❖Electronic centralized aircraft monitoring warning	Cognitive load, Performance	
[94]	❖Check a medication order for a 68-year-old female	Cognitive load, Performance	
[80]	❖Pop-up picture	Cognitive load, Emotion	
[72]	❖Auditory alerts	Cognitive load	
[87]	❖Auditory interruptive arithmetic task	Cognitive load, Performance	
[73]	❖Five math tasks	Working Memory, Fatigue, Cognitive load	
[88]	❖Generic conversation (Verbal)	Cognitive load, Performance	
[95]	❖Filing of the NASA TLX	Working Memory, Cognitive load	
[96]	❖Different rotation angles ❖Durations of the mental rotation	Resumption lag, Cognitive load	
[78]	❖Three pop-up message questions	Cognitive load	
[74]	❖Filler task	Cognitive load	
[49]	❖Pop-up response	Cognitive load, Performance, Resumption lag	
[90]	❖Computer-based unique question response	Cognitive load, Performance	
[84]	❖Chat-answering	Cognitive load, Performance	
[92]	❖Audio ❖Visual	Performance	
[98]	❖Spatial ❖non-spatial	Performance	
[100]	❖Math task	Performance, Working memory, Resumption lag	
[86]	❖Pop-up ❖Warning message	Cognitive load, Performance	
[77]	❖Answering a phone call ❖Answering questions from colleagues ❖Answering question from Residents ❖Updating a patient's condition to the doctor ❖Answering questions from the patients' family members	Cognitive load	Real-world
[46]	❖Intrusion ❖Break ❖Distraction ❖Rumination	Stress, Cognitive load	
[89]	❖Verbally answering a question ❖external stimulus	Working Memory, Cognitive load, Performance	
[79]	❖Communication from doctor ❖Communication from nurse ❖Social communication ❖Phone call ❖Other prompt	Cognitive load	
[70]	❖Passenger list ❖Calculating performance	Cognitive load	
[97]	❖Process-driven ❖Social ❖Nested ❖Notification	Cognitive load	
[91]	❖Alarm	Cognitive load, Performance	
[75]	❖Perform another Lego assembly ❖Pencil-and-paper math problems	Resumption lag, Cognitive load	Controlled real-world
[81]	❖Pop-up messages	Resumption lag, Cognitive load	
[82]	❖New message reminders pop-up	Cognitive load, Performance, Emotion	
[83]	❖Dialog message	Cognitive load, Performance	
[93]	❖Feedback ❖Guidance to trainees	Cognitive load, Performance, Frustration	
[85]	❖Notification ❖Message	Cognitive load	

When interruptions occurred at precise breakpoints, participants reported a higher mental effort [81]. Participants who experienced interruptions during the problem identification phase reported having a heavier mental effort [82]. In the literature the camera views used in laparoscopic surgery provide different views of the anatomy and have different cognitive costs and associated levels of workload. Longer resumption delays were an outcome of the side view of peg transfer task, which was also considered mentally demanding [96]. In the interruption situation, the mental effort of nurses was 2.04 times elevated, while patient care tasks were 4.72 times higher for electronic medical record (EMR) charting [77]. Even small changes (such as displaying alarms on an integrated workstation) can reduce the workload in a complex work environment such as an operating room, improving patient safety [91]. Interruptions during multi-robot supervision tasks increase the perceived workload, with extrinsic interruptions having a more negative effect on the workload than intrinsic interruptions [85]. Contrary to question cues, visual cues can efficiently decrease the MWL of workers over the course of learning by encouraging them to focus upon the regions that contain safety hazards and the mental effort involved in accurately identifying them. As a result, workers can find safety hazards easier [102]. *Task performance* refers to the efficiency of a participant in completing a task [103]. Students in interrupted situations execute less effectively and more slowly than in uninterrupted settings [54]. During the interruption, compared to emergency physicians who used fewer cues, those who used more cues scored much better in the simulation test [94]. After interruptions, the participants' attention was redirected in a different direction due to the heavy workload, which further impaired their performance on the core task [87]. Only those individuals who performed worse on the test were interrupted during the evaluation and selection phases [82]. In the instant and scheduled modes, the job efficiency of the skill primary task settings was lower than that of the cognitive primary task, settings. However, in a planned manner, there was no distinction between the skill/mental work set and the skill/skill task set. In the immediate method, the time performance ratio of the main cognitive task sets was noticeably higher than that of the primary skill task sets [49]. There was a slightly lower deterioration in the performance of the primary task due to the interruption of alarm handling with the integrated workstation [91]. Attending to notifications during a complex sensorimotor task negatively impacts primary task performance, regardless of the modality used to present the notification [92]. Task interruption disrupted post-interruption performance and accuracy, with larger P3 amplitudes and alpha power after interruption than after suspension [100]. A two-stage warning system enhanced situation awareness reduced MWL, and improved takeover performance compared with a single-stage warning system [86].

The Resumption lag is the time between the cessation of the second job and restarting the primary job [104]. Resumption lags among primary and secondary jobs take longer in the negotiated mode than in the other modes [49]. Participants took much longer to revert to their primary jobs, while interruptions occurred at precise breakpoints [81]. When interrupted for a prolonged period and utilizing a side view instead of a top view, the subjects needed considerably more time to restart the peg transfer job. However, it took less time to restart the peg transfer activity across trials from both viewpoints [96]. More time is needed to resume the primary visual-manual assembly performance after interrupting a comparable activity [75]. Increased P3 amplitudes and alpha power in the interruption tasks suggest that the interference of irrelevant information has a stronger effect on resumption lag [100]. The Memory for Goals theory is a cognitive model that seeks to explain how goal-directed activity and memory recall are affected by interruptions affect people's ability to execute goal-directed activity and recall information. It postulates that performance declines due to a reduction in memory engagement of the main task relying upon the *working memory* process [104]. Chen et al.'s findings show how suppressing unimportant information influences memory performance following an interruption [73].

Participants who demonstrated a more vital ability to use cues also demonstrated less performance loss after interruptions [95]. Interruptions increased the WML, as evidenced by an increase in theta power [100]. *Emotions* are cognitive states triggered by neurophysiological shifts; they are linked differently to ideas, sensations, behavioral reactions, and levels of pleasure or discomfort [105]. Researchers have attempted to evaluate emotional status from a multidimensional viewpoint due to the complexity of emotional states. The valence-arousal-dominance paradigm identifies three aspects of an individual's emotional state: a dominance dimension from being in charge to being ruled by emotions, an arousal component from not being stimulated to excitement, and a valence dimension from dislike to pleasure [106]. Compared to positive material, negative content elicited less cognitive burden, visual attention, and annoyance [80]. In the problem identification phase, interrupted participants reported having a more unfavorable attitude toward interruptions [82]. *Fatigue* is a reduction in the capability and effectiveness of mental, physical, or both tasks caused by an excess of either task [107]. Mental fatigue after an interruption significantly affected the performance of key tasks, workload, and resumption lag [71]. *Mental stress* is an unreasonable and adverse psychological reaction to variations in obligations that affect the human nervous system [108]. Distractions extend beyond the common variance described by intrusions, breaks, and disparities to explain a distinct fraction of the variation in occupational stress [46]. Abdalhadi et al. study the effect of acute stress on decision making using Function Near Infrared Spectroscopy [109].

3) ENVIRONMENT

The evaluated studies included a range of environmental settings, including laboratory settings and real-world scenarios (in the wild) (see Figure 5). In *laboratory research*,

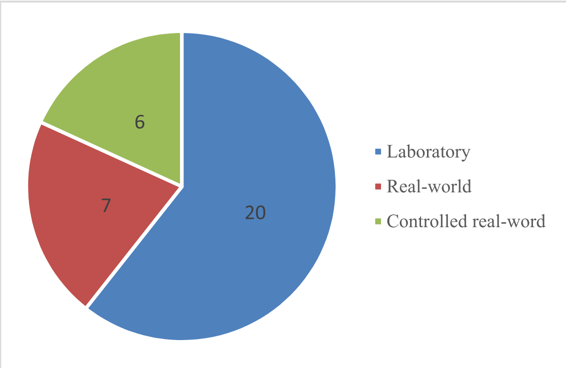


FIGURE 5. Environmental setting of evaluated studies.

factors under investigation are carefully controlled to determine whether they affect individual characteristics [110]. The experimenter attempted to simulate a natural environment using various simulation techniques, including the A320 flight simulator [76] aerodrome control simulator [93], automated car simulator [72], [92], emergency dispatch simulator [87], rail control simulator [95], S-CCS microworld (a simulation of single-ship naval anti-air warfare) [78], and dynamic position simulator [74]. The laboratory computer experiments stood out the most during the screening process. Research conducted in an actual or realistic world is known as a "*Real-world study*." Rather than changing the elements under investigation, it seeks to observe, examine, and elucidate what currently exists [58]. The experimenter conducted the study in various settings, including the foundry industry [97], airports [70], hospital emergency rooms [77], [79], [89], [91], and engineering workplaces [46].

C. EXPERIMENT DESIGN

High-quality raw data are required to evaluate the participants' cognitive states practically. However, measuring various metrics is frequently problematic because the findings are susceptible to participant errors, irrelevant variables, and environmental interventions. The participants' predictive validity strongly influenced the generalizability of the generated models and results. Even if these errors and interferences are unavoidable, the experimental design can be improved to reduce their impact. In other words, a logical experimental strategy is necessary to enhance the accuracy of the raw statistics gathered and used to evaluate participants' cognitive states. A good experiment must consider various factors, including subject assortment, task layout, and measurement techniques (Table 2).

1) PARTICIPANT SELECTION

Selecting a prominent and representative sample of subjects is essential to gathering high-quality data [111]. The effects of individual variations on the findings were eliminated, and the dependability of the conclusions improved when the participant size was sufficiently large. Table 2 shows the number of subjects analyzed to determine how interruptions affecting participants are depicted. According to Brysbaert [112], an effect size of $d = 0.4$ is a reasonable initial estimation of the minimum effect size that matters in psychological tests. Therefore, no fewer than 50 participants were required to compare the two within-participant variables with 80% power. Moreover, 100, 200, or more individuals were required when a between-group variable or interface was added. Since they were collected under perfect scenarios, these data constituted the absolute minimum for an 80%-powered study within the restriction set. The cognitive state of a subject is more nuanced in the real world. Related data often defy specific statistical test criteria and require clarification (e.g., normal distribution, balanced designs, no extraneous fluctuating sources of noise, and independence of observations). In addition to the number of subjects, the representation of the chosen participants is essential for the accuracy of the measured data. Participants must be carefully chosen to portray the audience's vital demographic traits and other features. A total of 33 studies were assessed, 16 of which selected employees to participate in the experiments [54], [70], [73], [76], [79], [80], [83], [84], [88], [89], [90], [91], [93], [94], [97], [98]. Alternative studies used undergraduate students.

2) EXPERIMENT TASK DESIGN

The critical component of the experimental design is the experimental task. The required tasks must be as practical as feasible to reflect worker actions in the natural realm. Six studies employed research activities from the real world, six utilized simulation tasks, and fourteen used computer tasks in the publications under evaluation. Kim et al. conducted a study in the Rochester, Minnesota, and Mayo Clinic emergency rooms in 2019 [77]. A self-report assessment of four categories of perceived interruptions was developed and validated in two populations (working undergraduate students and engineers) as part of Fletcher et al.'s research [46]. Walter et al. and Westbrook et al. studied emergency doctors employed in a teaching hospital in Sydney [79], [89]. In the study by Gontar et al., all turnaround jobs related to short-haul flights were scheduled on an Airbus of the A320 series [70]. The heavy-duty diesel engine foundry for automobiles, buses, and trucks served as the site for Andreasson et al., where regular worker duties were monitored [97]. Some studies used typical operations, including driving [72], [92], [98], operating rail equipment [95], flying an A320 [76], and performing navy missions [74], [78] and duplicated them in a laboratory setting (a simulator) as experiments. Only eight of the 33 studies reviewed the assigned cognitive levels. Doost et al. examined the degree of cognitive tasks based on skills, rules, and knowledge [54]. Chen et al. examined system monitoring, resource management, and task-level tracking [71]. Midha et al. investigated three levels of reading and writing assignments [88]. The research under consideration also used sentence copying as a physical task and mathematical question solving as a cognitive activity [49], [100]. Campoe et al. defined *task levels* as patient-controlled analgesia (PCA) with the bolus, basal PCA with the bolus, and continuous PCA with the bolus [90]. Verb creation and noun repetition were

TABLE 2. Experiment designs in the publications under examination.

Articles	Experiment Design: Number /Nature of Subjects	Tasks	Measuring method	Measuring technique
[54]	30/Workers	❖Skill-based (visual tracking) ❖Rule-based (size classification) ❖Knowledge-based (problem-solving)	Subjective and Objective	❖EDA ❖ECG ❖Primary task performance index ❖NASA-TLX
[76]	20/Workers	❖Non-pilot (Dual N-Back Test ❖Pilot Flying (A320 flight simulator)		❖ EEG ❖NASA-TLX
[81]	54/Adult students	Weekend trip to Chicago: ❖Select a rental car ❖Create an activity itinerary ❖Select a restaurant		❖GoPro HERO7 Black ❖NASA-TLX
[87]	41/Adult students	Emergency Dispatch Simulator: Monitor moving emergency aid vehicles on a map		❖Eye-tracking test ❖NASA-TLX ❖Performance scale
[82]	60/Adult students	Online procurement of domestic medical equipment		❖Performance scale ❖NASA-TLX ❖EEG
[73]	34/Adults	Spatial 2-back task		❖SSS ❖NASA-TLX ❖EEG
[88]	20/Workers	Reading (three levels) and Writing (three levels)		❖fNIRS (Octamon) ❖GoPro ❖NASA-TLX
[96]	32/Adult students	Peg transfer (Laparoscopy simulator)		❖HRC-470 Color CCD Bullet Cameras ❖NASA-TLX
[78]	36/Workers	S-CCS microworld (Simulation of single-ship naval antiair warfare)		❖Eye-tracking test ❖NASA-TLX
[49]	40/Adult students	❖Mathematical question solving (cognitive task) ❖Sentence copying (physical task)		❖Eye-tracking test ❖NASA-TLX
[90]	9/Workers	Patient-controlled analgesia (PCA) pump programming		❖NASA-TLX ❖Performance
[98]	56/Workers	SensoMotoric Instruments, driving task		❖Eye-tracking test ❖NASA-TLX
[100]	33/Adults	Spatial 2-back		❖EEG ❖NASA-TLX ❖Performance
[85]	39/Adult students	Multi-robot supervision task		❖NASA-TLX ❖post-experiment questionary ❖Performance
[86]	38/Adult students	Multi-Attribute Task Battery (MATB)		❖NASA-TLX ❖SART ❖Performance
[71]	34/Adult students	MATB-II (computer flight simulator)	Subjective	❖NASA-TLX ❖SSS
[94]	39/Workers	Expert Skills Evaluation (EXPERTise 2.0) emergency medicine simulation		❖EXPERTise 2.0 ❖Septris
[95]	46 & 52/Adult students	Rail control simulator		❖EXPERTise 2.0 ❖OSPAN ❖NASA-TLX
[83]	36/Adults	Reading a news article on a smartphone		NASA-TLX
[77]	7/Workers	Hospital Emergency department regular tasks		Observation (NGOMSL)
[74]	22/Adult students	Dynamic Positioning Simulator: Monitor the parameters during a monitoring interval for surpassing threshold values		❖RSME ❖SART
[46]	229/Adult students and workers	Daily engineer job work (1 week study)		SMEs
[89]	36/Workers	Writing one or more medication or fluid orders for administration to a patient while in hospital		❖OSPAN ❖WOMBAT
[79]	36/Workers	Hospital Emergency department: Direct care, indirect care, communication, documentation, in transit and other task		❖WOBAT ❖Modified EDWIN
[70]	160/Workers	❖Reading briefing package ❖Ordering fuel ❖Programming Flight management system		NASA-TLX
[97]	5/Workers	Foundry work for casting of heavy-duty diesel engines		DCog lens
[91]	7/Workers	Operating room surgical workstation		❖Survey ❖Questionnaires ❖Surgery Task Load Index
[93]	3/Adult students	Aerodrome control simulator		NASA-TLX
[80]	33/Adults	Read a set of 10 different texts	Objective	❖Eye-tracking test ❖iMotions
[72]	24/Adults	Automated driving simulator (Green Dino 3): From auditory stimuli generate a verb and repeat a noun		BioSemi Active Two EEG
[75]	18/Adult students	36-layer Lego assembly operation		Polar S810i (HR)
[84]	22/Adults	Email-answering		Eye-tracking test
[92]	20/Workers	Continuous Tracking and Reaction task		Performance

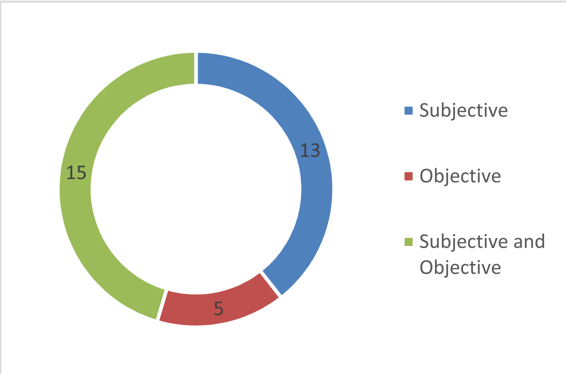


FIGURE 6. Measuring methods.

cognitive task levels in the automated driving simulation [72]. The Feature Recognition Task, Feature Association Task, Feature Prioritization Task, Feature Identification Task, and Feature Discrimination Task were the task levels executed in the study by Falkland et al. study [94]. Most of the reviewed papers conducted studies using computer-based activities.

3) MEASUREMENT METHODS

Many measurement techniques have been examined in the publications under evaluation. These fall into the following categories: objective (Physiological and Performance) and subjective as illustrated in Figure 6. Objective measures are based on quantifiable criteria, which makes them more consistent and precise. Whereas subjective measures which are based on personal opinions or feelings are susceptible to bias and variability due to intrinsic and extrinsic factors [113]. The subjective measurement techniques used in the reviewed articles included the following: NASA-TLX, expert observations (WOMBAT software, Distributed cognition (DCog) lens), Rating Scale Mental Effort (RSME), Derogatis and Melisaratos' Brief Symptom Inventory, Operation Span Task (OSPAN), modified Emergency Department Work Index (EDWIN) score, and Situation Awareness Rating Technique (SART). In addition, heart rate (electrocardiogram (ECG), Polar S810i), skin conductance (electrodermal activity (EDA)), task performance (GoPro, time of completion error, Septris, EXPERTise 2.0, E-Prime 3), electroencephalogram (EEG), eye tracking, iMotion, hierarchical task analysis (HTA), and functional near-infrared spectroscopy (fNIRS) were among the several objective methods used in the reviewed publications.

Only fifteen studies under review used all the matrices (subjective and objective). EDA and ECG (physiological) measurements, primary task performance (error and time completion), and NASA-TLX (subjective) assessments were employed by Doost et al. [54]. An eye tracker (Tobii X3-120) (physiological), NASA-TLX (subjective), and a performance scale were used by Kanaan et al. [87]. Wu et al.'s study used EEG (Emotive EPOC Neuroheadset) (physiological), NASA-TLX (subjective), and a performance scale (error and time) [82]. The performance index (E-Prime 3.0), EEG (64-channel Neuroscan SynAmps2) (physiological), NASA-TLX (subjective), and sleepiness/fatigue (Stanford Sleeping Scale (SSS)) were used by Chen et al. [73]. Midha et al. used fNIRS (Octamon, Artinis Medical Systems) (physiological), NASA-TLX (subjective), and a performance scale (GoPro) [88]. Lee et al. used an eye tracker (SMI RED 250) (physiological), NASA-TLX (subjective), and a performance scale [49].

Among the reviewed papers, thirteen used only subjective measurement methods. Kim et al. used natural goal operator methods, selection rules language (a simulation model to measure the impact of interruption on MWL), and the observation method (two observers followed the participants) [77]. The RSME for measuring MWL and Situation Awareness Rating Technique (SART) [86] were used by Van der Kleij et al. [74]. Fletcher et al. used four subject matter experts (SMEs) (three doctoral candidates and one master's student) and Derogatis and Melisaratos' Brief Symptom Inventory (physical stress measures) [46]. The workflow time study (shadowed over 120 h; direct observation study), OSPAN; working memory capacity and multitasking, WOMBAT software (observer used), and workload index (measured using a modified version of an existing metric) were utilized by Westbrook et al. and Walter et al. [79], [89]. Gontar et al. used the NASA-TLX (subjective) and paper-based observation measure methods in this study [70]. Andreasson et al. used the DCog lens (observational study); the primary resources for information collection were contestant observations, photographs, and field reports [97]. In the performance-based study, Falkland et al. used "Septris" to assess performance following an interruption [94]. The Tobii Pro X3 120 Hz eye tracker (gaze point and eye movements) and iMotions platform (recording face muscles) were used by Lewandowska et al., [80]. Power et al. used HTA (an objective approach to determining breakpoints) and GoPro HERO7 Black (to record the computer screen and keyboard) [81].

D. DATA PRE-PROCESSING

Gathering high-quality data from wearable devices performing real-world activities is complex because of external hardware artifacts, internal signal artifacts from eye movements and flickering, and muscle movements caused by physically challenging duties [114], [115]. Before calculating the participants' cognitive levels, high-quality data must be obtained by eliminating artifacts through adequate pre-processing [101]. The extrinsic artifacts were eliminated using filtering techniques. In particular, a notch filter (60 Hz) was employed to eliminate electrode wire noise and essential artifacts. Another name for a notch filter is a band reject filter or band stop filter. These filters let signals above and below the stop band frequency range through while rejecting signals in the band. For instance, eye blinking, movement, and muscle motion were dislodged using independent component analysis (ICA) and a low-pass filter (cutoff 65 Hz) to eliminate motion artifacts and information aliasing. ICA aims to decompose the recorded EEG signals into a set of statistically independent components, each of which is presumed to correspond to the activity of different brain sources. The fundamental assumption behind ICA is that the EEG signals recorded at different scalp electrodes represent linear mixtures of the underlying neural activities originating from various brain regions. A high-pass filter (cutoff > 0.5 Hz) was used to eliminate artifacts from gradual signal changes, sweat, and electrode drifts [73], [76], [82], [100]. The data were downsampled to 256 Hz for further calculations [72], [84]. The interquartile range [76] and wavelet filtering [73], [88] were used to remove muscular and motion artifacts. Wavelet filter operates based on the principles of wavelet analysis, which involves decomposing a signal into its constituent frequency components at different scales. Wavelet filters can effectively remove noise from signals by exploiting the differences in frequency content between the signal and the noise. The signal is decomposed into wavelet coefficients representing various frequency bands, and then a thresholding technique is applied to remove coefficients associated with noise. Gratton and Coles' ocular correction was used to compensate for eye movements [72], [84], and Goldberg and Kotval used a fixation algorithm for blinks [78], [84], [87]. A one-dimensional data interpolation method for a uniform sampling rate and a single imputation method for filling in missing data were also used [49]. Cohen's kappa was used to divide the data into 1 second [89]. Cohen's kappa is widely used in various fields, including psychology, medicine, and machine learning, to assess the reliability and agreement between raters or classification systems, particularly when dealing with categorical data or nominal variables. It provides a more robust measure of agreement than simple percent agreement, especially when dealing with imbalanced datasets or when chance agreement is a concern. Homer2 was used to process raw fNIRS data [88]. Homer2 is a software package widely used in the field of fNIRS data analysis. It provides a comprehensive set of tools for preprocessing, analyzing, and visualizing fNIRS data. Homer2 offers various preprocessing functions to clean and prepare fNIRS data for analysis. This includes procedures for motion artifact correction, signal filtering, and baseline correction. It also provides tools for conducting statistical analyses on fNIRS data, such as general linear model analysis, correlation analysis, and group-level statistics. These analyses help researchers identify significant changes in brain activity in response to experimental conditions or stimuli.

E. COMPUTATION MODELS

Importance was assessed to ascertain whether there were meaningful variations between the control and experimental groups. Numerous significance test techniques, including the Kruskal-Wallis test, Analysis of the variance (ANOVA) test [80], k-means cluster test [94], [95], Mann-Whitney inverted U-test [83], emotion analysis (facial actions coding system) [80], time analysis [77], confirmatory factor analysis [46], bivariate correlation [70], Wilcoxon signed-rank test [87], and paired sample t-test, have been applied to analyze participants' mental status and performance calculation. A one-way ANOVA was utilized to analyze the difference wave (novel standard, expressed in μV) [72], examine the difference in perceptions of the interruptions that occurred in different phases, and examine whether there was a difference in overall workload across the four conditions [82], test whether there were significant differences in brain activity between the easy, medium, and hard difficulty reading and writing tasks [88], determine the relationship between driving experience and cue utilization [95], compare nurse frustration (NASA-TLX Subscale 6) across the three interruption conditions, compare the total task completion times, and compare full cognitive workload scores for the NASA-TLX across the three interruption conditions [90]. A two-way repeated ANOVA was used to evaluate the impacts of the task performance condition and task cognitive level on dependent variables; [54] task performance time on the system dialogue messages; [83] find interactions among camera view, rotation angle, and interruption duration; [96] ascertain the impacts of influences on dependent variables [74], performed for subjects' answers to the NASA-TLX questionnaire; and determine [49] the relationship between the interruption management system and workload [84]. Mixed repeated measures ANOVA was used to examine whether differences in cue utilization were associated with differences in Septris scores before and after the interruption [94], to explore the interaction effects between the interruption phase (before the interruption, after interruption) and workload (low, high) [87], and to examine whether differences in cue utilization were associated with differences in response latencies before and after disruption in the train control simulation [95].

The paired sample t-test was used to analyze the following: variations in time spent attending to interruptions due to the order of presentation [81], Stanford Sleepiness Scale [73] score, pupil diameter between different conditions [78], effect of interruptions across conditions on post-interruption time per layer [75], and resumption lag difference between the IMS and random conditions [84]. Statistical analysis was performed using IBM SPSS version SAS V.9.4 and R software. The eye-tracking analysis models used were MATLAB (eye-tracking metrics) [87] and ClearView [78]. The different cognitive signal analysis models used were the BrainVision Analyzer 2.1 [72], Fast Fourier transformation [82], EEGLAB toolbox of MATLAB R2020a [73], [100], Modified Beer-Lambert Law (fNIRS) [88], and machine learning [76].

F. EFFECT OF INTERRUPTIONS

Result outcomes of the interruptions effect on the cognitive states and measuring methods evaluated from the articles under review are illustrated in Table 3. Doost et al observed that social media interruptions had a significant impact on HR and skin conductance levels, indicating increased MWL and stress, as well as decreased performance in knowledge-based tasks when interrupted by social media [54]. The Functional Resonance Analysis Method designed scenarios showed that the models' capability for task transfer was demonstrated through the results, which revealed that the detected workload increased with interruptions; in addition, the convolutional neural network model demonstrated high sensitivity and specificity in workload detection during n-back-test evaluations, suggesting its ability to generalize across different subjects and transfer tasks to new environments [76]. Interruptions at fine breakpoints resulted in longer resumption lags and higher MWL compared to coarse breakpoints, indicating that the disruptiveness of an interruption is tied to the point within the task hierarchy where it occurs [81]. Participants experienced a greater decline in primary task performance due to interruptions under high workload conditions compared to low workload conditions, as evidenced by decreased scanpath length per second and mean saccade amplitude; this effect was not observed in participants with low workload, indicating a more focused but narrower visual search strategy in high workload scenarios. Participants reported better performance under low workload conditions based on the Performance subscale of the NASA-TLX [87].

Participants who were interrupted during the evaluation-and-selection phase had worse task performance, while interruptions during the problem-identification phase resulted in increased MWL and negative perceptions of interruptions. Interruptions during the alternative-development phase caused changes in arousal and valence, and EEG measures showed that interruptions affected MWL and emotional states, particularly in the frontal lobes' theta band. Despite poorer performance, participants interrupted during the evaluation-and-selection phase still had positive feelings and confidence towards interruptions, suggesting a disconnect between subjective experiences and actual performance impact [82]. Interruptions were found to increase alpha activity and P300 amplitude, indicating improved inhibitory control and attentional reallocation, as well as increased theta power and a speeding-up effect post-interruption, although fatigue negatively affected cognitive abilities and worsened the impact of interruption on working memory and behavioral performance; moreover, interruptions were shown to impose a higher cognitive burden compared to suspension and baseline tasks [73]. Interruptions during tasks were found to cause noticeable changes in MWL, as indicated by significant differences in oxygenated hemoglobin (O2Hb) and deoxygenated hemoglobin (HHb) levels, however, there was no significant difference in brain activity detected by fNIRS, highlighting a discrepancy between subjective and objective assessments of workload during writing tasks [88]. The length of interruptions significantly affects task resumption, with shorter interruptions resulting in quicker resumption; the side view resulted in longer resumption times and was rated as higher in mental demand, aligning with subjective workload ratings; participants perceived better performance with the top view [96].

Participants exhibited more dilated pupils in cognitive primary task sets compared to skill primary task sets, suggesting higher cognitive workload in the former. Participants reported (NASA-TLX) the highest subjective workload in the immediate interruption mode and the lowest in the negotiated mode [49]. The study revealed that an increase in interruptions led to a significant increase in completion time for nurses using PCA pumps, and although there was a positive correlation between interruptions and cognitive workload, the results were not statistically significant due to limited sample size and data variability; furthermore, the study found that post-interruption errors resulted in narcotic overdosing, a serious concern for patient safety, and familiarity with the PCA pump did not reduce the likelihood of errors [90]. Participants' eye movements revealing that visual interruptions negatively impacted the ability to anticipate hazards. NASA-TLX scale showed that higher MWL and effort were reported by participants in the visual interruption conditions. Spatial and non-spatial task conditions resulted in a higher number of glances needed to identify the hazard compared to the gray-screen condition [98]. EEG data revealed a significant increase in P2 and P3 amplitudes after interruptions during the 2-back task, suggesting enhanced cognitive processing to reestablish task goals. There was an observed increase in theta and alpha power spectra post-interruption, which may reflect the cognitive demands associated with reorienting attention and re-engaging working memory resources. Participants reported a higher subjective workload following interruption conditions compared to suspension and baseline conditions [100].

A significant relationship was found in the 2-Way ANOVA between the number of robots and the percentage of faults reported, indicating that as the number of robots increased, performance decreased; participants consistently reported a higher perceived workload during extrinsic interruptions regardless of the number of robots being monitored [85]. Chen et al. discovered that interruptions have a greater negative effect on performance when tasks are complex, and mental fatigue worsens the negative impact of interruptions by affecting primary task performance, subjective workload, and resumption lag. Additionally, resuming complex tasks requires more time due to the increased cognitive load [71]. The duration of the interruption task and the resumption lag were positively correlated with the subjective workload upon resumption for older participants, implying that interruptions exacerbate the mental load during task switching [83]. The simulation results revealed that interruptions significantly increased nurses' MWL, with 2.04 times increase during patient care activities and 4.72 times increase during EMR charting [77]. The results indicated that change detection support is beneficial in recovering situation awareness after an interruption, but it also increased workload in non-interrupted conditions, suggesting that continuous support may not always be advantageous [74]. The subjective experience of interruptions has a significant impact on an individual's stress levels, supporting the theory that perception of a stressor is a more proximal predictor of stress outcomes than the objective stimulus [46].

Working memory capacity was found to have a protective effect against errors, with a 10-point increase on the OSPAN

TABLE 3. Effect of interruptions in the publications under examination.

Articles	Effect of Interruptions	Effect on Measuring method	
		Objective	Subjective
[54]	Increase cognitive load	Higher level ECG and EDA in workload task. Lower performance when interrupted	Interrupted task was rated as having the highest cognitive load
[76]		EEG showed high sensitivity and specificity in workload detection during n-back-test evaluations	Interrupted n-back-test was rated as having the highest cognitive load
[81]		Interruption increased resumption lab and decreased performance	Cognitive load was high during interruptions
[87]		The eye tracking results revealed in high workload conditions noted shorter saccades. High workload conditions negatively impact interruptions on primary task performance	Performed decreased under high workload conditions
[82]		EEG revealed that interruptions induce changes in MWL and emotional states. Exhibited poorer task performance	Despite poorer task performance
[73]		Increased theta power (EEG) during interruptions, implies a greater cognitive demand	Interruptions impose a greater cognitive burden
[88]		fNIRS data indicated significant sensitivity to MWL variations with increased O2Hb and decreased HHb	Hard condition rated significantly higher in mental demand
[96]		Requiring less time to resume the peg transfer and perform better when using the top view	The side view as higher in mental demand, and frustration compared to the top view
[49]		Participants exhibited more dilated pupils in cognitive primary task sets suggesting higher cognitive workload	Highest workload in the immediate interruption mode and the lowest in the negotiated mode
[90]		Performance decreased with increase in interruptions	Cognitive load increased with interruptions
[98]		Visual interruptions negatively impacted the ability to anticipate hazards	Higher MWL and effort were reported
[100]		Increase in P2 and P3 amplitudes (EEG) after interruptions suggesting enhanced cognitive load. Task performance decreased	Higher subjective workload following interruption conditions
[85]		Performance decreasing as the number of robots increased	Higher workload during extrinsic interruptions
[71]			Mental fatigue negatively affects primary task performance and workload
[83]			Interruptions exacerbate the mental load during task switching
[77]			Interruptions on nurses' MWL in emergency departments
[74]			Change detection increased workload in non-interrupted conditions
[46]			Interruptions can have a considerable impact on an individual's stress levels
[89]			Clinical error rates increased
[70]			Interruption increases operator workload
[97]			Interruption decreases performance
[93]			Interruption increases workload
[80]		Negative affective content shows lower cognitive load (Eye-tracking) and elicited stronger emotional (iMotions)	
[72]		fP3 neurophysiological (EEG) response was lower when engaged in automated driving with an additional cognitive load	
[75]		Complex interruptions might promote cognitive arousal	
[92]		Visual notifications decrease performance during high workloads	
[84]	Decreases cognitive load	IMS can effectively minimize the disruptiveness of interruptions	
[78]		Cognitive load was higher during interruption that was foreseen and was lower following a forewarned interruption	
[86]		Situation awareness and takeover performance was enhanced with the two-stage warning system	MWL scores were significantly lower when using the two-stage warning system
[94]			Higher cue utilization scores demonstrated better performance after an interruption
[95]			Higher cue utilization was associated with a reduced cognitive load
[79]			Communicative prompts decreased cognitive load
[91]			Integrated workstation for displaying alarms reduces workload and minimizing errors

resulting in a 19% decrease in prescribing errors; clinical error rates increased with patient age and physician age, but decreased with doctor seniority, and resident medical officers had the highest error rate compared to consultants; junior doctors had lower rates of legal procedural errors compared to senior colleagues, and a 10-point increase in OSPAN performance led to a 19% reduction in legal procedural error rate [89]. A high frequency of disruptions that pilots must manage was observed, with an average of 7.9 interruptions per pilot per turn-around, suggesting a substantial volume of interruptions; the relationship between interruptions and operator workload accounted for 11.3% of its variance, but factors like time pressure, weather conditions, and non-routine events increased the variance explained in operator workload significantly [70]. Interruptions were found to significantly affect the distributed workload within the socio-technical system, impacting the overall production performance at the casting line [97]. Özdemir et al. found positive correlations between the number of aircraft, total time, exercise duration, and instruction count with many of the task load indexes of pseudo-pilots, indicating that an increase in these factors leads to an increase in workload. The study successfully identified the number of aircraft, air traffic control trainee performance, and the interruption duration as the key factors affecting the overall workload of pseudo-pilots [93]. The negative content received more attention, and had less mental effort compared to positive content, as detected by the iMotions Facial Expression Analysis Module, which showed significant emotional response differences such as greater anger, sadness, and contempt for negative pictures and higher levels of joy and surprise for positive pictures [80]. The frontal P3 (fP3) neurophysiological response was significantly decreased during automated driving with additional cognitive load, indicating reduced susceptibility to auditory stimuli, regardless of the difficulty level of the cognitive task [72]. The potential increase in productivity resulting from interruptions in assembly operations due to cognitive arousal was not directly measured in terms of error rates or time metrics [75]. The notifications, regardless of how they were presented, had a detrimental effect on the performance of the main task, with visual notifications causing more distraction than audio notifications [92].

The IMS accurately identified periods of high workload 62.2% of the time and low workload 37.8% of the time, even with a 1-second delay in pupil dilation response, indicating its ability to detect brief decreases in pupil size during high workload moments and strategically time interruptions based on real-time workload assessment through pupil dilation [84]. The cognitive load associated with task resumption was reduced when participants received a warning before the interruption, as indicated by pupil diameter measurements taken before the first post-interruption decision [78]. The utilization of a two-stage warning system resulted in a decrease in mental strain and an improvement in situation awareness compared to a single-stage system, indicating that the more complex warning system may better prepare individuals for the takeover task [86]. The physicians with higher cue utilization scores demonstrated significantly better performance on the simulation task after an interruption, compared to their lower cue utilization counterparts [94]. Participants who made better use of cues were able to perform consistently even when interrupted, while those who did not make good use of cues experienced longer response times and more errors, and also reported feeling a higher workload [95]. The positive effects of communicative prompts, such as timely information transfer and advice provision, suggest that while they contribute to cognitive load, they also play a crucial role in workload management and patient treatment [79]. The use of an integrated workstation for displaying alarms may enhance patient safety by reducing workload and minimizing errors in the operating room, provided that the alarm system design is further optimized [91].

V. FURTHER DISCUSSION

A. IMPLICATIONS

Human cognitive capacity is important because safety and efficiency in the workplace are highly regarded. Therefore, it is essential to precisely measure and calculate workers' cognitive conditions to improve job management. Self-reported questionnaires have been extensively used to measure employees' cognitive health; however, occasionally, they cannot provide accessible, unbiased, and accurate findings. In interrupted work situations, employees' cognitive states are anticipated to be computed precisely and objectively due to improvements in wearable physiological monitoring devices and related calculation algorithms [88]. The ability of physiological equipment to calculate the distinct cognitive states of employees under varied job cognitive levels has been demonstrated in research articles. The analysis supports using physiological measuring equipment to determine workers' cognitive conditions in an interrupted situation.

This study evaluated tension, exhaustion, emotion, working memory performance, task performance, and resuming lag with MWL in interrupted work settings. The investigator should choose subjects with an adequate sample size and generalizability while planning an experiment, and they should receive proper pre-experiment instructions. fNIRS, a non-invasive and movement-friendly brain scanner, is a potential method for measuring brain activity [88]. fNIRS uses near-infrared light to track changes in blood oxygenation in the brain. Neurovascular coupling can be used to evaluate cognitive performance when active brain areas require more blood flow to fulfill their higher energy needs. Task design must consider the accuracy and resemblance of the data to the actual job tasks. Sometimes, it may be helpful to use virtual reality and augmented reality to create hazardous or other activities that cannot be conducted in a lab. Several tools can be used to identify physiological parameters, including self-reported surveys, reaction times, and perception task precision. Before conducting data analysis, improving the raw data quality by using a bandpass filter, eliminating outliers, data downsampling, and ICA is possible. In addition, statistical and machine learning models perform well when computing employee cognitive states in interrupted real-world work environments.

Simulating the impact of interruptions and job complexity on cognitive states has three key benefits:

1. Cognitive activity can be measured directly using physiological equipment. The most direct measurement techniques are voltage fluctuations recorded by the EEG and blood hemoglobin variations identified by fNIRS. Although this process is not fully understood, the EEG and fNIRS signal patterns have been extensively studied and can be predicted accurately.
2. Physiological tools make it possible to track the evolution of cognitive processes. Most cognitive processes occur in consecutive sequences across timescales ranging from a few seconds to milliseconds.
3. The physiological device's signal contains more detailed information because it includes many dimensions, such as magnitude, frequency, power, and phase.

Two or more physiological tools, such as EEG and eye tracking, can be used concurrently to gather additional cognitive data. This rich information simplifies the construction and exercise of numerous computational models [116].

An experiment design framework was built to investigate the effects of interruptions and cognitive task levels on mental status. Some tasks, particularly those with high workloads, are unacceptable to workers with impaired cognitive function. Real-time cognitive status monitoring and identification of workers with poor mental health are possible using physiological methods. Managers can alter their work to avoid accidents and underperformance. In addition, prior research has shown that the type of interruption, task level, and work environment are all directly connected to the cognitive conditions of workers [54], [73], [80], [88]. Physiological monitoring tools can provide unbiased evaluations of these working circumstances and useful guidelines for improving job management.

B. LIMITATIONS

Because studies on the effects of interruptions and cognitive task levels on mental status using physiological measuring equipment are still in their early stages, important limitations of existing studies must be addressed. These restrictions include ambiguity in the study question, poor design of the experiments, the use of subpar data processing techniques, and a lack of sufficient computer capacity. The following describes these details:

1) Different approaches have been used to compute the various characteristics of employees affected by job interruptions, such as working memory performance, MWL, emotions, resumption lags, performance, stress, and fatigue. Nevertheless, employees' cognitive processes are complicated and impacted by various internal and external stimuli (e.g., interruption type, task level, and working environment). There is much connectivity among these cognitive states. However, earlier research did not sufficiently explain the problem and offered a framework for computing workers' mental states. How do these characteristics affect employee cognitive processes? How do the related cognitive states interact with one another? However, these questions remain to be answered.
2) Measuring the cognitive status of workers can greatly benefit from various computational approaches (both subjective and objective). However, their practical applicability may be questioned if real-world events cannot be accurately replicated. However, the experimental design used in this study had several limitations. First, most current research enrolls only undergraduates with small subject populations. They did not check whether the findings and conclusions were appropriate for the workers. In addition, many researchers have focused on sanitized activities, such as computer-based simulations. Consequently, the experiment lasted only for a short time. The atmosphere in a lab is quite different from that at work. Such study designs cannot replicate the problematic, diverse, and demanding labor that employees regularly perform, which affects the accuracy of the computed findings.
3) An effective preprocessing approach for physiological devices was employed in earlier research to obtain high-quality data while working in an interrupted environment and to eliminate artifacts. It still does not work well when the scalp sweats significantly and the head needs to move regularly, which is unavoidable in real-world activities. Physiological signal pre-processing techniques are an additional barrier to determining the cognitive conditions of workers during work interruptions.
4) Although machine learning algorithms are effective for analyzing large amounts of data, they still need to be adequately applied to determine how interruptions affect workers' cognitive states. Traditional machine learning methods cannot be trained independently because their effectiveness frequently depends on manually created features [117]. Consequently, deep learning algorithms that can classify and understand features together have been developed [118], [119]. Deep neural networks can complete various tasks, including classification, regression, and generation, using manually derived features as input variables or directly ingesting raw data [120]. Deep algorithmic approaches for analyzing and decoding physiological data have recently gained popularity as a research area.

C. FUTURE RESEARCH RECOMMENDATIONS

The following is a summary of the recommendations for future studies to fully understand the impact of interruptions on cognitive states in a real-world environment.

1) It is essential to consider the important variables (e.g., interruption type, task level, and working environment) and how they affect the cognitive state and other parameters (MWL, emotions, resumption lag, performance, stress, fatigue, and working memory performance). These considerations could offer managers helpful

guidelines for enhancing working circumstances and maintaining employees' cognitive health, necessitating theoretical and experimental neuroscientific studies.

2) Future research should incorporate more accurate simulations of real-world events into experiments to calculate workers' cognitive states in interrupted work contexts. To select adequate and representative subjects for research, the researcher must first address the limitations of selecting participants from prior studies. The impact of interruptions on individuals should be studied using physiological, behavioral, and subjective metrics. Instead of using streamlined laboratory activities, experimental tasks and settings should reflect actual working situations by providing a more compelling and realistic setting for exploratory experimentation.
3) Future research should concentrate on physiological measurement device data preprocessing under interrupted work scenarios. Due to the near impossibility of avoiding head movements and perspiration in workers' everyday jobs, improved algorithms are required to lessen their adverse effects on the quality of physiological data. This approach can also improve preparedness for research tasks and everyday activities.
4) Future research should focus on enhancing the precision and effectiveness of the computing model to fulfill the demands of real-time supervision of individuals' mental statuses established using physiological measurement methods. Numerous deep learning methods, including convolutional neural networks [121], [122], recurrent neural networks [123], [124], and graph neural networks [125], [126], have been used in EEG-based computations in different disciplines and have outperformed conventional machine learning methods in terms of performance. Future research should consider implementing standard deep learning techniques to increase model precision and efficacy in physiological measure-based assessments of employee cognitive state computations in an interrupted workplace environment.

VI. CONCLUSION

This manuscript conducted a thorough examination of the existing literature concerning the effects of interruptions and mental task intensity on cognitive function, with a specific focus on the evaluation of cognitive state, experimental methodologies, and analytical strategies. The incorporated studies affirmed the prospects of physiological, behavioral, and subjective techniques to measure various parameters for workers affected by work interruptions, such as MWL, emotional state, resumption lag, performance, stress, fatigue, and working memory performance. A good and practical experimental design should consider participant assortment, task layout, and measurement technique assortment to improve the characteristics of the data gathered. The dataset was pre-processed before calculating employees' cognitive levels to eliminate the included distortions. The cognitive states of the workers can then be calculated using data from both conventional statistical approaches and sophisticated machine learning modes.

Additionally, it identified critical problems with the experimental design (participant selection, task model, and measurement method) and data analysis (pre-processing and cognitive status computation). The analysis also highlights areas of research that need to be addressed in future investigations. Studies on workers' cognitive status computation during work interruptions are presented in the current review. It may be a systematic direction for researchers, particularly those new to the topic, to understand how to begin work interruption studies. It offers valuable advice to project managers for enhancing their work management.

Nevertheless, previous studies have not adequately explained the connection between cognitive states influenced by internal and external stimuli. Existing experimental research designs cannot replicate interrupted work scenarios. Relevant algorithms cannot address the efficiency of real-time interrupted scenario monitoring or the vulnerability of physiological signal measurements. In the interrupted work scenario, further research is required for theory development, experimental optimization, and algorithm development.

REFERENCES

[1] J. Laarni, "Multitasking and interruption handling in control room operator work," in *Human Factors in the Nuclear Industry: A Systemic Approach To Safety*. Cambridge, U.K.: Woodhead Publishing, Jan. 2021, pp. 127–149, doi: 10.1016/B978-0-08-102845-2.00007-7.

[2] D. A. Boehm-Davis and R. Remington, "Reducing the disruptive effects of interruption: A cognitive framework for analysing the costs and benefits of intervention strategies," *Accident Anal. Prevention*, vol. 41, no. 5, pp. 1124–1129, Sep. 2009, doi: 10.1016/j.aap.2009.06.029.

[3] L. Coraggio. (1990). *Deleterious Effects of Intermittent Interruptions on the Task Performance of Knowledge Workers: A Laboratory Investigation*. Accessed: Nov. 19, 2022. [Online]. Available: https://search.proquest.com/openview/010d0c5973d158d04148e289a70f06c3/1?pq-origsite=gscholar&cbl=18750&diss=y

[4] S. Spiekermann, "Attention & interruption management for systems design—A research overview," *SSRN Electron. J.*, vol. 1, pp. 1–74, Aug. 2008, doi: 10.2139/ssrn.1244622.

[5] V. M. González and G. Mark, "'Constant, constant, multi-tasking craziness': Managing multiple working spheres," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Austria, Apr. 2004, pp. 113–120, doi: 10.1145/985692.985707.

[6] S. Leroy and T. M. Glomb, "Tasks interrupted: How anticipating time pressure on resumption of an interrupted task causes attention residue and low performance on interrupting tasks and how a 'ready-to-resume' plan mitigates the effects," *Org. Sci.*, vol. 29, no. 3, pp. 380–397, Jun. 2018, doi: 10.1287/orsc.2017.1184.

[7] J. Wajcman and E. Rose, "Constant connectivity: Rethinking interruptions at work," *Org. Stud.*, vol. 32, no. 7, pp. 941–961, Jul. 2011, doi: 10.1177/0170840611410829.

[8] J. B. Spira and J. B. Feintuch. (2005). *The Cost of Not Paying Attention: How Interruptions Impact Knowledge Worker Productivity*. Rep. From Basex. Accessed: Nov. 19, 2022. [Online]. Available: http://iorgforum.org/wp-content/uploads/2011/06/CostOfNotPayingAttention.BasexReport1.pdf

[9] P. Carayon, "Stressful jobs and non-stressful jobs: A cluster analysis of office jobs," *Ergonomics*, vol. 37, no. 2, pp. 311–323, Feb. 1994, doi: 10.1080/00140139408963648.

[10] Y.-Y. Yeh and C. D. Wickens, *An Investigation of the Dissociation Between Subjective Measures of Mental Workload and Performance*. Urbana, IL, USA: Univ. Illinois, 1984.

[11] P. Carayon, T. B. Wetterneck, A. S. Hundt, M. Ozkaynak, J. DeSilvey, B. Ludwig, P. Ram, and S. S. Rough, "Evaluation of nurse interaction with bar code medication administration technology in the work environment," *J. Patient Saf.*, vol. 3, no. 1, pp. 34–42, Mar. 2007, doi: 10.1097/pts.0b013e3180319de7.

[12] E. B. Cutrell, M. Czerwinski, and E. Horvitz, "Effects of instant messaging interruptions on computing tasks," in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst.*, Apr. 2000, pp. 99–100, doi: 10.1145/633292.633351.

[13] T. Gillie and D. Broadbent, "What makes interruptions disruptive? A study of length, similarity, and complexity," *Psychol. Res.*, vol. 50, no. 4, pp. 243–250, Apr. 1989, doi: 10.1007/bf00309260.

[14] E. Cutrell, M. Czerwinski, and E. Horvitz, "Notification, disruption, and memory: Effects of messaging interruptions on memory and performance," in *Proc. INTERACT IFIP TC Int. Conf. Hum.-Comput. Interact.*, Jul. 2001, p. 263.

[15] H. Eyrolle and J.-M. Cellier, "The effects of interruptions in work activity: Field and laboratory results," *Appl. Ergonom.*, vol. 31, no. 5, pp. 537–543, Oct. 2000, doi: 10.1016/s0003-6870(00)00019-3.

[16] H. Alm and L. Nilsson, "The effects of a mobile telephone task on driver behaviour in a car following situation," *Accident Anal. Prevention*, vol. 27, no. 5, pp. 707–715, Oct. 1995, doi: 10.1016/0001-4575(95)00026-v.

[17] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An integrated theory of the mind," *Psychol. Rev.*, vol. 111, no. 4, pp. 1036–1060, Oct. 2004, doi: 10.1037/0033-295x.111.4.1036.

[18] B. P. Bailey and S. T. Iqbal, "Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management," *ACM Trans. Comput.-Hum. Interact.*, vol. 14, no. 4, pp. 1–28, Jan. 2008, doi: 10.1145/1314683.1314689.

[19] S. K. Card, T. P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*, 1st ed. Boca Raton, FL, USA: Lawrence Erlbaum Associates, 1986, doi: 10.1201/9780203736166.

[20] H. Puranik, J. Koopman, and H. C. Vough, "Pardon the interruption: An integrative review and future research agenda for research on work interruptions," *J. Manage.*, vol. 46, no. 6, pp. 806–842, Jul. 2020, doi: 10.1177/0149206319887428.

[21] C. Speier, I. Vessey, and J. S. Valacich, "The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance," *Decis. Sci.*, vol. 34, no. 4, pp. 771–797, Nov. 2003, doi: 10.1111/j.1540-5414.2003.02292.x.

[22] F. Nachreiner, "Standards for ergonomics principles relating to the design of work systems and to mental workload," *Appl. Ergonom.*, vol. 26, no. 4, pp. 259–263, Aug. 1995, doi: 10.1016/0003-6870(95)00029-c.

[23] H. A. Maior, M. Pike, S. Sharples, and M. L. Wilson, "Examining the reliability of using fNIRS in realistic HCI settings for spatial and verbal tasks," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, Apr. 2015, pp. 3039–3042, doi: 10.1145/2702123.2702315.

[24] Y. Lean and F. Shan, "Brief review on physiological and biochemical evaluations of human mental workload," *Hum. Factors Ergonom. Manuf. Service Industries*, vol. 22, no. 3, pp. 177–187, May 2012, doi: 10.1002/hfm.20269.

[25] M. A. Vidulich and P. S. Tsang, "The confluence of situation awareness and mental workload for adaptable human-machine systems," *J. Cogn. Eng. Decision Making*, vol. 9, no. 1, pp. 95–97, Mar. 2015, doi: 10.1177/1555343414554805.

[26] K. Ryu and R. Myung, "Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic," *Int. J. Ind. Ergonom.*, vol. 35, no. 11, pp. 991–1009, Nov. 2005, doi: 10.1016/j.ergon.2005.04.005.

[27] N. Dadashi, A. W. Stedmon, and T. P. Pridmore, "Semi-automated CCTV surveillance: The effects of system confidence, system accuracy and task complexity on operator vigilance, reliance and workload," *Appl. Ergonom.*, vol. 44, no. 5, pp. 730–738, Sep. 2013, doi: 10.1016/j.apergo.2012.04.012.

[28] B. Cinaz, B. Arnrich, R. La Marca, and G. Tröster, "Monitoring of mental workload levels during an everyday life office-work scenario," *Pers. Ubiquitous Comput.*, vol. 17, no. 2, pp. 229–239, Feb. 2013, doi: 10.1007/s00779-011-0466-1.

[29] A. Johnson and A. Widyanti, "Cultural influences on the measurement of subjective mental workload," *Ergonomics*, vol. 54, no. 6, pp. 509–518, Jun. 2011, doi: 10.1080/00140139.2011.570459.

[30] S. Jo, R. Myung, and D. Yoon, "Quantitative prediction of mental workload with the ACT-R cognitive architecture," *Int. J. Ind. Ergonom.*, vol. 42, no. 4, pp. 359–370, Jul. 2012, doi: 10.1016/j.ergon.2012.03.004.

[31] D. K. Mitchell, *Mental Workload and ARL Workload Modeling Tools*. Aberdeen Proving Ground MD, USA: Army Research Lab, Apr. 2000, doi: 10.21236/ada377300.

[32] C. Lebiere, "A theory-based model of cognitive workload and its applications," in *Proc. Interservice/Ind. Training, Simulation Educ. Conf. (I/ITSEC)*, Arlington, TX, USA: National Training Systems Association, 2001, pp. 1–5. Accessed: Nov. 19, 2022. [Online]. Available: http://www.iitsecdocs.com/volumes/2001

[33] L. L. Di Stasi, A. Antolí, and J. J. Cañas, "Main sequence: An index for detecting mental workload variation in complex tasks," *Appl. Ergonom.*, vol. 42, no. 6, pp. 807–813, Nov. 2011, doi: 10.1016/j.apergo.2011.01.003.

[34] P. Lehrer, M. Karavidas, S.-E. Lu, E. Vaschillo, B. Vaschillo, and A. Cheng, "Cardiac data increase association between self-report and both expert ratings of task load and task performance in flight simulator tasks: An exploratory study," *Int. J. Psychophysiol.*, vol. 76, no. 2, pp. 80–87, May 2010, doi: 10.1016/j.ijpsycho.2010.02.006.

[35] B. Z. Allison and J. Polich, "Workload assessment of computer gaming using a single-stimulus event-related potential paradigm," *Biol. Psychol.*, vol. 77, no. 3, pp. 277–283, Mar. 2008, doi: 10.1016/j.biopsycho.2007.10.014.

[36] T. Oron-Gilad, J. L. Szalma, S. C. Stafford, and P. A. Hancock, "The workload and performance relationship in the real world: A study of police officers in a field shooting exercise," *Int. J. Occupational Saf. Ergonom.*, vol. 14, no. 2, pp. 119–131, Jan. 2008, doi: 10.1080/10803548.2008.11076757.

[37] C. D. Wickens, J. G. Hollands, S. Banbury, and R. Parasuraman, *Engineering Psychology and Human Performance*, 4th ed. New York, NY, USA: Psychology Press, 2013, doi: 10.4324/9781315665177.

[38] C. D. Wickens, "Multiple resources and mental workload," *Human Factors: J. Human Factors Ergonom. Soc.*, vol. 50, no. 3, pp. 449–455, Jun. 2008, doi: 10.1518/001872008x288394.

[39] S. Sharples and T. Megaw, "Definition and measurement of human workload," in *in Evaluation of Human Work*, 4th ed., J. R. Wilson and S. Sharples, Eds., Boca Raton, FL, USA: CRC Press, 2015, pp. 515–549. Accessed: Nov. 19, 2022. [Online]. Available: https://www.perlego.com/book/1603082/evaluation-of-human-work-pdf

[40] J. Engström, G. Markkula, T. Victor, and N. Merat, "Effects of cognitive load on driving performance: The cognitive control hypothesis," *Human Factors, J. Hum. Factors Ergonom. Soc.*, vol. 59, no. 5, pp. 734–764, Feb. 2017, doi: 10.1177/0018720817690639.

[41] J.-F. Stich, "A review of workplace stress in the virtual office," *Intell. Buildings Int.*, vol. 12, no. 3, pp. 208–220, Jul. 2020, doi: 10.1080/17508975.2020.1759023.

[42] Q. R. Jett and J. M. George, "Work interrupted: A closer look at the role of interruptions in organizational life," *Acad. Manage. Rev.*, vol. 28, no. 3, p. 494, Jul. 2003, doi: 10.2307/30040736.

[43] S. Leroy, A. M. Schmidt, and N. Madjar, "Working from home during COVID-19: A study of the interruption landscape," *J. Appl. Psychol.*, vol. 106, no. 10, pp. 1448–1465, Oct. 2021, doi: 10.1037/apl0000972.

[44] G. Mark, V. M. Gonzalez, and J. Harris, "No task left behind? Examining the nature of fragmented work," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Portland, OR, USA: Association for Computing Machinery (ACM), Apr. 2005, pp. 321–330, doi: 10.1145/1054972.1055017.

[45] F. Sasangohar, B. Donmez, A. C. Easty, and P. L. Trbovich, "Effects of nested interruptions on task resumption: A laboratory study with intensive care nurses," *Human Factors, J. Hum. Factors Ergonom. Soc.*, vol. 59, no. 4, pp. 628–639, Jan. 2017, doi: 10.1177/0018720816689513.

[46] K. A. Fletcher, S. M. Potter, and B. N. Telford, "Stress outcomes of four types of perceived interruptions," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 60, no. 2, pp. 222–235, Mar. 2018, doi: 10.1177/0018720817738845.

[47] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognit. Sci.*, vol. 12, no. 2, pp. 257–285, Jun. 1988, doi: 10.1016/0364-0213(88)90023-7.

[48] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, Apr. 1988, doi: 10.1016/S0166-4115(08)62386-9.

[49] B. Lee, K. Chung, and S.-H. Kim, "Interruption cost evaluation by cognitive workload and task performance in interruption coordination modes for human–computer interaction tasks," *Appl. Sci.*, vol. 8, no. 10, p. 1780, Sep. 2018, doi: 10.3390/app8101780.

[50] K. A. Brookhuis and D. de Waard, *Assessment of Drivers' Workload: Performance and Subjective and Physiological Indexes*, 1st ed. Boca Raton, FL, USA: CRC Press, 2001, doi: 10.1201/B12791-2.5.

[51] A. H. Roscoe, "Assessing pilot workload. Why measure heart rate, HRV and respiration?" *Biol. Psychol.*, vol. 34, nos. 2–3, pp. 259–287, Nov. 1992, doi: 10.1016/0301-0511(92)90018-p.

[52] D. W. Rowe, J. Sibert, and D. Irwin, "Heart rate variability: Indicator of user state as an aid to human-computer interaction," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (CHI)*, Los Angeles, CA, USA: ACM Press/Addison-Wesley Publishing, Jan. 1998, pp. 480–487, doi: 10.1145/274644.274709.

[53] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005, doi: 10.1109/TITS.2005.848368.

[54] E. Zahmat Doost and W. Zhang, "Mental workload variations during different cognitive office tasks with social media interruptions," *Ergonomics*, vol. 66, no. 5, pp. 592–608, May 2023, doi: 10.1080/00140139.2022.2104381.

[55] J. Rasmussen, "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models," *IEEE Trans. Syst. Man, Cybern.*, vol. SMC-13, no. 3, pp. 257–266, May 1983, doi: 10.1109/TSMC.1983.6313160.

[56] R. W. Swezey and R. E. Llaneras, *Models in Training and Instruction and Instruction*, 2nd ed. Hoboken, NJ, USA: Wiley, 1997.

[57] T. B. Sheridan, "Human supervisory control," in *Handbook Human Factors Ergonomics*, 2nd ed., G. Salvendy, Ed. Hoboken, NJ, USA: Wiley, 1997, pp. 990–1013.

[58] R. B. Haynes, "Forming research questions," *J. Clin. Epidemiol.*, vol. 59, no. 9, pp. 881–886, Sep. 2006, doi: 10.1016/j.jclinepi.2006.06.006.

[59] S. B. Hulley, S. R. Cummings, W. S. Browner, D. G. Grady, and T. B. Newman, *Design Clinical Research*, 3rd ed. Philadelphia, PA, USA: Lippincott Williams and Wilkins, 2007.

[60] C. G. Fisher and K. B. Wood, "Introduction to and techniques of evidence-based medicine," *Spine*, vol. 32, pp. S66–S72, Sep. 2007, doi: 10.1097/brs.0b013e318145308d.

[61] B. Freedman, "Equipoise and the ethics of clinical research," *New England J. Med.*, vol. 317, no. 3, pp. 141–145, Jul. 1987, doi: 10.1056/nejm198707163170304.

[62] J. M. Bland and D. G. Bland, "Statistics notes: One and two sided tests of significance," *Brit. Med. J.*, vol. 309, no. 6949, p. 248, Jul. 1994, doi: 10.1136/bmj.309.6949.248.

[63] H. A. Aziz, "Comparison between field research and controlled laboratory research," *Arch. Clin. Biomed. Res.*, vol. 1, no. 2, pp. 101–104, Apr. 2017, doi: 10.26502/acbr.50170011.

[64] A. Liberati, D. G. Altman, J. Tetzlaff, C. Mulrow, P. C. Gøtzsche, J. P. A. Ioannidis, M. Clarke, P. J. Devereaux, J. Kleijnen, and D. Moher, "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration," *J. Clin. Epidemiol.*, vol. 62, no. 10, pp. e1–e34, Oct. 2009, doi: 10.1016/j.jclinepi.2009.06.006.

[65] B. Zeigarnik, "On finished and unfinished tasks," in *Psychol Forsch*. London, U.K.: Kegan Paul, Trench, Trubner & Company, 1927, pp. 1–85, doi: 10.1037/11496-025.

[66] M. B. Edwards and S. D. Gronlund, "Task interruption and its effects on memory," *Memory*, vol. 6, no. 6, pp. 665–687, Nov. 1998, doi: 10.1080/741943375.

[67] *National Transportation Safety Board Aviation Accident Final Report*, Nat. Transp. Saf. Board (NTSB), New Orleans, LA, USA, May 1988.

[68] G. Mark, *Multitasking in the Digital Age*, vol. 8. San Rafael, CA, USA: Morgan & Claypool, 2015.

[69] R. Gatzounis, M. G. S. Schrooten, G. Crombez, and J. W. S. Vlaeyen, "Interrupted by pain: An anatomy of pain-contingent activity interruptions," *Pain*, vol. 155, no. 7, pp. 1192–1195, 2014, doi: 10.1016/j.pain.2014.03.017.

[70] P. Gontar, S. A. E. Schneider, C. Schmidt-Moll, C. Bollin, and K. Bengler, "Hate to interrupt you, but…analyzing turn-arounds from a cockpit perspective," *Cognition, Technol. Work*, vol. 19, no. 4, pp. 837–853, Sep. 2017, doi: 10.1007/s10111-017-0440-4.

[71] Y. Chen, W. Fang, B. Guo, and H. Bao, "The moderation effects of task attributes and mental fatigue on post-interruption task performance in a concurrent multitasking environment," *Appl. Ergonom.*, vol. 102, Jul. 2022, Art. no. 103764, doi: 10.1016/j.apergo.2022.103764.

[72] R. M. A. Van der Heiden, J. L. Kenemans, S. F. Donker, and C. P. Janssen, "The effect of cognitive load on auditory susceptibility during automated driving," *Hum Factors*, vol. 64, no. 7, pp. 1195–1209, Nov. 2022, doi: 10.1177/0018720821998850.

[73] Y. Chen, W. Fang, B. Guo, and H. Bao, "Fatigue-related effects in the process of task interruption on working memory," *Front Hum. Neurosci.*, vol. 15, p. 697, Nov. 2021, doi: 10.3389/FNHUM.2021.703422/BIBTEX.

[74] R. van der Kleij, T. Hueting, and J. M. Schraagen, "Change detection support for supervisory controllers of highly automated systems: Effects on performance, mental workload, and recovery of situation awareness following interruptions," *Int. J. Ind. Ergonom.*, vol. 66, pp. 75–84, Jul. 2018, doi: 10.1016/j.ergon.2018.02.010.

[75] C. Pankok, M. Zahabi, W. Zhang, I. Choi, Y.-F. Liao, C. S. Nam, and D. Kaber, "The effects of interruption similarity and complexity on performance in a simulated visual-manual assembly operation," *Appl. Ergonom.*, vol. 59, pp. 94–103, Mar. 2017, doi: 10.1016/j.apergo.2016.08.022.

[76] A. Hernández-Sabaté, J. Yauri, P. Folch, M. À. Piera, and D. Gil, "Recognition of the mental workloads of pilots in the cockpit using EEG signals," *Appl. Sci.*, vol. 12, no. 5, p. 2298, Feb. 2022, doi: 10.3390/app12052298.

[77] J. H. Kim, N. Parameshwara, W. Guo, and K. S. Pasupathy, "The impact of interrupting nurses on mental workload in emergency departments," *Int. J. Hum.–Comput. Interact.*, vol. 35, no. 3, pp. 206–217, Feb. 2019, doi: 10.1080/10447318.2018.1447421.

[78] K. Labonté, S. Tremblay, and F. Vachon, "Forewarning interruptions in dynamic settings: Can prevention bolster recovery?" *J. Experim. Psychol., Appl.*, vol. 25, no. 4, pp. 674–694, Dec. 2019, doi: 10.1037/xap0000219.

[79] S. R. Walter, M. Z. Raban, W. T. M. Dunsmuir, H. E. Douglas, and J. I. Westbrook, "Emergency doctors' strategies to manage competing workload demands in an interruptive environment: An observational workflow time study," *Appl. Ergonom.*, vol. 58, pp. 454–460, Jan. 2017, doi: 10.1016/j.apergo.2016.07.020.

[80] A. Lewandowska, I. Rejer, K. Bortko, and J. Jankowski, "Eye-tracker study of influence of affective disruptive content on User's visual attention and emotional state," *Sensors*, vol. 22, no. 2, p. 547, Jan. 2022, doi: 10.3390/s22020547.

[81] S. A. Powers and M. W. Scerbo, "Examining the effect of interruptions at different breakpoints and frequencies within a task," *Hum. Factors: J. Hum. Factors Ergonom. Soc.*, vol. 65, no. 1, pp. 22–36, Apr. 2021, doi: 10.1177/00187208211009010.

[82] M. Wu, Q. Gao, and Y. Liu, "Exploring the effects of interruptions in different phases of complex decision-making tasks," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 65, no. 3, pp. 450–481, Jun. 2021, doi: 10.1177/00187208211018882.

[83] J. Choi and M. Park, "System dialog message and its effect on task switching by age," *Int. J. Human–Computer Interact.*, vol. 36, no. 17, pp. 1646–1657, Jul. 2020, doi: 10.1080/10447318.2020.1785152.

[84] I. Katidioti, J. P. Borst, D. J. Bierens de Haan, T. Pepping, M. K. van Vugt, and N. A. Taatgen, "Interrupted by your pupil: An interruption management system based on pupil dilation," *Int. J. Hum.–Comput. Interact.*, vol. 32, no. 10, pp. 791–801, Oct. 2016, doi: 10.1080/10447318.2016.1198525.

[85] A. Dahiya, Y. Cai, O. Schneider, and S. L. Smith, "On the impact of interruptions during multi-robot supervision tasks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2023, pp. 9771–9777, doi: 10.1109/ICRA48891.2023.10160323.

[86] S. Wang, Y. Liu, S. Li, Z. Liu, X. You, and Y. Li, "The effect of two-stage warning system on human performance along with different takeover strategies," *Int. J. Ind. Ergonom.*, vol. 97, Sep. 2023, Art. no. 103492, doi: 10.1016/j.ergon.2023.103492.

[87] D. Kanaan and N. M. Moacdieh, "Eye tracking to evaluate the effects of interruptions and workload in a complex task," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 64, no. 7, pp. 1168–1180, Feb. 2021, doi: 10.1177/0018720821990487.

[88] S. Midha, H. A. Maior, M. L. Wilson, and S. Sharples, "Measuring mental workload variations in office work tasks using fNIRS," *Int. J. Hum.-Comput. Stud.*, vol. 147, Mar. 2021, Art. no. 102580, doi: 10.1016/j.ijhcs.2020.102580.

[89] J. I. Westbrook, M. Z. Raban, S. R. Walter, and H. Douglas, "Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: A prospective, direct observation study," *BMJ Quality Saf.*, vol. 27, no. 8, pp. 655–663, Aug. 2018, doi: 10.1136/bmjqs-2017-007333.

[90] K. R. Campoe and K. K. Giuliano, "Impact of frequent interruption on Nurses' patient-controlled analgesia programming performance," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 59, no. 8, pp. 1204–1213, Sep. 2017, doi: 10.1177/0018720817732605.

[91] E. M. Zeißig, A. Janß, J. Dell'Anna-Pudlik, M. Ziefle, and K. Radermacher, "Development and experimental evaluation of an alarm concept for an integrated surgical workstation," *Biomedizinische Technik*, vol. 61, no. 2, pp. 199–209, Apr. 2016, doi: 10.1515/BMT-2015-0235.

[92] R. Rajan, T. Selker, and I. Lane, "Effects of mediating notifications based on task load," in *Proc. 8th Int. Conf. Automot. User Interface Interact. Veh. Appl.*, Oct. 2016, pp. 145–152, doi: 10.1145/3003715.3005413.

[93] M. Özdemir, K. Dönmez, and S. Demirel, "Determining the key factors affecting pseudo-pilot workload based on real-time simulations," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2677, no. 7, pp. 718–732, May 2023, doi: 10.1177/03611981231169529.

[94] E. C. Falkland, M. W. Wiggins, H. Douglas, D. Sturman, J. C. Auton, L. Shieh, and J. I. Westbrook, "Explaining emergency physicians' capacity to recover from interruptions," *Appl. Ergonom.*, vol. 105, Nov. 2022, Art. no. 103857, doi: 10.1016/j.apergo.2022.103857.

[95] E. C. Falkland, M. W. Wiggins, and J. I. Westbrook, "Cue utilization differentiates performance in the management of interruptions," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 62, no. 5, pp. 751–769, Aug. 2020, doi: 10.1177/0018720819855281.

[96] N. R. Lodinger and P. R. DeLucia, "Angle of camera view influences resumption lag in a visual-motor task," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 61, no. 5, pp. 793–804, Aug. 2019, doi: 10.1177/0018720818794526.

[97] R. Andreasson, J. Lindblom, and P. Thorvald, "Interruptions in the wild: Portraying the handling of interruptions in manufacturing from a distributed cognition lens," *Cognition, Technol. Work*, vol. 19, no. 1, pp. 85–108, Feb. 2017, doi: 10.1007/S10111-016-0399-6/FIGURES/5.

[98] A. Borowsky, W. J. Horrey, Y. Liang, A. Garabet, L. Simmons, and D. L. Fisher, "The effects of brief visual interruption tasks on drivers' ability to resume their visual search for a pre-cued hazard," *Accident Anal. Prevention*, vol. 93, pp. 207–216, Aug. 2016, doi: 10.1016/j.aap.2016.04.028.

[99] N. Koundal, N. Bencheva, M. N. M. Saad, I. Elamvazuthi, M. S. Al-Quraishi, and A. Abdalhadi, "Investigating the effect of nested interruptions and levels of cognitive demand on mental states," in *Proc. 32th Int. Sci. Conf. Electron. (ET)*, Sep. 2023, pp. 1–6, doi: 10.1109/et59121.2023.10279605.

[100] Y.-Y. Chen, W.-N. Fang, H.-F. Bao, and B.-Y. Guo, "The effect of task interruption on working memory performance," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 66, no. 4, pp. 1132–1151, Nov. 2022, doi: 10.1177/00187208221139017.

[101] B. Cheng, C. Fan, H. Fu, J. Huang, H. Chen, and X. Luo, "Measuring and computing cognitive statuses of construction workers based on electroencephalogram: A critical review," *IEEE Trans. Computat. Social Syst.*, vol. 9, no. 6, pp. 1644–1659, Dec. 2022, doi: 10.1109/TCSS.2022.3158585.

[102] H. Fu, Z. Xia, Y. Tan, and X. Guo, "Influence of cues on the safety hazard recognition of construction workers during safety training: Evidence from an eye-tracking experiment," *J. Civil Eng. Educ.*, vol. 150, no. 1, Aug. 2023, Art. no. 04023009, doi: 10.1061/JCEECD.EIENG-1882.

[103] W. C. Borman and S. M. Motowidlo, "Expanding the criterion domain to include elements of contextual performance," in *Personnel Selection in Organizations*, W. C. Borman and N. Schmitt, Eds. San Francisco, CA, USA: Jossey-Bass: Psychology Faculty Publications, 1993, pp. 71–98. Accessed: Dec. 15, 2022. [Online]. Available: https://digitalcommons.usf.edu/psy_facpub/1111

[104] E. M. Altmann, J. G. Trafton, and D. Z. Hambrick, "Momentary interruptions can derail the train of thought," *J. Experim. Psychol., Gen.*, vol. 143, no. 1, pp. 215–226, Feb. 2014, doi: 10.1037/a0030986.

[105] M. Cabanac, "What is emotion?" *Behavioural Processes*, vol. 60, no. 2, pp. 69–83, Nov. 2002, doi: 10.1016/s0376-6357(02)00078-5.

[106] S. Gannouni, A. Aledaily, K. Belwafi, and H. Aboalsamh, "Adaptive emotion detection using the valence-arousal-dominance model and EEG brain rhythmic activity changes in relevant brain lobes," *IEEE Access*, vol. 8, pp. 67444–67455, 2020, doi: 10.1109/ACCESS.2020.2986504.

[107] A. Ishii, M. Tanaka, and Y. Watanabe, "Neural mechanisms of mental fatigue," *Rev. Neurosci.*, vol. 25, no. 4, pp. 469–479, Aug. 2014, doi: 10.1515/REVNEURO-2014-0028.

[108] I. J. Deary, "Measuring stress: A guide for health and social scientists," *J. Psychosomatic Res.*, vol. 41, no. 2, p. 186, Aug. 1996, doi: 10.1016/0022-3999(96)00066-9.

[109] A. Abdalhadi, N. Bencheva, N. M. Saad, M. S. Al-Quraishi, and N. Koundal, "Study the effect of acute stress on decision making using function near infrared spectroscopy (fNIRS)," in *Proc. Asia Simulation Conf.*, 2024, pp. 452–463, doi: 10.1007/978-981-99-7243-2_37.

[110] S. B. Hulley, S. R. Cumming, W. S. Browner, D. S. Grady, and T. B. Newman, *Designing Clinical Research*, 4th ed. Philadelphia, PA, USA: Lippincott Williams and Wilkins (LWW), 2007.

[111] M. Teplan, "Fundamentals of EEG measurement," *Meas. Sci. Technol.*, vol. 2, no. 2, pp. 1–11, 2002.

[112] M. Brysbaert, "How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables," *J. Cogn.* , vol. 2, no. 1, Jul. 2019, Art. no. PMC6640316, doi: 10.5334/JOC.72.

[113] C. Rusnock, B. Borghetti, and I. McQuaid, "Objective-analytical measures of workload—The third pillar of workload triangulation?" in *Foundations of Augmented Cognition*, vol. 9183, D. D. Schmorrow and C. M. Fidopiastis, Eds. Cham, Switzerland: Springer, 2015, pp. 124–135, doi: 10.1007/978-3-319-20816-9_13.

[114] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—State-of-the-art and guidelines," *J. Neural Eng.*, vol. 12, no. 3, Apr. 2015, Art. no. 031001, doi: 10.1088/1741-2560/12/3/031001.

[115] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—Practices and possibilities," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 488–500, May 2012, doi: 10.1109/TITB.2012.2188536.

[116] J. Chen, X. Song, and Z. Lin, "Revealing the 'invisible gorilla' in construction: Estimating construction safety through mental workload assessment," *Autom. Construct.*, vol. 63, pp. 173–183, Mar. 2016, doi: 10.1016/j.autcon.2015.12.018.

[117] S. A. Bini, "Artificial intelligence, machine learning, deep learning, and cognitive computing: What do these terms mean and how will they impact health care?" *J. Arthroplasty*, vol. 33, no. 8, pp. 2358–2361, Aug. 2018, doi: 10.1016/j.arth.2018.02.067.

[118] Y. Peng, Y. Lin, C. Fan, Q. Xu, D. Xu, S. Yi, H. Zhang, and K. Wang, "Passenger overall comfort in high-speed railway environments based on EEG: Assessment and degradation mechanism," *Building Environ.*, vol. 210, Feb. 2022, Art. no. 108711, doi: 10.1016/j.buildenv.2021.108711.

[119] S. Zhang, W. Ren, X. Tan, Z.-J. Wang, Y. Liu, J. Zhang, X. Zhang, and X. Cao, "Semantic-aware dehazing network with adaptive feature fusion," *IEEE Trans. Cybern.*, vol. 53, no. 1, pp. 454–467, Jan. 2023, doi: 10.1109/TCYB.2021.3124231.

[120] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[121] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Jul. 2018, Art. no. 056013, doi: 10.1088/1741-2552/aace8c.

[122] Y. Hou, L. Zhou, S. Jia, and X. Lun, "A novel approach of decoding EEG four-class motor imagery tasks via scout ESI and CNN," *J. Neural Eng.*, vol. 17, no. 1, Feb. 2020, Art. no. 016048, doi: 10.1088/1741-2552/ab4af6.

[123] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "LSTM-based EEG classification in motor imagery tasks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2086–2095, Nov. 2018, doi: 10.1109/TNSRE.2018.2876129.

[124] S. Alhagry, A. Aly, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, pp. 1–4, Nov. 2017, doi: 10.14569/ijacsa.2017.081046.

[125] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1290–1301, Jul. 2022, doi: 10.1109/TAFFC.2020.2994159.

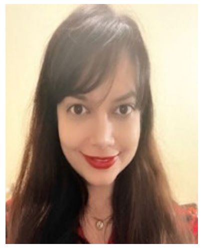
[126] Z. Wang, Y. Tong, and X. Heng, "Phase-locking value based graph convolutional neural networks for emotion recognition," *IEEE Access*, vol. 7, pp. 93711–93722, 2019, doi: 10.1109/ACCESS.2019.2927768.

NITIN KOUNDAL received the B.Tech. degree from Punjab Technical University (PTU) and the M.E. degree in mechanical engineering from NITTTR Chandigarh (Panjab University), India, in 2019. He is currently pursuing the Ph.D. degree with the Centre for Intelligent Signal and Imaging Research, Universiti Teknologi PETRONAS (UTP), Malaysia. His research interests include biomedical signal processing, rehabilitation device design, finite element analysis, and additive manufacturing.

ABDUALRHMAN ABDALHADI received the B.S. degree from Asia Pacific University of Technology and Innovation (APU), Malaysia, and the M.S. degree in mechatronic and automatic control from Universiti Teknologi Malaysia, Malaysia, in 2021. He is currently pursuing the Ph.D. degree in brain physiological signal processing, virtual reality and user experience with the Centre for Intelligent Signal and Imaging Research, Universiti Teknologi PETRONAS, Malaysia. His research interests include virtual reality, neuro-signal processing, control systems, and automatic control.

MAGED S. AL-QURAISHI (Member, IEEE) received the B.Sc. degree in biomedical engineering from Baghdad University, Iraq, in 2005, the M.Sc. degree in biomedical engineering from Universiti Putra Malaysia, in 2015, and the Ph.D. degree from Universiti Teknologi PETRONAS, Malaysia, in 2021. He is currently a Postdoctoral Fellow with the Smart Mobility and Logistics Center, King Fahd University of Petroleum and Minerals. He has published several articles in refereed journals, such as IEEE, Springer, and MDPI. His research interests include biomedical signal processing, neuroengineering, machine learning, deep learning, instrumentation, and rehabilitation robotics.

IRRAIVAN ELAMVAZUTHI (Senior Member, IEEE) received the Ph.D. degree from the Department of Automatic Control and Systems Engineering, The University of Sheffield, U.K., in 2002. He is currently an Associate Professor with the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS (UTP), Malaysia. His research interests include control, robotics, mechatronics, power systems, and biomedical applications. He is the Chair of the IEEE Robotics and Automation Society (Malaysia Chapter).



MAHDIYEH SADAT MOOSAVI received the Ph.D. degree from the National Polytechnical Institute of Grenoble (INPG) in 1996. She has been a Professor with the Arts et Métiers Institute of Technology, LISPEN, HESAM Université, Chalon-sur-Saône, France, since 2004. She has authored many scientific papers in virtual reality and related disciplines. She is involved in different projects with industrial partners and initiated international collaborative projects in the area of virtual reality with universities in the USA, Australia, Colombia, and Malaysia. Her research interests are focused on virtual immersion linked with engineering, cultural heritage, and health applications.

CHRISTOPHE GUILLET is currently an Associate Professor with Université Bourgogne Franche-Comté. He is a member of the Methods and Uses Cases of Virtual Reality Team, University of Burgundy, LISPEN, Chalon-sur-Saône, France. His main research interests include the development of mathematical methods for multidimensional data analysis recorded in virtual environments for human movement, brain activity, and cognitive psychology studies and involving the combination of probabilistic, topological, machine learning, and neural network tools.

FRÉDÉRIC MERIENNE (Member, IEEE) received the Ph.D. degree from the National Polytechnic Institute of Grenoble (INPG), in 1996. He has been a Professor with the Arts et Métiers Institute of Technology, LISPEN, HESAM Université, Chalon-sur-Saône, France, since 2004. He has authored many scientific articles in virtual reality and related disciplines. His research interests include virtual immersion linked with engineering, cultural heritage, and health applications. He is involved in different projects with industrial partners and initiated international collaborative projects in the area of virtual reality with universities in USA, Australia, Colombia, and Malaysia.

NAUFAL M. SAAD (Member, IEEE) received the master's degree from École Nationale Supérieure d'Ingénieurs de Limoges, France, and the Ph.D. degree in telecommunication from Université de Limoges, France, in 2005. He is currently an Associate Professor and the Chair of the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS (UTP), Malaysia, and a Research Member with the Centre for Intelligent Signal and Imaging Research, UTP. His research interests include neuro-signal processing, medical imaging, and communication.