2023

WORKING MEMORY PERFORMANCE: IS SUBJECTIVE MEASUREMENT A BETTER PREDICTOR THAN COGNITIVE LOAD?

Megan McCray

WORKING MEMORY PERFORMANCE: IS SUBJECTIVE MEASUREMENT A BETTER PREDICTOR THAN COGNITIVE LOAD?

By

Megan M. McCray

B.A., Marshall University, 2019

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Master of Arts

Department of Psychology

Human Factors Program
In the Graduate School
The University of South Dakota
May, 2023

The members of the Committee appointed to examine
the Thesis_____ of Megan M. McCray
find it satisfactory and recommend that it be accepted.

DocuSigned by:
Timothy Ricker
AE99786091264D2...
Chairperson

DocuSigned by:
B583D32E63AB431...

DocuSigned by:
9B159DA81980438...

DocuSigned by:
Lu Baugh
05454119D36C468...

ABSTRACT

We rely on our capacity for rapid attention switching to conduct multiple tasks simultaneously. Leading working memory models assume that memory maintenance and attention-demanding secondary task processing cannot coincide. Any reduction in memory maintenance activities occurring due to secondary task processing leads to impaired recall. This temporal relationship is typically characterized through the proportion of time spent attending to the concurrent processing task, also called cognitive load. Although the primary determinant of forgetting in leading models, recent findings show limitations to cognitive load effects in multitasking. We investigated whether the effects of cognitive load are a byproduct of subjective task difficulty assessments by participants during a visuospatial working-memory dual-task by asking participants to complete subjective workload measurement (NASA-TLX). Results were compared to objective cognitive load to determine which measurement is a better model for predicting multitasking effects. The present findings inform our understanding of human working memory capabilities and the role of both subjective workload and objective cognitive load in driving memory performance during multitasking.

Thesis/Dissertation Advisor DocuSigned by: Timothy Ricker
BCEDBC8246A44E1...
Dr. Timothy Ricker

Table of Contents

List of Tables

List of Figures

Introduction

Working memory is a crucial component in understanding human information processing. Imagine not being able to do something as simple as follow a recipe, compute mental math, or even remember the first half of this sentence. These tasks are among millions of others that are trivial to complete but still require working memory. Working memory is a limited capacity system that temporarily stores and maintains information in an accessible state for cognitive processing (Baddeley, 1986; Baddeley & Hitch, 1974; Cowan, 1988, 1995; Oberauer, 2002). Working memory performance predicts aptitude in many higher-order tasks (Barrouillet & Lecas, 1999; Daneman & Carpenter 1980; Kyllonen & Crystal, 1990; Süß et al., 2002; Engle et al., 1999), making it central to human cognition.

The term multitasking is used among the general population, often listed within job descriptions as a desirable skill despite common misunderstandings about how multitasking functions. The predominant assumption is that multitasking reflects concurrent attention-based processing of multiple tasks, but this is incorrect. One can only hold a single task in the focus of attention at any given time (Pashler, 1994). Instead of processing items from different tasks simultaneously, individuals engage in the rapid switching of attention between tasks, using brief instances of free time to maintain individual memory items. Strategies must be employed to engage in multiple tasks requiring working memory concurrently. One can take three general approaches to multitasking, balancing both tasks (equal dedication of resources), prioritizing one (unequal distribution of resources), or executing the tasks serially (completing one task before allocating any resources to the secondary task; Hart & Wickens, 1990).

Current theories argue that memory performance decreases when one shifts attention away from memory-related processing (Barrouillet et al., 2007; Portrat et al., 2008; Lemaire &

Portrat, 2018; Lewandowsky et al., 2004; Lewandowsky et al., 2009; Oberauer et al., 2012; Portrat & Lemaire, 2015). One performs attention-based maintenance during short pauses between concurrent processing tasks to prevent forgetting. This brief allocation of attention strengthens memory traces for the duration that the item is in the focus of attention. If there is insufficient time remaining for maintenance following an attention shift to a concurrent processing task then memory performance suffers.

Two main models characterize how splitting attention leads to forgetting: Time-Based Resource-Sharing (TBRS) Model (Barrouillet et al., 2004) and Serial Order in a Box – Complex Span (SOB-CS) Model (Oberauer et al., 2012). Existing literature shows both support and criticism for both models. The TBRS and SOB-CS models both address the concept that memory traces are maintained within working memory through the allocation of attention. In both models, individuals are more likely to successfully recall all the items when presented with few items versus many memory items. However, the specific details defining the mechanisms of forgetting and maintenance vary across models. The following section details both models along with the main historical developments that lead to their creation.

Decay Theories and the Time-Based Resource-Sharing (TBRS) Model

Debate surrounding forgetting from immediate memory can be traced back to Brown (1958) who advanced the understanding of forgetting by providing empirical evidence for short-term forgetting and establishing rules governing time-based decay. He manipulated time and distractors in dual task experiments to study the impacts on recall performance, ultimately explaining his findings through the idea that the memory trace naturally decays over time. Brown observed heightened levels of forgetting when increasing the number of memory items, introducing a secondary task, and delaying memory recall by seconds. Observations held true in single- and dual-task trials, even when the number of memory items was well within working memory span. Performance was not impacted by stimuli that shared high similarity with the memory items. This was in conflict with current thought that recall would only be impaired in conditions in which distractors had a high similarity factor when compared to the memory stimuli, as it was suspected that similarity caused confusion at recall. Brown also revealed that presentation position of distractor items had a differential effect on recall performance. Distractors presented before the memory items had a lesser effect on recall than those presented immediately after the memory item. Brown's trace decay hypothesis explained these results by stating that only distractors presented after the memory item directly impacted the time one had available to conduct necessary maintenance activities to counteract decay. Increased time between memory item and distractor reduced the effect of decay at test due to short-term learning effects that transferred items into long-term memory. Brown provided a strong foundation for studying time-based decay as a forgetting mechanism by manipulating memory items, distractors, and retention periods to determine the impact on recall performance.

Baddeley and Hitch (1974) brought further support for decay theory, introducing the first version of Baddeley (1986)'s Multi-Component Model of working memory. Baddeley and Hitch demonstrated impaired working memory recall due to concurrent task processing. Performance in single-tasks was found to always be better than performance in dual-tasks. Further investigation of dual-tasks showed that when primary and secondary tasks had high-similarity recall was worse than when tasks had low-similarity. Interference increased when increasing the number of tasks. Increased task similarity also increased interference between tasks, negatively impacting recall. To account for these findings they proposed a limited-capacity system in which control and storage processes share the same pool of limited resources (i.e., attention). The three model components responsible for control and storage in the model were the phonological store, visuospatial sketchpad, and central executive. The phonological store temporarily holds auditory information and supports the rehearsal of verbal information, maintaining it in working memory via a process termed the phonological loop. The visuospatial sketchpad is a temporary storage and manipulation area for visual and spatial information. Lastly, the central executive controls attention allocation and coordinates information processing, interfacing with perceptual input, other working memory components, and long-term memory. The components of the model function to counteract passive trace decay that is thought to occur with time. Memory performance is dependent on trace activation being above the threshold of forgetting at time of recall.

Interest in further understanding working memory performance while dual-tasking led to the development of the popular complex span paradigm (Daneman & Carpenter, 1980). In these tasks memory items are presented one at a time and must be remembered for immediate recall at the end of the sequence. A secondary processing task is presented after each individual memory item. During these processing tasks one must make a simple judgment and response, such as verifying the accuracy of an addition problem or judging whether a shape is symmetrical. The memory items are recalled after all of the memory-presentation and processing task pairings are complete. The design evokes resource sharing by requiring memory item maintenance to occur while processing an additional task. Within this basic structure many variants have become popular including variations such as the reading, counting, and operation span tasks (Daneman & Carpenter, 1980; Case, Kurland, & Goldberg, 1982; Case, 1985; Turner & Engle, 1989).

Resources can be shared across the memory and processing tasks in two ways, (1) resource-sharing, allocating a portion of the resources to each task and complete all tasks simultaneously, or (2) splitting the resources temporally, first devoting all resources to one task then switching all resources to the next task. Towse and Hitch explored this question of how resources are split (Towse & Hitch, 1995; Towse et al., 1998; 2000) ultimately proposing a task-switching model where one does not simultaneously engage in processing and storage but in resource-switching between separate pools of resources for processing and storage. This model was directly in contrast to resource-sharing theories that suggested dependence between processing and storage resources. By manipulating task difficulty and time in a counting span dual-task Towse and Hitch (1995) showed that working memory recall suffered as the duration of a concurrent task that occurred between memory item presentations increased, independent of concurrent task difficulty. These findings were replicated and extended to operation and reading span tasks in both children and adults (Towse et al., 1998; 2000). The task-switching model explained that secondary task execution required resource-switching from maintenance to processing. This switch was stated to halt memory item maintenance for the duration of the retention interval (RI). Increasing the length of the RI by increasing the length of time to complete the concurrent task meant increasing the time when maintenance could not occur. Increased RIs meant increased time for memory decay, leading to increased forgetting.

Barrouillet and Camos (2001) addressed limitations to Towse and Hitch's (1995) task-switching model by manipulating difficulty while controlling other parameters such as stimulus delay, stimulus presentation duration, and retention interval duration. They demonstrated that secondary task difficulty plays an essential role in memory performance. This finding showed that RI duration alone could not consistently account for forgetting. Barrouillet et al. (2004) replicated these findings in adults and proposed the Time-Based Resource-Sharing (TBRS) model to address oversimplifications of the task-switching model by accounting for task difficulty and a shared-resource system.

The TBRS model assumes one can attend to only one task at a time, and instances of shifting attention away from memory maintenance lead to rapid decay of memory activation. Forgetting occurs once memory trace activation drops below a threshold value (Barrouillet et al., 2004; Barrouillet et al., 2007; Portrat et al., 2008). The TBRS model explains multiple-item maintenance using reactivation strategies, refreshing or rehearsing, and rapidly cycling through memory items to counteract decay. Rehearsal is the vocal or subvocal repetition of memory items resulting in the increased activation of rehearsed memory traces. Refreshing is the covert allocation of attention to a memory item for a brief instance to replenish memory trace activation. Refreshing was proposed by Barrouillet et al. (2004) to account for memory maintenance when rehearsal is unlikely. Demanding concurrent processing experienced in multitasking reduces the free time to engage in reactivation strategies used to counteract decay. Forgetting occurs as trace activations fall below threshold due to unchecked decay (Barrouillet et al., 2004; Barrouillet et al., 2007; Barrouillet & Camos, 2007).

Barrouillet et al. (2004) tested the TBRS model proposal by developing the continuous operation span task and systematically manipulating secondary task parameters. The continuous operation span task replaced operation span tasks with simple tasks that harness low individual variability, such as adding a digit to previously presented number and speaking the solution aloud or reading a digit on the screen to ensure controlled retention interval duration. The control gained from continuous operation span tasks allows for more direct manipulation of secondary task difficulty while controlling RI duration. Within the setting of a continuous complex span task, the task-switching model of Towse and Hitch (Towse & Hitch, 1995; Towse et al., 1998) predicts that increasing the RI duration will negatively impact performance and that the type of secondary task is irrelevant. In contrast the TBRS model predicts that the duration of the RI is unimportant and that the amount of free time available during the RI predicts memory performance. Results (Barrouillet et al., 2004) directly conflicted with the task-switching model (Towse & Hitch, 1995; Towse et al., 1998) by confirming the effect of secondary task type on recall performance due to changes in the proportion of free time. Manipulation of RI duration did not impact recall performance when cognitive cost (i.e., task type and pace) was held constant. These findings demonstrated that forgetting is not simply due to the passage of time, as suggested by a pure memory-decay hypothesis like the task-switching model. What is key to describing forgetting is the proportion of available free time during memory retention which influences the proportion of free time available for reactivation. The TBRS model reevaluated the role of time and proposed that the proportion of occupied to total time, termed cognitive load, predicts forgetting.

The Time-Based Resource-Sharing (TBRS) model comprehensively explains multitasking effects on working memory performance. Cognitive load captures the constraints of shared resources (i.e., a tradeoff between processing and storage), passive time-based decay, and free time used for reactivation to determine forgetting induced by a concurrent processing task. The TBRS model builds upon the resource-sharing hypothesis to better explain the relationship between processing and storage by predicting a linear relationship between recall accuracy and cognitive load. Task difficulty is assumed to modify the slope of the linear relationship between cognitive load and memory performance. This addition predicts performance in many, if not all, tasks even when the secondary task is removed (i.e., no-load condition, the intercept; Barrouillet et al., 2004).

Ample evidence supporting the TBRS model has accumulated, showing the importance of time in forgetting (Lépine et al., 2005; Barrouillet et al., 2004; Barrouillet, et al., 2007; Portrat, et al., 2008) and arguing for the trace reactivation explanation of cognitive load. Lépine, Bernardin, and Barrouillet (2005) concluded that even very simple secondary tasks could impair recall performance if presented at a fast pace. Recall performance decreased as the time during which attention engaged in processing increased (Portat et al., 2008), supporting the hypothesis that a primary determinant of forgetting is the time available for refreshing. An additional investigation varied the total retention interval while controlling for free time, and thus reactivation.

Barrouillet et al. (2007) extended the findings of decreased recall under conditions with a simple secondary task (Lépine et al., 2005) to reading span and response selection tasks, finding that memory retrieval and response selection have similar effects on concurrent maintenance. An additional finding suggests that various processes can have comparable effects on recall if they demand similar levels of executive processing. Tasks that do not require significant amounts of executive processing may have minimal impact on concurrent activities, resulting in no or minimal impairment to recall. The TBRS model accounts for this finding by including RI duration and task difficulty in the cognitive load formula.

The TBRS model evolved through the development of increasingly controlled experiments and continues to explain a growing body of research. While this model offers a valuable framework for studying the impact of multitasking on working memory and can help researchers better understand the complex interplay between cognitive processes, limitations have been found surrounding the TBRS model (Ricker & Vergauwe, 2020, 2022) and the role of time in the TBRS model (Oberauer et al., 2012; Oberauer & Lewandowsky, 2011, 2013) extensive enough to warrant another competing set of models to exist.

Interference Theories and the Serial Order in a Box – Complex Span Model

Case et al. (1982) demonstrated that the increased memory span while multitasking observed as one ages into adulthood is not due to increase in resources or processing space. Instead one becomes more efficient at tasks requiring fewer resources for processing and leaving more available for storage. Case et al. found a linear relationship between speed of repeating words and recall performance. Span was a direct function of speed which varied by developmental group and recall performance could be made equal across age groups by controlling repetition speed through replacing familiar stimuli with unfamiliar stimuli. Case et al. argued that these results support interference-based models, a group of models that propose forgetting occurs because of the overlap of stimulus representations. This means that processing multiple items in a brief period of time can result in interference with one another at the representation level (Hudjetz & Oberauer, 2007; Oberauer & Kliegl, 2001; Oberauer et al., 2004). Unlike temporal models, which focus on the role of time in causing forgetting, interference models concentrate on the number and content of processing items (Saito & Miyake, 2002; Oberauer et al., 2012).

Nairne (1990) proposed the highly influential Feature Model stating that memory traces may change due to interference from outside events and cognitive activities. This computational model represents memory traces as lists of features. Outside events can disrupt memory representation, causing interference among contiguous traces in primary memory, by overwriting feature bindings of preexisting memory traces. Model fits correctly capture a number of benchmark findings including serial position curves, similarity effects, and articulatory suppression effects. For example, simulation accurately predicted similarity effects by capturing decreased performance under conditions with high similarity memory items.

Saito & Miyake (2002) linked the idea of interference with the task-switching model of Towse and Hitch (1995; Towse et al.1998) by analyzing sentence order and reading time effects. Differences in memory recall were observed when holding RI constant and varying the amount of processing required, however, the effect disappeared when varying the RI while holding constant the amount of sentence processing (Towse et al. 1998, 2000). These results undermine time-based decay by suggesting that the amount of processing activities is what impacts sentence order effect. Decreased performance was observed with increased memory items, supporting a shared pool of resources used for processing and storage. The findings suggested task difficulty (or the amount of processing needed) influenced task performance, not just retention time. Before the proposal of TBRS, the conclusion supported modifications to the task-switching model (Towse & Hitch 1995; Towse 1998) in which time-based forgetting was replaced with interference-based forgetting and a degree of dependency between processing and storage was included.

At a similar time, Farrell and Lewandowsky (2002) introduced a new memory model for serial order called Serial Order in a Box (SOB) which emphasized the role of encoding and retrieval processes in determining forgetting. Encoding refers to the process of taking in sensory information and converting it into a meaningful representation that can be stored in the brain. Retrieval is the process of accessing and recalling information that has been previously stored in memory. In SOB, retrieval involves searching for information in long-term memory and bringing it into working memory for use in the current moment. Farrell and Lewandowsky used five simulations to demonstrate SOB's ability to explain several benchmark findings including the shape of the serial position curve and error patterns. This model is able to account for many response patterns, however, several limitations were found including the inability to account for probed recall performance and backward recall. This was later updated to the C-SOB (Contextual-Serial Order in a Box; Farrell, 2006) model to account for findings of similarity-sensitive encoding and the existence of the mixed-list phonological similarity advantage in serial recall.

Oberauer and Lewandowsky (2008) tested time-based decay and interference models in a serial recall task by introducing articulatory suppression during delays between item presentations and delays between item recalls. Extending the delay before retrieval left memory unchanged regardless of whether participants were engaging in a simple response task or a more difficult choice response task throughout the retention interval. Evidence supported that a single distractor event can cause much forgetting, but additional distractor events did not cause much more forgetting. Additionally, the findings suggested that rehearsal may function as a retrieval and re-encoding episode and not by counteracting decay. These results challenged models of decay and rehearsal, suggesting that interference is the primary determinant of recall performance in complex span paradigms.

Lewandowsky & Oberauer (2009) continued to build support for the interference model by conducting a study on how people respond to stimuli after making an error. The study found slower, less accurate, memory responses after errors on secondary task performance. This finding indicated an attentional postponement due to error-related processing. The increase in error rates observed in more difficult conditions was likely a result of more post-error processing reducing the free time for interference removal, leading to the appearance of time-based forgetting. The findings show it is essential to control the timing of processing events and error rates on the processing task when studying the causes of forgetting in paradigms that combine memory maintenance with a concurrent processing task.

Oberauer et al. (2012) proposed a new iteration of the SOB model titled Serial-Order-in-a-Box Complex-Span (SOB-CS) to address the ongoing conflicting evidence between the time-based and interference-based models. This model expanded in two major ways. First, this model stated that encoded distractors create interference. The duration of attentional processing and novelty determine distractor strength. A distractor that receives considerable attention will have a more substantial interference effect. This approach emphasizes that memory performance decreases as the number of representations or interference between representations increases. Second, interference removal occurs during free time by “unbinding” distractor items. One can use free time not to refresh decaying memories but to remove interfering representations. This leads to the prediction of an effect of cognitive load, the proportion of occupied time relative to total retention time, on memory performance similar to that made by the TBRS model (Oberauer et al., 2012). Here free time is crucial because it determines how much interference removal can occur following each episode of interference.

Simulations of this model were benchmarked against several consistent findings, including the relationship between storage and processing, serial positioning curves, similarity effects, and individual differences. For example, the SOB-CS model explained that decreased performance with increased pace was due to both less free time to conduct removal activities and longer total distractor presentation time, which led to a more robust encoding of distractors. The SOB-CS model states that operation durations greater than 500ms (duration needed for encoding) do not further impact interference strength, supported by evidence that manipulation of item duration beyond 500ms plays a lesser role than free time. SOB-CS predicts serial position curve data better than TBRS* (a computational model of TBRS; Oberauer et al., 2012) due to its ability to capture distractor similarity effects with higher rates of interference in conditions with high item similarity. Data detailing item and order errors for simple and complex span tasks, item-distractor similarity effects, and patterns of individual differences, are all uniquely predicted by SOB-CS, generating compelling support for this computational model.

Summary Comparison of Time-Based and Interference-Based Models

While both model types utilize cognitive load to calculate free time, the underlying mechanisms of maintenance and forgetting are different. The TBRS model explains forgetting as decay during occupied time, preventing reactivation. In contrast, the interference-based models suggest the active removal of memory items during free time. Deliberate removal practices cannot occur when free time is restricted and items are misremembered, impairing recall.

Outside of the directly conflicting findings of interference-based and time-based models, researchers have also observed limitations to both models of cognitive load effects (Ricker & Cowan, 2010; Vergauwe et al., 2014). These findings suggest limitations or at least variability in the predictiveness of cognitive load. Cowan & Ricker (2010) employed several different secondary tasks, thereby varying the cognitive load while also varying RI duration. A combination of the TBRS model and passive decay across the entire retention duration was required to account for the observed data. This is in conflict with the predictions of TBRS that the amount of free time determines forgetting, not the length of the retention interval itself. Interference models predict increased interference across retention intervals because the number of interfering events increases. The data supported an effect of retention interval independent of the number of interfering events.

Recent experiments by Ricker & Vergauwe (2020, 2022) demonstrate key boundary conditions for observing cognitive load effects in working memory. Memory for orientation was tested by reproducing the location of a dot around the circumference of a circle. Pace of the secondary task was manipulated to induce varying cognitive load levels. No cognitive load effect was observed. Ricker & Vergauwe (2022) explored these findings by changing the task structure from a Brown-Peterson style task, in which all memory items are presented before a single retention interval filled with a secondary task, to a complex span task, as the majority of cognitive load literature leverages the latter study design during experiments. Two different complex span designs manipulated the amount of free time after stimulus presentation. Results showed a cognitive load effect in the experiment with brief amounts of free time, however the effect was not present when the duration was longer. An additional experiment was conducted using a Brown-Peterson design with a short-consolidation period in which cognitive load effects were present. These findings conflict predictions by both the TBRS and SOB-CS models. The inconsistent presence of cognitive load effects across experiments contradicts the core of the TBRS theory. When comparing performance of Brown-Peterson and complex-span tasks under the short free time condition, performance differed greatly. This counterintuitive finding of better performance in a complex span task compared to a Brown-Peterson task challenges the idea that memory performance decreases gradually as interference increases gradually.

Ricker and Vergauwe (2022) proposed a new approach to understanding cognitive load effects. Rather than tradeoff between forgetting and maintenance mechanisms, the memory enrichment approach suggests decreased memory performance associated with increased cognitive load is due to a subjective assessment of difficulty. When individuals feel a task is too difficult they do not engage in strategic memory enrichment processes. Ricker and Vergauwe described enrichment as a strategic process used during free time to enrich poor quality memory representations for better recall. If this theory is correct, then subjective assessment of task difficulty should predict memory performance better than objective cognitive load. Subjective differences in task difficulty are commonly investigated using workload measures.

Workload

Workload is a multidimensional construct used to describe the associated costs of task performance (Hart & Staveland, 1988; Hart, 2006). Context and individual differences can influence perceived task difficulty and thus workload. The definition of workload encompasses subjectivity. It is up for individual interpretation and varies widely across experience levels, external factors, and tasks. A task executed in one context may have a different perceived workload than the same task performed under different conditions (e.g., environment, stress, instruction, etc.; Hart & Wickens, 1990). Similarly, the same task may be effortful for one person but not for another due to practice.

Workload can be measured using physiological, subjective, and performance-based approaches such as heart rate variability, validated rating scales, and task performance accuracy (Hartman & McKenzie, 1979). Subjective ratings are the most common workload measurement (Hart & Staveland, 1988; Hart & Wickens, 1990). Subjective measures probe the individual using indexes and scales to gain subjective insight into their perceived workload. Researchers prefer to utilize subjective indexes instead of physiological measurements, especially in applied settings, because of their quick implementation, associated low cost, sufficient workload predictability, and the ability to withstand changes in numerous environmental variables.

Despite the interrater variability subjective workload correlates to physiological measures of workload as shown through the validation of several scales (Hart & Staveland, 1988; Hart &

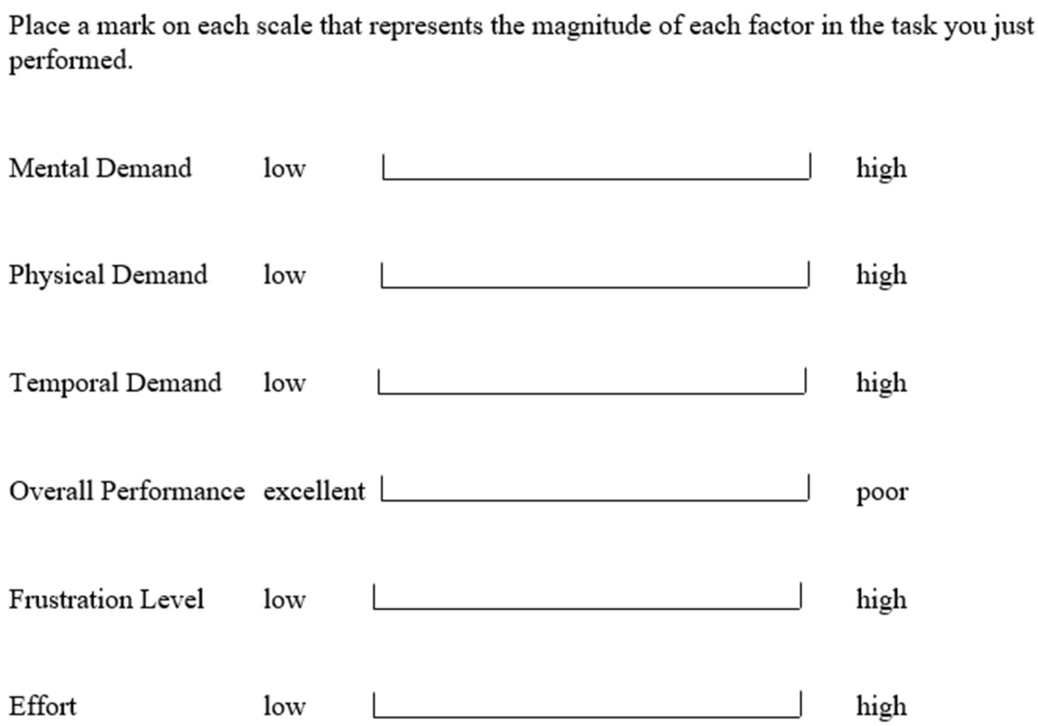
Wickens, 1990). Unlike physiological measurements, subjective indexes do not obtain responses in real-time. Researchers typically administer probes during a forced pause in the task, between task components, or after the task has been completed. This requires one to reflect on their experience instead of capturing real-time responses, relying on memory, and causing one to transition from executing the task to reading probes, potentially affecting response accuracy (Hart & Wickens, 1990).

Perceived workload scales vary in their composition. Some are unidimensional, meaning they can measure perceived workload through a single scale, while others utilize multiple subscales and a weighting calculation to determine perceived workload. These subscales target different underlying themes of perceived workload. Many fields aim to improve workload measurements to meet their specific needs (e.g., users, environments, and scenarios). Researchers have justified the need for designing multiple, unique, subjective indexes that measure perceived workload through slightly different scales due to specific procurement needs (Hart & Wickens, 1990). Some examples include Overall Workload (Vidulich & Tsang, 1987), Modified Cooper-Harper Scale (Wierwille & Casali, 1983), and Subjective Workload Assessment Technique (SWAT; Reid & Nygren, 1988). Some researchers advise using these tools under their intended circumstances for optimal perceived workload measurements, while others suggest that a good tool should be able to measure perceived workload independent of the specific discipline (Hart, 2068).

In this study we measured perceived workload through a subjective measure: the NASA-TLX (See Figure 1; Hart & Staveland, 1988; Hart, 2006). The NASA-TLX is a short, easy-to-administer subjective workload measurement which includes a weighting task that aids in balancing the scores to account for individual differences in perceived task difficulty. This is important in understanding how each participant interprets subscale descriptions and relates them to the specific task. The weighting task utilizes 15 pairwise questions and six subscales to calculate a mean perceived weighted workload score. The pairwise questions (See Table 1) compare each dimension of perceived workload against all other dimensions. The individual indicates responses for the questions after understanding the task but before being asked to record their subscale ratings. The NASA-TLX subscales include physical demand, mental demand, temporal demand, frustration level, overall performance, and effort (See Table 1).

Figure 1

NASA-TLX Rating Scales



Place a mark on each scale that represents the magnitude of each factor in the task you just performed.

Mental Demand	low	high
Physical Demand	low	high
Temporal Demand	low	high
Overall Performance	excellent	poor
Frustration Level	low	high
Effort	low	high

Note. Adapted from Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 139–183. https://doi.org/10.1016/s0166-4115(08)62386-9.

Table 1

NASA-TLX Sources-of-Workload Comparison Table

Column A		**Column B**
Physical Demand	OR	Mental Demand
Temporal Demand	OR	Mental Demand
Overall Performance	OR	Mental Demand
Frustration Level	OR	Mental Demand
Effort	OR	Mental Demand
Temporal Demand	OR	Physical Demand
Overall Performance	OR	Physical Demand
Frustration Level	OR	Physical Demand
Effort	OR	Physical Demand
Temporal Demand	OR	Overall Performance
Temporal Demand	OR	Frustration Level
Temporal Demand	OR	Effort
Overall Performance	OR	Frustration Level
Overall Performance	OR	Effort
Effort	OR	Frustration Level

Note. Adapted from Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 139–183. https://doi.org/10.1016/s0166-4115(08)62386-9. Instructions include "Select the member of each pair that provided the most significant source of workload variation in these tasks."

Table 2

NASA-TLX Rating Scale Definitions

Title	Endpoints	Descriptions
Mental Demand	Low/High	How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?
Physical Demand	Low/High	How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
Temporal Demand	Low/High	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
Performance	Good/Poor	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
Effort	Low/High	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration Level	Low/High	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Note. Adapted from Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 139–183. https://doi.org/10.1016/s0166-4115(08)62386-9.

Current Study

Memory Enrichment Theory (Ricker & Vergauwe, 2022) predicts that memory performance during multitasking is driven by task engagement determined by subject perception of task difficulty. The present work investigated the influence of perceived workload, a measure of subjective difficulty, on forgetting during multitasking as a potential alternative to cognitive load explanations. Individual differences in perceived difficulty can influence how one approaches, prioritizes, and ultimately executes a task. Fluctuations in psychological and physiological stress may alter the minimum threshold of acceptable performance quality and rate (Hart & Staveland, 1988). If perceived workload predicts memory performance across cognitive load conditions better than cognitive load itself, this would be strong support for Memory Enrichment Theory.

Existing cognitive theories of working memory performance during multitasking use highly controlled lab settings. In contrast, perceived workload measurements are common in applied settings such as airplanes (Hartman & McKenzie, 1979), operating rooms (Archampong et al., 2012), and vehicles (Recart & Nunes, 2003).The present study aimed to recontextualize cognitive load findings as a function of subjective experience rather than solely by objective conditions. Ricker and Vergauwe (2020, 2022) demonstrated that both the time-based and interference models are fundamentally flawed explanations of multitask forgetting. This work proposes exploring support for an alternative explanation. We investigated the measurement of perceived workload as a predictor for dual-task performance. Perceived workload is expected to increase when free time during the secondary task is reduced, ultimately predicting poor recall performance. Two research questions will be addressed by the proposed study. Is subjective workload a better predictor of working memory performance than cognitive load in a dual-task with low-level perceptual stimuli? Does overload condition performance (both primary and secondary) reflect “giving up” as characterized by poor performance and high perceived workload levels?

Methods

Participants

Data were collected from 114 participants (54% female, age *M* = 20.01, *SD* = 3.73) who consented to participate in exchange for course credit or on a voluntary basis. Two participants were not included in the analyses due to below chance performance in one of the tasks. One participant fell below chance in the primary task while the other fell below chance during the secondary task. This experiment was approved by the Institutional Review Board at the University of South Dakota.

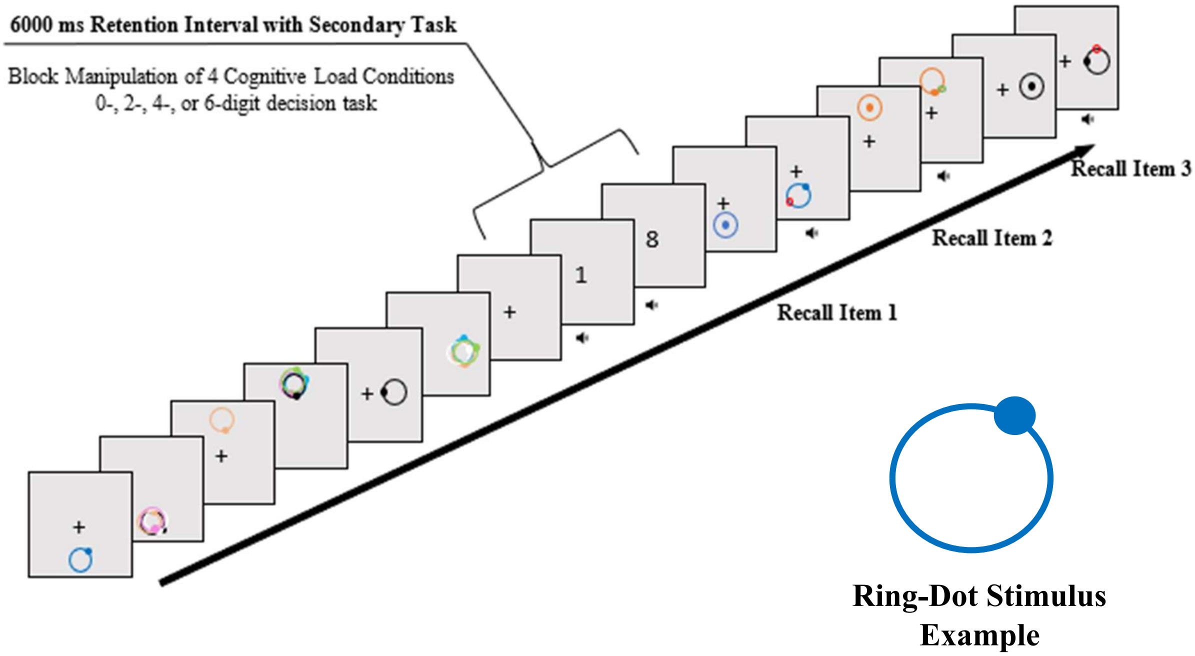
Materials & Equipment

The experiment was conducted locally at the University of South Dakota via PsychoPy3 (version 2020.2.10; Peirce et al., 2019). This study utilized desktop computers with all stimuli presented within the center 95.25 mm by 95.25 mm of the screen. Participants were seated at a comfortable distance, generally about 700 mm from the screen.

Stimuli included black text and colored ring-dot figures on a gray background. Ring diameter was 29.5 mm, while the dot diameter was 4.35 mm (see Figure 2 for an example). The presentation locations for the ring-dot stimuli included four different positions: above, right, below, and left of the center of the screen, with the ring centers located 29.5 mm from center of screen. The stimulus colors included blue, black, orange, lime, white, purple, and cyan. Colors were shuffled each trial to ensure randomization of stimulus color. The masks were composed of four colored ring-dot figures presented concurrently so that each offset the stimulus by 1.45 mm. The masking rings were different colors than the previously displayed stimulus selected at random without replacement.

Figure 2

An Example of a Single Experimental Trial with Enlarged Stimulus Image



Note. Trial proceeds through time from left to right. Three memory items are present, immediately followed by a perceptual mask. Following stimuli presentation, the participant is presented with the secondary task, a digit decision task. Load condition is controlled by block. Recall probes and feedback on all three memory items conclude each trial. An enlarged ring-dot stimulus is shown in the bottom right corner of this figure.

Procedure

A one-factor design tested response performance and perceived workload under varying cognitive load conditions (no, low, medium, and high load).

The dual-task design presented a primary memory span task with a secondary task during the retention interval. An example of a single trial is presented in Figure 2. Each trial began with a 500ms fixation cross, after which the first memory item was presented and immediately followed by a post-perceptual mask. The memory stimulus was a ring located randomly in one of four positions (above, right, below, left of center). The ring had a circle placed along its circumference at a random degree location (1 - 360), displayed for 250ms. This was followed by the post-perceptual mask, lasting 250 ms. The memory item and mask pair was repeated two additional times, each time with unique ring and dot locations and colors, resulting in a total of three memory items per trial. The secondary task occurred next (described in the following paragraph) during a 6000 ms retention period. After the retention period, the participant recalled the dot location for each stimulus presented in the current trial in the order they were presented.

Primary task feedback was provided immediately after each item response. If the response angle was within 30 degrees of the stimulus, a green ring replaced the selected location, a dot appeared in the original stimulus location, and a positive auditory tone sounded. If the response angle was greater than 30 degrees, a red ring presented at the response location, followed by corrective auditory feedback.

The secondary task was presented during the retention interval, immediately following the masking of the third memory item. The secondary task duration was held constant across all loads (6000ms) and consisted of the presentation of a 0-, 2-, 4-, or 6-digit sequence. The researcher instructed each participant to respond as quickly and accurately as possible through a keypress indicating if the displayed digit is even or odd. Labels were placed on the keyboard covering the original markings and labeling the ‘S’ key as ‘EVEN’ and the ‘D’ key as ‘ODD’. The manipulation of the number of digits over a constant period allowed for the induction of differing cognitive load levels: no (0 digits), low (2 digits), medium (4 digits), and high (6 digits) cognitive loads. In the no load condition, a fixation cross remained on screen for the duration of the secondary task in place of a digit. For the remaining conditions, stimuli were digits ranging in value between one and nine. Digits were drawn from a randomized list with replacement after each presentation, allowing the same number to be presented multiple times in a sequence, even consecutively. A 100ms blank screen was presented between each digit to discriminate consecutive digit presentations visually.

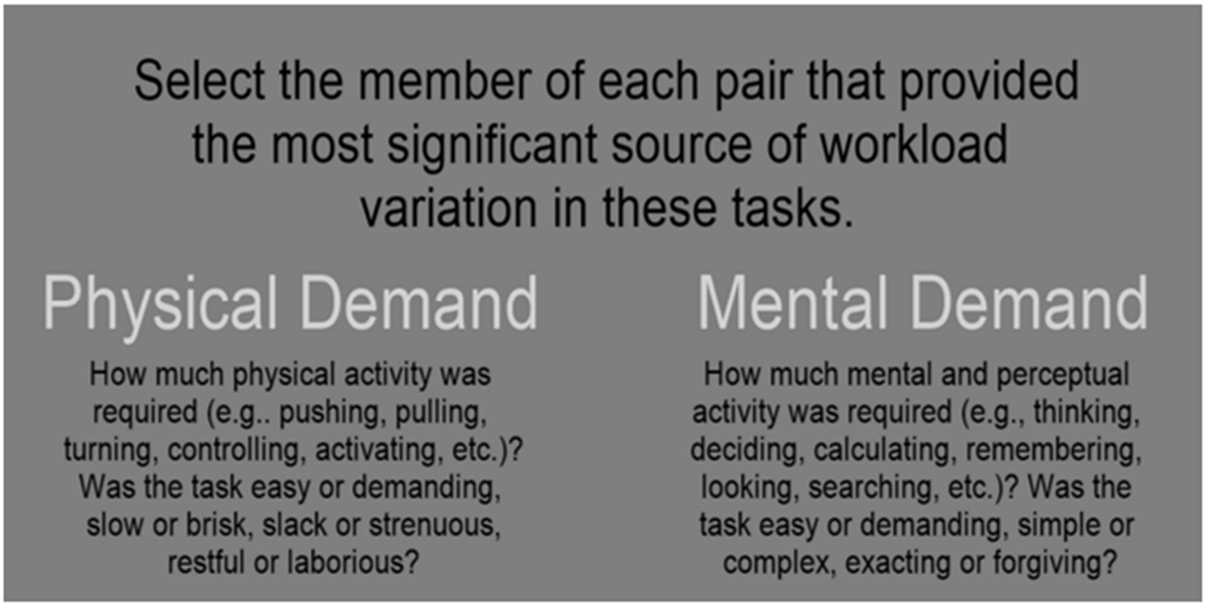
Digit presentation time for each condition was 2900ms, 1400ms, and 900ms for the low, medium, and high load conditions, respectively. Each digit was followed by a 100ms blank screen. Auditory feedback sounded at the end of each digit presentation to indicate response accuracy, not at the keyboard response time. A high pitched, rising tone sequence was played if the indicated response was within 15 degrees of stimulus, otherwise a low pitched, falling tone was played.

The experiment began with three different practice blocks. Practice block 1 consisted of four primary memory task alone trials in which participants recalled three visual memory items. Practice task 2 consisted of six secondary digit decision task alone trials (two trials per three conditions; no-load condition not included). Practice task 3 consisted of eight dual-task trials identical to the experimental task (two trials per each of the four conditions), randomized across participants.

The NASA-TLX weighting task (fifteen pairwise comparison choices; See Figure 3 for an example) followed the practice blocks. Participants indicated which pair contributed most to the perceived workload that they experienced with the final practice block in mind. The subscale responses were multiplied with their associated weight and normalized.

Figure 3

Example of Experimental Presentation of NASA-TLX Pairwise Comparison of Factors

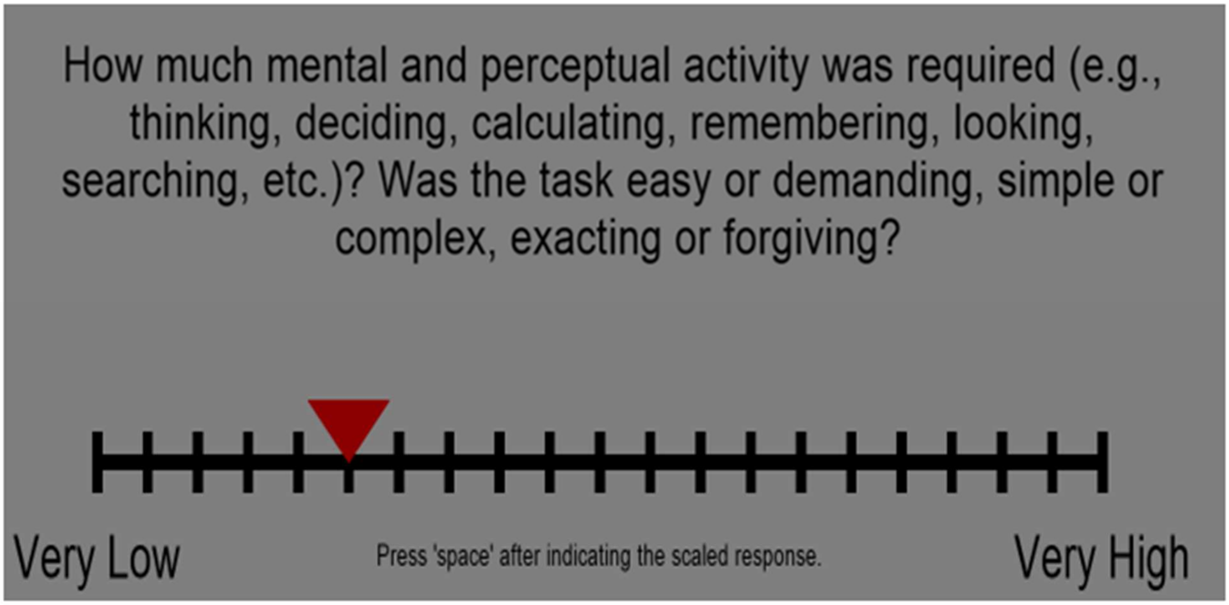


Note. This figure is a screen capture showing an example of how the pairwise questions from Hart & Staveland (1998) were presented in this experiment.

Experimental trials began after the completion of the weighing task. The experiment consisted of 4 blocks. The cognitive load level (none, low, medium, high) of the secondary task was manipulated by block with block order randomized. Each block consisted of 20 trials. This was reduced to 18 trials after the first 9 participants to allow more time for debriefing post-experiment. Presentation of the NASA-TLX workload index occurred after each block of experimental trials. During the NASA-TLX the participant moved a red triangle along sliding scales to reflect their perceived workload for each probe (See Figure 4 for an example).

Figure 4

Example of Experimental Presentation of NASA-TLX Rating Scales



Note. This figure is a screen capture showing an example of how subscale indices from Hart & Staveland (1998) were presented in this experiment.

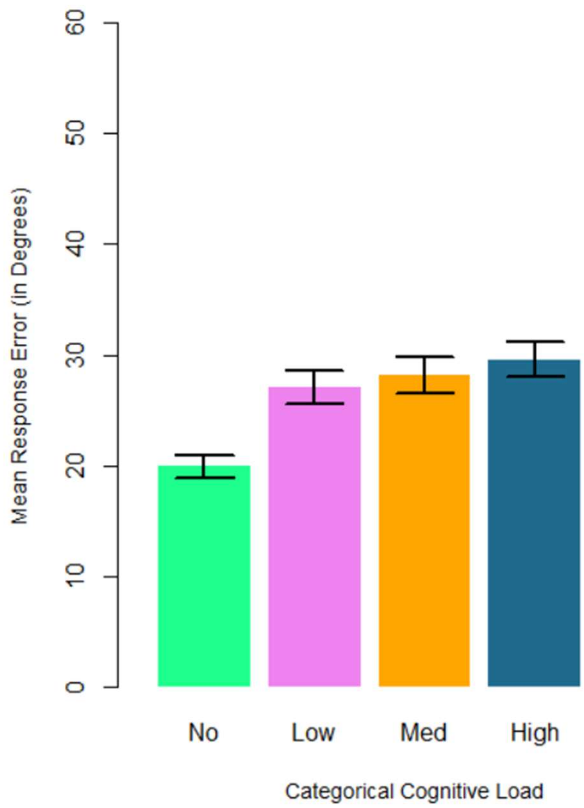
Results

Primary Task Error and Cognitive Load

Mean reproduction error in circular degrees was used to determine primary task accuracy. An ANOVA of categorical cognitive load (no, low, medium, and high load conditions; See Figure 5) on response error produced an effect of cognitive load, $F(3, 333) = 25.62$, Bayes factor $= 8.81 \times 10^{11}$ in favor of the alternative with means ascending from no load to high load conditions (means: no = 19.94, low = 27.10, medium = 28.22, high = 29.58). When excluding the no load condition the results reversed, $F(2, 222) = 1.93$, Bayes factor = 8.89 in favor of the null.

Figure 5

Average Primary Task Error by Categorical Cognitive Load

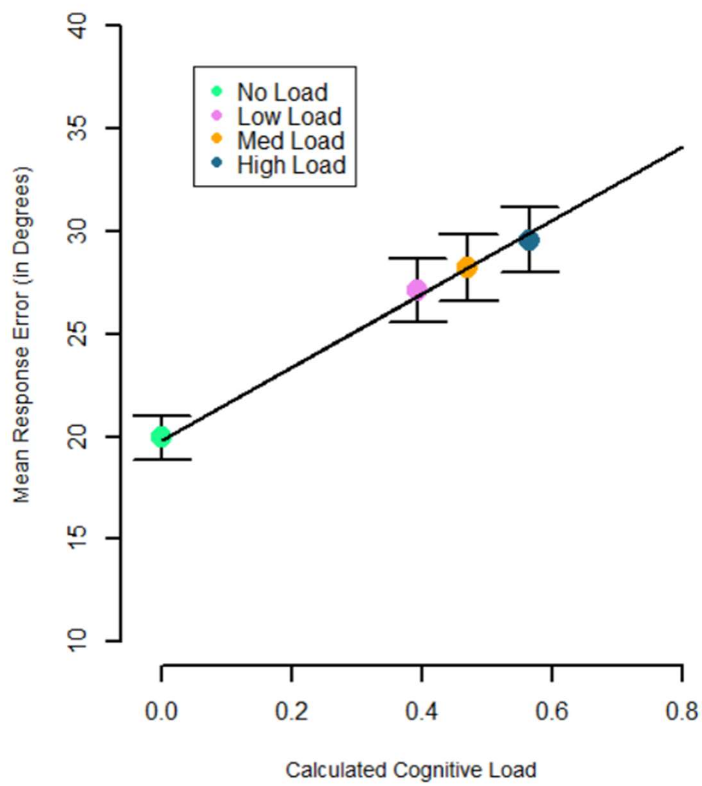


Note. Error bars represent standard error of the mean.

Two regressions of calculated cognitive load on primary task response error were conducted. The calculated cognitive load was operationalized as the mean participant response time in each condition multiplied by the number of processing task iterations in that condition. This was done to more accurately capture the cognitive load experienced by the participant as defined in the original TBRS calculations. The first analysis collapsed across participants. This regression predicted 99.75% of the variation in primary task error $F(1, 2) = 795.7$, Bayes factor = 1.59×10^{1}, $R^2 = 1.00$ (See Figure 6). The second regression used the calculated cognitive load of each individual participant in each condition. Participant was included as a main effect in this second model. Results indicated cognitive load and subject predicted 76.23% of the variation in primary task error $F(112, 335) = 9.59$, Bayes factor = 7.23×10^{56}, $R^2 = 0.76$.

Figure 6

Mean Response Error by Calculated Cognitive Load



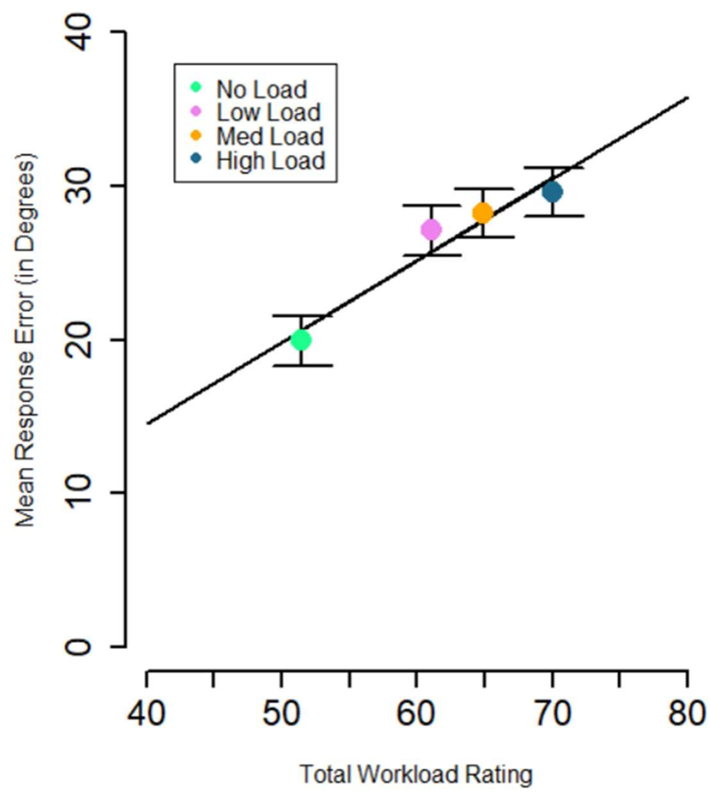
Note. Response times measured in milliseconds. Error bars represent standard error of the mean.

Primary Task Error and Perceived Workload

Two regressions were conducted of total unweighted perceived workload on primary task response error. The first analysis collapsed across participants. The results of the regression indicated mean total unweighted perceived workload explained 93.77% of the variation in primary task error $F(1, 2) = 30.1$, Bayes factor = 3.05, $R^2 = 0.94$ (See Figure 7). The second regression was of mean performance in each condition for each participant. Participant was included as a main effect in this second model. The results of the regression indicated total unweighted perceived workload and participant explained 77.13% of the variation in primary task error $F(112, 335) = 10.09$, Bayes factor = $8.02 \times 10^{56,}$ $R^2 = 0.77$.

Figure 7

Mean Response Error by Total Unweighted Perceived Workload

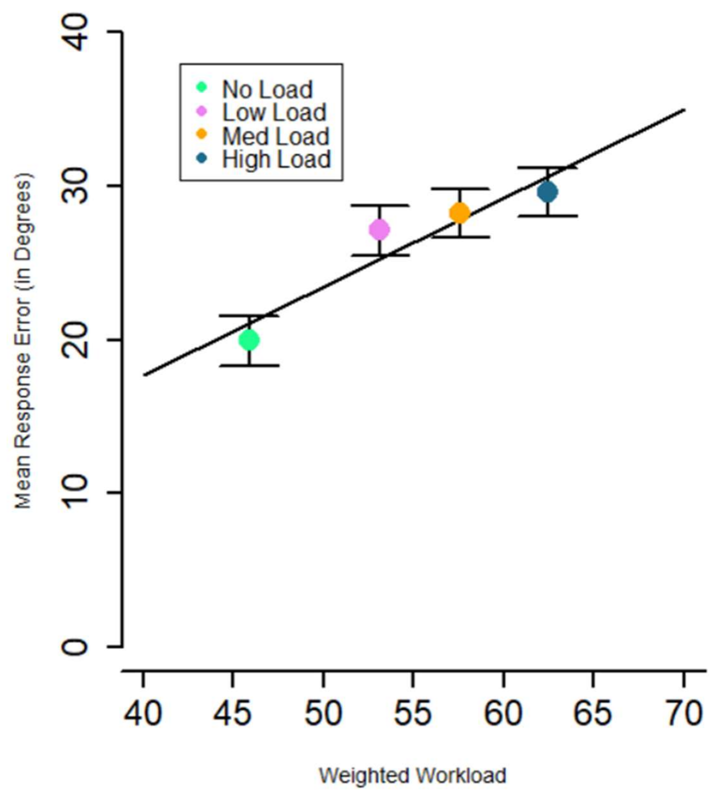


Note. Error bars represent standard error of the mean. Unweighted (raw) NASA-TLX scores range from 0 – 120.

Two regressions were conducted of weighted perceived workload on primary task response error. The first analysis collapsed across participants. The results of the regression indicated mean weighted perceived workload explained 83.9% of the variation in primary task error, $F(1, 2) = 16.68$, Bayes factor = 2.26, $R^2 = 0.84$ (See Figure 9). The second regression was of mean performance in each condition for each participant. Participant was included as a main effect in this second model. The results of the regression indicated weighted perceived workload and participant explained 69.4% of the variation in primary task error $F(112, 335) = 10.09$, Bayes factor = 5.34×10^{57}, $R^2 = 0.69$.

Figure 8

Mean Response Error by Weighted Perceived Workload



Note. Error bars represent standard error of the mean. Weighted NASA-TLX scores range from 0 – 100.

Perceived Workload and Cognitive Load

Repeated-Measures ANOVA of mean total unweighted perceived workload as a function of cognitive load condition and participant shows an effect, $F(3,333) = 64.88$, Bayes factor = 6.17×10^{29} in favor of the alternative (means: no = 51.50, low = 61.06, medium = 64.92, high = 70.07). The results of the regression indicated cognitive load and participant explained 95.97% of the variation in total unweighted perceived workload $F(1,2) = 47.57$, Bayes factor = 3.85, $R^2 = 0.96$. While calculated cognitive load and participant explained 78.82% of the variation in total unweighted perceived workload $F(112, 335) = 11.13$, Bayes factor = 6.13×10^{65}, $R^2 = 0.79$.

Repeated-Measures ANOVA of mean weighted perceived workload as a function of cognitive load condition and participant shows an effect, $F(3,333) = 47.76$, Bayes factor = 3.49×10^{22} in favor of the alternative (means: no = 45.87, low = 53.11, medium = 57.57, high = 62.40). The results of the regression indicated cognitive load and participant explained 92.18% of the variation in weighted perceived workload $F(1, 2) = 23.58$, Bayes factor = 2.69, $R^2 = 0.92$. While calculated cognitive load and participant explained 71.78% of the variation in weighted perceived workload $F(112, 335) = 7.61$, Bayes factor = 1.91×10^{50}, $R^2 = 0.72$.

Analysis of the relationship between individual subscales and cognitive load is detailed in Table 3. Mean workload weightings were calculated from the pairwise response questions before experimental trials (See Table 4).

References

Archampong, D., Borowski, D., Wille-Jørgensen, P., & Iversen, L. H. (2012). Workload and surgeon´s specialty for outcome after colorectal cancer surgery. *Cochrane Database of Systematic Reviews*, *3*. https://doi.org/10.1002/14651858.CD005391.pub3

Baddeley, A. (1986). *Working memory* (pp. xi, 289). Clarendon Press/Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(3), 570–585. https://doi.org/10.1037/0278-7393.33.3.570

Barrouillet, P., & Camos, V. (2001). Developmental increase in working memory span: Resource sharing or temporal decay? *Journal of Memory and Language*, *45*, 1–20. https://doi.org/10.1006/jmla.2001.2767

Barrouillet, P., & Camos, V. (2007). The time-based resource-sharing model of working memory. *The Cognitive Neuroscience of Working Memory*. https://doi.org/10.1093/acprof:oso/9780198570394.003.0004

Barrouillet, P., De Paepe, A., & Langerock, N. (2012). Time causes forgetting from working memory. *Psychonomic Bulletin & Review*, *19*(1), 87–92. https://doi.org/10.3758/s13423-011-0192-8

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working Memory Span Development: A Time-Based Resource-Sharing Model Account. *Developmental Psychology*, *45*, 477–490. https://doi.org/10.1037/a0014615

Barrouillet, P., & Lecas, J.-F. (1999). Mental models in conditional reasoning and working memory. *Thinking & Reasoning*, *5*, 289–302. https://doi.org/10.1080/135467899393940

Brown, G. (1986). Case, R. (1985). Intellectual Development: Birth to Adulthood. *British Journal of Educational Psychology*, *56*(2), 220–222. https://doi.org/10.1111/j.2044-8279.1986.tb02666.x

Brown, J. (1958). Some Tests of the Decay Theory of Immediate Memory. *Quarterly Journal of Experimental Psychology*, *10*(1), 12–21. https://doi.org/10.1080/17470215808416249

Casali, J. G., & Wierwille, W. W. (1983). A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in a simulated flight task emphasizing communications load. *Human Factors*, *25*, 623–641.

Case, R., Kurland, D. M., & Goldberg, J. (1982a). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386–404. https://doi.org/10.1016/0022-0965(82)90054-6

Case, R., Kurland, D. M., & Goldberg, J. (1982b). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386–404. https://doi.org/10.1016/0022-0965(82)90054-6

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163–191. https://doi.org/10.1037/0033-2909.104.2.163

Cowan, N. (1995). *Attention and memory: An integrated framework* (pp. xv, 321). Oxford University Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309. Farrell, S., & Lewandowsky, S. (2002). An endongeous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, *9*, 59–79. https://doi.org/10.3758/BF03196257

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908. https://doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human mental workload* (pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hart, S. G., & Wickens, C. D. (1990). Workload Assessment and Prediction. In H. R. Booher (Ed.), *Manprint: An Approach to Systems Integration* (pp. 257–296). Springer Netherlands. https://doi.org/10.1007/978-94-009-0437-8_9

Hartman, B. O., & McKenzie, R. E. (1979). *Survey of methods to assess workload*. Advisory Group for Aerospace Research and Development Neuilly-Sur-Seine (France).

Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm. *Memory & Cognition*, *35*(7), 1675–1684. https://doi.org/10.3758/BF03193501

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433. https://doi.org/10.1016/S0160-2896(05)80012-1

Lemaire, B., & Portrat, S. (2018). A computational model of working memory integrating time-based decay and interference. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00416

Lépine, R., Bernardin, S., & Barrouillet, P. (2005). Attention switching and working memory spans. *European Journal of Cognitive Psychology*, *17*(3), 329–345. https://doi.org/10.1080/09541440440000014

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review*, *11*(5), 771–790. https://doi.org/10.3758/bf03196705

Lewandowsky, S., & Oberauer, K. (2009). No evidence for temporal decay in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1545–1551. https://doi.org/10.1037/a0017010

Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences*, *13*(3), 120–126. https://doi.org/10.1016/j.tics.2008.12.003

McKeown, D., & Mercer, T. (2012). Short-term forgetting without interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1057–1068. https://doi.org/10.1037/a0027749

Morey, C. C., & Cowan, N. (2005). When Do Visual and Verbal Memories Conflict? The Importance of Working-Memory Load and Retrieval. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *31*, 703–713. https://doi.org/10.1037/0278-7393.31.4.703

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251–269. https://doi.org/10.3758/BF03213879

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 411–421. https://doi.org/10.1037/0278-7393.28.3.411

Oberauer, K., & Kliegl, R. (2001). Beyond resources: Formal models of complexity effects and age differences in working memory. *European Journal of Cognitive Psychology*, *13*, 187–215. https://doi.org/10.1080/09541440042000278

Oberauer, K., & Kliegl, R. (2004). Simultaneous Cognitive Operations in Working Memory After Dual-Task Practice. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 689–707. https://doi.org/10.1037/0096-1523.30.4.689

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review*, *115*, 544–576. https://doi.org/10.1037/0033-295X.115.3.544

Oberauer, K., & Lewandowsky, S. (2013a). Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General*, *142*, 380–411. https://doi.org/10.1037/a0029588

Oberauer, K., & Lewandowsky, S. (2013b). Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General*, *142*, 380–411. https://doi.org/10.1037/a0029588

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*(5), 779–819. https://doi.org/10.3758/s13423-012-0272-4

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. Psychological bulletin, 116(2), 220.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods, 51,* 195-203.

Portrat, S., Barrouillet, P., & Camos, V. (2008). Time-related decay or interference-based forgetting in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1561–1564. https://doi.org/10.1037/a0013356

Portrat, S., & Lemaire, B. (2015). Is attentional refreshing in working memory sequential? A computational modeling approach. *Cognitive Computation*, *7*, 333–345.

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied*, *9*, 119–137. https://doi.org/10.1037/1076-898X.9.2.119

Reid, G. B., & Nygren, T. E. (1988). The Subjective Workload Assessment Technique: A scaling procedure for measuring mental workload. In *Human mental workload* (pp. 185–218). North-Holland. https://doi.org/10.1016/S0166-4115(08)62387-0

Ricker, T. J., & Cowan, N. (2010a). Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1355–1368. https://doi.org/10.1037/a0020356

Ricker, T. J., & Cowan, N. (2010b). Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1355–1368. https://doi.org/10.1037/a0020356

Ricker, T. J., Cowan, N., & Morey, C. C. (2010). Visual working memory is disrupted by covert verbal retrieval. *Psychonomic Bulletin & Review*, *17*(4), 516–521. https://doi.org/10.3758/PBR.17.4.516

Ricker, T. J., & Vergauwe, E. (2020). Consistent Failure to Produce a Cognitive Load Effect in Visual Working Memory Using a Standard Dual-Task Procedure. *Journal of Cognition*, *3*(1), 15. https://doi.org/10.5334/joc.108

Ricker, T. J., & Vergauwe, E. (2022). Boundary conditions for observing cognitive load effects in visual working memory. *Memory & Cognition*, *50*(6), 1169–1185. https://doi.org/10.3758/s13421-022-01320-3

Saito, S., & Miyake, A. (2004). On the nature of forgetting and the processing–storage relationship in reading span performance. *Journal of Memory and Language*, *50*(4), 425–443. https://doi.org/10.1016/j.jml.2003.12.003

Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—And a little bit more. *Intelligence*, *30*(3), 261–288. https://doi.org/10.1016/S0160-2896(01)00100-3

Towse, J., & Hitch, G. J. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *48A*, 108–124. https://doi.org/10.1080/14640749508401379

Towse, J. N., Hitch, G. J., & Hutton, U. (1998). A Reevaluation of Working Memory Capacity in Children. *Journal of Memory and Language*, *39*(2), 195–217. https://doi.org/10.1006/jmla.1998.2574

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5

Vergauwe, E., Barrouillet, P., & Camos, V. (2009). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1012–1028. https://doi.org/10.1037/a0015859

Vergauwe, E., Barrouillet, P., & Camos, V. (2010). Do mental processes share a domain-general resource? *Psychological Science*, *21*(3), 384–390.

Vergauwe, E., Camos, V., & Barrouillet, P. (2014). The impact of storage on processing: How is information maintained in working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1072–1095. https://doi.org/10.1037/a0035779

Vergauwe, E., Dewaele, N., Langerock, N., & Barrouillet, P. (2012). Evidence for a central pool of general resources in working memory. *Journal of Cognitive Psychology*, *24*(3), 359–366. https://doi.org/10.1080/20445911.2011.640625

Vidulich, M. A., & Tsang, P. S. (1987). Absolute Magnitude Estimation and Relative Judgement Approaches to Subjective Workload Assessment. *Proceedings of the Human Factors Society Annual Meeting*, *31*(9), 1057–1061. https://doi.org/10.1177/154193128703100930