Cognitive Load and the Role of Cue Utilisation during Sustained Attention Tasks

Daniel Sturman

PhD/Master of Organisational Psychology Candidate

Macquarie University, Sydney, Australia

BPsych (Hons), Department of Psychology,

Macquarie University, Sydney, Australia

Thesis submitted on the 4th day of March, 2019

Table of Contents

Declaration

This thesis has not been submitted for a higher degree to any other university or institution. The work is predominantly that of the PhD candidate. Components of the thesis that involved collaboration have been noted as such. The Macquarie University Ethics Committee approved the research reported in this thesis on the 28th of February 2017 (Reference No 5201700078), 21st November 2017 (Reference No 5201700154), and 26th March 2018 (Reference No 5201800150).

Daniel Sturman

4th March 2019

Acknowledgments

I would like to take this opportunity to thank my supervisor, Professor Mark Wiggins, for his significant contribution to this thesis. Mark's enthusiasm for research is contagious, and his knowledge, guidance, and support have helped make me a better researcher. Mark has been an amazing mentor, and this thesis would not have been possible without him. I would also like to thank Dr. Jaime Auton. Her insight, practical knowledge, and encouragement have been invaluable. Lastly, I would like to thank my partner, Catherine Gaussen, for the many hours of proof reading and her unwavering love and support.

Summary

One of the key roles of system controllers in high risk environments, such as rail control, power control, and driving, is to recognise, anticipate and respond to changes in the system state. This requires that operators sustain a visual search to monitor and control their operating systems, often for extended periods. However, attentional resource theory posits that sustaining attention over extended periods results in the consumption of cognitive resources, thereby reducing the residual resources available to manage changes in the system state. The utilisation of cues is a cognitive strategy that operators engage to reduce the rate at which cognitive resources are consumed. The aim of this thesis was to examine whether differences in cue utilisation are associated with differences in the rate at which cognitive resources are consumed across a range of operational settings, under a range of operational conditions, and using operators of varying levels of operator experience.

Studies 1 and 2 were conducted to establish whether a general capacity for cue utilisation predicts differences in the consumption of cognitive resources during sustained attention tasks. The results of Study 1 indicated that participants with higher cue utilisation recorded smaller increases in mean response latency during a novel 30-minute rail control simulation, compared to participants with lower cue utilisation. Study 2 replicated and extended these results using a 45-minute rail control simulation, demonstrating greater decreases in fixation rates, smaller changes in cerebral oxygenation in the prefrontal cortex, and smaller increases in mean response latency for participants with higher cue utilisation, compared to participants with lower cue utilisation. These results are consistent with the proposition that cue utilisation is associated with the allocation of fewer cognitive resources to sustained attention tasks.

Study 3 was designed to validate a newly adapted sustained visual search task for process control environments. The results revealed changes in response latency throughout the sustained visual search task that were positively associated with changes in response latency

during a 30-minute low signal probability rail control task, a 45-minute low signal probability rail control task, and a 45-minute high signal probability rail control task. The findings suggest that the sustained visual search task is a valid alternative to a longer-duration process control task for experimental studies.

Study 4 was intended to examine whether experienced operators' cue utilisation differentiates performance during domain-relevant sustained attention tasks. In two experiments, power distribution operators with higher cue utilisation demonstrated shorter mean response latencies during a power control sustained visual search task, compared to operators with lower cue utilisation. These results support the view that experienced operators with higher cue utilisation adopt strategies during operational tasks that reduce the demands on cognitive load.

Study 5 was designed to establish whether differences in cognitive load based on cue utilisation are also evident in more dynamic operational environments. Using motor vehicle driving as a context, drivers' consumption of cognitive resources were examined during a 20-minute, simulated driving task. Qualified drivers with higher cue utilisation demonstrated smaller mean visual saccade amplitudes, smaller mean fixation dispersions, smaller increases in cerebral oxygenation and recorded fewer missed traffic signals during the simulated driving task, compared to participants with lower cue utilisation. These results are consistent with the broader proposition that experienced operators adopt more efficient search patterns, and consume fewer cognitive resources, during dynamic operational tasks.

Extending these findings, Study 6 assessed physiological measures of cognitive resource consumption during periods of power distribution operators' regular workdays. Across two testing sessions, and controlling for subjective measures of workload, higher cue utilisation was associated with smaller increases in cerebral oxygenation in the prefrontal cortex, which is indicative of lower cognitive load. These results indicate that experienced operators with higher cue utilisation are consuming fewer cognitive resources during typical operational tasks, compared to operators with lower cue utilisation.

The outcomes from this programme of research provide a number of theoretical contributions that advance an understanding of the relationship between cue utilisation and sustained attention. These contributions include support for the proposition that cue utilisation is associated with the consumption of fewer cognitive resources during sustained attention tasks, the identification of search pattern efficiency as a potential mediator for the relationship between cue utilisation and cognitive load, and evidence to support the resource depletion theory of the vigilance decrement. These findings have applied implications for the selection, management, and training of operators in high risk environments.

Chapter One

Introduction

Sustained attention

In contemporary high risk industrial environments, such as power control, rail control and aviation, system operators are required to sustain attention for extended periods of time (Edkins & Pollock, 1997; Lee, Kim, & Jang, 2011; Mackworth, 1948; Vicente, Roth, & Mumaw, 2001). Sustained attention refers to the ability of operators to remain alert and focused, and respond rapidly and accurately to deviations in the system state (Langner & Eickhoff, 2013). Sustained attention allows operators to take timely and appropriate actions, helping to ensure optimal efficiency and the safety of the system (O'Hara & Hall, 1992; Stanton, Salmon, Walker, & Jenkins, 2009). The failure to sustain attention can result in inaccurate or delayed responses, increasing the likelihood of negative outcomes ranging from system inefficiency to catastrophic disasters (Helton & Warm, 2008; Reason, 2000; Warm, 1984).

An operator's capacity to sustain attention is influenced by the characteristics of the operational task (Helton & Russell, 2013; Thiffault & Bergeron, 2003). For instance, tasks requiring less operator interaction are associated with a reduced capacity to sustain attention (Molloy & Parasuraman, 1996; Parasuraman, Molloy, & Singh, 1993; Warm, 1984). In contemporary industrial environments, the level of interaction between the operator and the operating system varies between workplaces (Anders, Seijmonsbergen, & Bouten, 2011; Hanauer, Englesbe, Cowan, & Campbell, 2009; Metzger & Polakow, 2011). Notably, more automated environments tend to require greater monitoring and less interaction by human operators (Hanauer et al., 2009; Parasuraman, Sheridan, & Wickens, 2000).

Monitoring tasks refer to system operations that require infrequent human interaction. In monitoring environments, system operations tend to be automated, and therefore, overt responses to critical signals are required infrequently (Davies & Parasuraman, 1982). Operators in monitoring environments act mainly in a fail-safe capacity, monitoring the system for deviations and intervening only when problems or system failures occur (Reinerman-Jones, Matthews, Langheim, & Warm, 2011; Sheridan, 1970).

Process control tasks refer to operational tasks that require routine adjustments in response to more regular system deviations (Cuny, 1979; Kaber & Endsley, 2004; Nachreiner, Nickel, & Meyer, 2006). Process control environments, where systems tend to be semi-automated, require a combination of monitoring and regular interaction with the system (Nachreiner & Nickel, 2006). For example, industrial environments such as air traffic control and energy transmission, require periods of monitoring and periods of frequent human intervention to ensure that the systems operate safely and at an optimal level (Amaldi & Leroux, 1995; Loft, Sanderson, Neal, & Mooij, 2007).

Both process control and monitoring environments require operators to sustain attention to rapidly identify changes in the system state. Failures to sustain attention can result in delayed responses or missed signals, reducing the likelihood that appropriate interventions will be initiated in sufficient time to maintain system efficiency and avoid system failure (Hitchcock et al., 2003; Warm, Parasuraman, & Mathews, 2008).

Sustained attention in monitoring environments

Sustained attention is typically assessed through measures of response latency and accuracy in response to critical signals (Lim & Dinges, 2008; Steinborn, Flehmig, Westhoff, & Langner, 2009). Smaller increases in mean response latency and the number of missed signals per unit time reflects a greater capacity to remain alert and maintain attention, and is therefore indicative of greater sustained attention (Davies & Parasuraman, 1982; Doran, Van Dongen, & Dinges, 2001; Gunzelmann, Gross, Gluck, & Dinges, 2009)

A considerable body of research exists regarding factors influencing sustained attention in monitoring environments (Endsley & Kiris, 1995; Molloy & Parasuraman, 1996; Shaw et al., 2010; Warm et al., 2008). A consistent finding in this body of literature is that, as operators sustain attention throughout a monitoring task, their performance typically decreases as a function of time (Mackworth, 1948; Parasuraman, 1979; Rodger, 1983; Teichner, 1974). Specifically, throughout a monitoring task, the mean response latency and/or the frequency of errors in response to critical signals tends to increase (Helton et al., 2007; Parasuraman, 1979; Teichner, 1974). This effect is referred to as the vigilance decrement (Parasuraman, 1979).

The vigilance decrement has been demonstrated in both laboratory settings and operational environments, and with a variety of visual and auditory vigilance tasks (Colquhoun, 1967; Lieberman et al., 2006; Pigeau, Angus, O'Neill, & Mack, 1995; Young, Robinson, & Alberts, 2009). Vigilance tasks in experimental studies typically require that participants sustain attention for a period of time, during which they are required to differentiate target from nontarget stimuli (Davies & Parasuraman, 1982; Helton & Russell, 2015; Parasuraman, Warm, & Dember, 1987; Shaw et al., 2010). A defining feature of traditional vigilance tasks is that they require participants to respond overtly to infrequent targets, while suppressing responses to frequently presented, nontarget stimuli (Davies & Parasuraman, 1982; Parasuraman, 1979; Teichner, 1974). Laboratory vigilance tasks have been designed to reflect monitoring tasks, which require similarly infrequent responses to critical signals (Kass, Vodanovich, Stanny, & Taylor, 2001; Mackworth, 1948; Parasuraman, Mouloua, Molloy, & Hilburn, 1996). Consequently, the outcomes of vigilance research tend to provide predictions that relate primarily to performance in monitoring environments.

Theories of vigilance decrement

A number of competing theories have been proposed to explain the vigilance decrement, including the resource depletion hypothesis, the underload account, and the

Malleable Attentional Resources Theory (MART). These theories are based on the proposition that the relative proportion of cognitive resources allocated to a task can influence the quantity and quality of information processing (Gopher & Donchin, 1986; Kahneman, 1973; Navon & Gopher, 1979).

Underload account

The underload account of the vigilance decrement is based on the premise that monitoring tasks are monotonous and boring (Manly, Robertson, Galloway, & Hawkins, 1999; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997; Scerbo, 1998). The underload account posits that, after extended periods of responding infrequently to critical signals separated by long intervals, participants' attentional resources cease to be allocated to the vigilance task as a result of understimulation (Norman & Shallice, 1986; Robertson et al., 1997; Stuss, Shallice, Alexander, & Picton, 1995). Consequently, cognitive resources are not directed towards the detection of system deviations, which results in delayed responses or missed targets (Manly et al., 1999; Robertson et al., 1997).

The mindlessness hypothesis is an underload account of the vigilance decrement, which holds that the withdrawal of attentional resources from an understimulating task results in a 'mindless' approach to the task (Robertson et al., 1997). As the time-on-task increases, the withdrawal of cognitive resources from the primary task results in an automatic, or 'mindless', approach to the task, whereby participants cease to pay attention to the task and respond to signals in a thoughtless manner (Thomson, Besner, & Smilek, 2015). Over time, this automatic responding results in responses to infrequent critical signals as though they were frequent nontargets, increasing the likelihood of missed signals (Manly et al., 1999; Smallwood & Schooler, 2006).

Empirical support for the mindlessness hypothesis can be derived from research demonstrating that participants consider vigilance tasks as 'boring' (Scerbo, 1998), and that participants report increased rates of daydreaming over time during vigilance tasks (Smallwood et al., 2004). Further support for the mindlessness hypothesis derives from research using the Sustained Attention to Response Task (SART; Robertson et al., 1997), which is a modified vigilance task designed to promote mindlessness. The SART is a short-duration sustained attention task that requires participants to *suppress* responses to infrequent critical signals, while *responding overtly* to frequent, neutral stimuli. The failure to detect critical signals is attributed to absent-mindedness and automaticity (Robertson et al., 1997). Robertson et al. (1997) provide support for the mindlessness hypothesis, demonstrating that absent-minded participants, as measured by the Cognitive Failures Questionnaire (CFQ; Broadbent, Cooper, FitzGerald, & Parkes, 1982), perform more poorly on the SART, compared to participants who are less absent-minded.

A shortcoming of the mindlessness account of the vigilance decrement is that, while it is based on the principle that attention is withdrawn from the monitoring tasks, it does not address whether attention is simply withdrawn, or whether attentional resources are redirected elsewhere. The mind-wandering hypothesis overcomes the shortfall of the mindlessness hypothesis, suggesting that attentional resources are redirected to task-unrelated thoughts (Smallwood & Schooler, 2006). Task-unrelated thoughts, or mind-wandering, tends to occur when a task is understimulating (McVay & Kane, 2012; Schooler et al., 2011). Consequently, the mind-wandering hypothesis posits that, throughout an understimulating, low demand monitoring task, the likelihood of mind-wandering increases, which decreases task performance as cognitive resources are directed away from the monitoring task (Schooler et al., 2011; Smallwood, Beach, Schooler, & Handy, 2008).

In comparison to the mindlessness hypothesis, relatively little vigilance-related research has been targeted towards mind-wandering. Empirical support for the mind-wandering hypothesis is based on evidence indicating that: (a) mind-wandering consumes the same attentional resources as those required for primary tasks (Smallwood, 2010; Smallwood

& Schooler, 2006), (b) low demand tasks result in high levels of mind-wandering over time (Giambra, 1995; Smallwood et al., 2004), and (c) mind wandering redirects attention away from the external environment towards task unrelated thoughts (Schooler et al., 2011; Smallwood et al., 2008). In combination, this evidence provides support for the proposition that understimulation will result in the attentional resources required for sustained attention being redirected away from the sustained attention task.

Despite the evidence, a number of empirical findings appear to challenge the underload account of vigilance decrement (Grubb et al., 1994; Helton et al., 2005). For instance, Helton et al. (2005) demonstrated that implicit, temporal patterns of target stimuli enhance performance during a traditional vigilance task. Helton et al. argue that, regardless of whether participants consciously identified the implicit pattern, they must have actively, or mindfully, attended to the target stimuli for the implicit pattern to influence their performance. Further, while absent-minded individuals perform relatively poorly on the SART (Robertson et al., 1997), absent-mindedness does not differentiate performance on a traditional vigilance task (Grubb et al., 1994). Nevertheless, absent-mindedness is associated with self-reports of higher cognitive demands during vigilance tasks (Grubb et al., 1994), suggesting that sustained attention may be cognitively demanding.

Kaplan's (1995) Attentional Resource Theory (ART) posits that attention comprises two distinct systems, consisting of direct (actively controlled) attention and effortless (passively controlled) attention. From this perspective, operational tasks require effortful, direct attention. Over time, this effort results in direct attention fatigue, whereby the capacity to direct and control attention is reduced (Kaplan, 1995). Kaplan proposes that recovery from direct attention fatigue is facilitated by time in a restorative environment, where one's attention is held effortlessly. Similarly, effortless attention can be engaged by mentally decoupling from the direct attention task through processes such as mind wandering (Baldwin & Lewis, 2017; Smallwood, 2011). Therefore, while the ART supports the notion that the vigilance decrement is associated with mind wandering, it implies that performance decrements and mind wandering result from the cognitive demands associated with sustained attention, rather than understimulation.

Malleable Attentional Resources Theory

The Malleable Attentional Resources Theory (MART) is based on the proposition that more cognitively demanding tasks lead to the mobilization of additional resources to meet increased task demands (Young & Stanton, 2002). Conversely, attentional resources are posited to shrink during periods of low workload to accommodate the reduction in task demands (Young & Stanton, 2002). During these periods of low workload, operators have insufficient resources to respond to sudden increases in task demands. Therefore, MART proposes that less stimulating tasks result in greater decrements in performance, as participants have fewer attentional resources available for the detection of critical signals (Young & Stanton, 2002).

The original evidence supporting the MART is derived from a driving simulator experiment, during which the level of vehicle automation was manipulated, while mental workload was assessed via a secondary task (Young & Stanton, 2002). Young and Stanton (2002) calculated attention ratio scores by dividing performance on the secondary task by the amount of visual attention directed at the secondary task, and observed that, as mental workload decreased, the allocation of visual attention to the secondary task became less efficient. Young and Stanton posit that this decrease in efficiency represents a shrinkage in attentional resources.

Despite these initial results, a key limitation of the MART is that empirical assessment of the theory is difficult. Measures of the availability of attentional resources typically rely on inferences derived from eye behaviour metrics and performance data (Stanton, Young, & Walker, 2007; Young & Stanton, 2002). However, these inferences are unable to differentiate the availability of attentional resources from the effort invested to perform the task. For instance, low demands may reduce operator effort, resulting in the allocation of fewer resources to the task, despite the availability of additional attentional resources (Matthews & Desmond, 1997). Desmond and Hoyes (1996) suggested that decreases in operator performance during periods of low task demands may be due to a failure to mobilise effort appropriately to match the task. Therefore, while MART could explain the performance decrements during low workload tasks, there is currently insufficient empirical evidence to test the theory conclusively.

A further limitation of the MART is that it does not as easily account for the increased rate of decline in performance commonly observed during more cognitively demanding vigilance tasks (Helton & Russell, 2013; Shaw, Funke, et al., 2013). Based on the MART, greater task demands should mobilise additional cognitive resources, either maintaining or increasing performance. Consequently, the MART does not adequately explain performance decrements during more cognitively demanding, sustained attention tasks.

Resource Depletion Theory

The Resource Depletion Theory, or overload account, of the vigilance decrement is based on the proposition that sustaining attention is cognitively demanding, and results in the consumption of attentional resources (Grier et al., 2003; Helton & Russell, 2012; Kahneman, 1973). These attentional resources are drawn from a limited pool of cognitive resources (Wickens, 1980; Young, Brookhuis, Wickens, & Hancock, 2015). Consequently, sustained attention reduces the availability of residual cognitive resources (Wickens, 1980). The associated depletion in cognitive resources that occurs over time eventually results in fewer resources than are necessary to remain alert and respond quickly and accurately (Grier et al., 2003). In turn, this leads to a decline in performance, which becomes evident in an increased response latency or decreased accuracy in response to critical signals (Helton et al., 2007; Parasuraman, 1979; Teichner, 1974).

Empirical support for the resource depletion account of the vigilance decrement is drawn from research demonstrating that more cognitively demanding tasks result in a more rapid decline in performance (Helton & Russell, 2013; Shaw, Funke, et al., 2013). For example, Helton and Russel (2011) demonstrated that the rate at which vigilance decrements occur during a standard vigilance task is increased with the addition of secondary working memory tasks. Conversely, the vigilance decrement decreases when the cognitive demands associated with a task are reduced (Helton & Warm, 2008; Parasuraman et al., 2009). These results are not consistent with the MART nor the underload account of vigilance decrement. Based on the underload account, the higher workload associated with additional task demands should reduce mind-wandering. However, according to the MART, additional task demands should increase the availability of cognitive resources.

Further empirical support for the resource depletion theory can be drawn from research demonstrating that a greater availability of residual cognitive resources is associated with a greater capacity for sustained attention (Matthews, Warm, Shaw, & Finomore, 2010). For instance, Helton and Warm (2008) demonstrated that increasing the salience of critical signals during a vigilance task resulted in a greater capacity to sustain attention. Presumably, increasing the salience of the signals effectively reduced the cognitive resources required to detect these signals, thereby reducing the rate at which cognitive resources were consumed (Helton & Warm, 2008). Similarly, MacLean et al. (2009) employed pre-event signals, which allowed participants to pre-empt responses, thereby reducing cognitive demands during a sustained attention task. This allowed participants to sustain attention for extended periods, presumably by reducing the rate at which cognitive resources were consumed (MacLean et al., 2009).

Based on the resource depletion account of the vigilance decrement, it might be inferred that there are two distinct components to sustained attention tasks which consume cognitive resources: the resources required to sustain attention to a task; and the resources required to engage in task-related activities (Caggiano & Parasuraman, 2004). Therefore, the rate at which cognitive resources are consumed should constitute a function of participants' effortful attention to the task, and the cognitive demands associated with task activities. Consequently, sustained attention tasks that are more engaging than monitoring tasks should consume cognitive resources at a slower rate, and therefore be associated with a relatively slower decline in performance.

Sustained attention in process control environments

Like monitoring environments, increases in mean response latency and error rates over time are also evident in process control environments (Edkins & Pollock, 1997; Hitchcock et al., 2003; Small, Wiggins, & Loveday, 2014). Process control and monitoring environments share a number of similarities. For instance, in both environments, operational demands tend to be relatively low and the tasks repetitive (Dunn & Williamson, 2011; Thackray, Bailey, & Touchstone, 1977). Consequently, performance decrements observed in process control environments may be explained by the same principles that are purported to explain the vigilance decrement in monitoring environments.

Based on a resource depletion account of the vigilance decrement, the repetitive operations required for process control tasks should consume operators' attentional resources. As with monitoring tasks, the continued consumption of cognitive resources will reduce the residual resources available to manage changes in the system state, increasing the likelihood of operator error. However, as the rate at which cognitive resources are consumed may differ in process control and monitoring environments, sustained attention may also differ based on the level of automation in the operational environment.

Process control tasks require a higher level of routine interaction by system operators in comparison to monitoring tasks (Carvalho, dos Santos, Gomes, Borges, & Guerlain, 2008; Nachreiner et al., 2006). Consequently, as process control environments necessitate more frequent engagement with the operating system, operators should exert relatively less effort to sustain their attention, reducing the rate at which their cognitive resources are consumed. Therefore, over an extended period of time, process control tasks should be associated with greater sustained attention, compared to monitoring tasks.

Chapter Two

Cue Utilisation

Delaying the onset of performance decrements in process control environments could help increase the efficiency of system operations and reduce the frequency of operational errors. Based on the resource depletion theory, reducing the cognitive load experienced by operators should delay the onset of performance decrements (Helton & Warm, 2008; Shaw, Funke, et al., 2013). From a human factors perspective, this can be achieved by changing either the environment or the operator (Hancock, Volante, & Szalma, 2016). While task design can be an effective way to reduce cognitive load, making changes to the operational environment is not always feasible. In these circumstances, interventions that are directed towards reducing cognitive load need to target individual operators. This requires an understanding of the strategies employed by operators to cope with different task demands.

Heuristic reasoning

One way that humans reduce cognitive load is through heuristics (Shah & Oppenheimer, 2008). Heuristics refer to the ‘fast and frugal’ reasoning that humans engage in familiar situations (Gigerenzer & Goldstein, 1996; Goldstein & Gigerenzer, 2002). Heuristic reasoning is argued to increase the efficiency of cognitive processing (Fishburn, 1982; Schoemaker, 1982).

Heuristics rely on associations formed in long-term memory through exposure to a situation or environment (Gigerenzer & Gaissmaier, 2011; Goldstein & Gigerenzer, 2002). The activation of heuristics is typically automatic and non-conscious, and consequently decreases cognitive load by reducing the demands placed on working memory (Brunswik, 1955; Klein, Calderwood, & Clinton-Cirocco, 1986). There are a number of different theoretical models that purport to describe the mechanisms underlying the formation and activation of heuristic associations, including the Lens Model (Brunswik, 1955), the Adaptive

Control of Thought - Rational (ACT-R) framework (Anderson, 1982), and the Recognition-Primed Decision (RPD) model (Klein, 1993, 1997).

The Lens Model

Brunswik (1955) proposed a cue-based perception model, suggesting that cognitive judgements take place through a process of cue utilisation. According to the model, cue utilisation is a process in which an operator will utilise different environmental cues to make a cognitive judgment about the true state of a situation. Brunswick's illustration of the model (see Figure 1), resembles light passing thorough a convex lens and thus, is referred to as the 'Lens Model'.

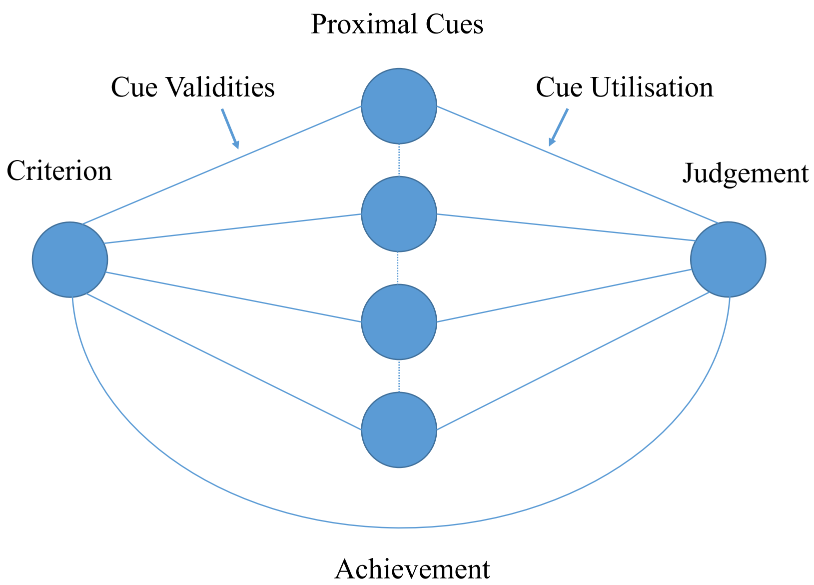

Figure 1. Human judgement modelled as the processing of imperfect sensory cues. Adapted from Brunswik (1952; 1956).

The criterion on the left-hand side of the lens (Figure 1), represents the environmental criterion or distal variable of interest (i.e., the true state of the situation). The environment also contains features that are differentially associated with the criterion, and which provide information about the criterion. Brunswick refers to these features as *proximal cues*, as they provide approximations about the true state of the situation. The judgement on the right-hand side of the lens (Figure 1), represents the judgements that people make about the criterion (i.e. their conclusions regarding the true state of the situation).

As the precise relationship between proximal cues and the criterion is not always evident, an individual's judgement regarding the true state of the situation will not necessarily be accurate (Brunswik, 1955). The correlation between the criterion and proximal cues, referred to as *cue validities* (Figure 1), describes the actual strength of the relationship between the proximal cue and the criterion. Similarly, the correlation between the proximal cues and judgement, referred to as cue utilisation, describes the relative importance (weights) of the cues to the individual making the judgement. The relative importance of cues varies between individuals, suggesting that operators may not always use the most predictive cues when making a judgement. The achievement index, represented by the overarching line between the criterion and judgement (Figure 1), reflects the accuracy with which an individual has judged the criterion.

The manner by which a judge will integrate and weigh the importance of cues to form a judgement is dependent upon the associations that have been established in memory between features and events (Balzer, Doherty, & O'Connor, 1989; Bisantz & Pritchett, 2003). For example, in the absence of a speed sign, an experienced automobile driver may take into account traffic conditions, proximity to schools, and the number of pedestrian crossings to judge the legal speed limit of a road. Although these cues may be the most predictive indicators of a speed limit as reflected in empirical data, a less experienced driver may rely on less predictive cues, such as the speed of other drivers on the road, to ascertain the speed limit.

The Lens Model has been used as a basic framework to investigate how operators use cues to make intuitive judgments regarding the system state (Bisantz & Pritchett, 2003;

Wigton, 1988; Wigton, Patil, & Hoellerich, 1986). There has been general support for the principles that underlie the Lens Model and the proposition that, during operational tasks, operators extract and utilise cues as a function of the associations formed in memory through experience between cues and feedback regarding the true system state (Balzer et al., 1989; Bisantz & Pritchett, 2003; Karelaia & Hogarth, 2008; Wigton et al., 1986). Therefore, the Lens Model suggests that operators rely on a range of idiosyncratic cues in the environment to derive meaning from uncertain situations.

Adaptive Control of Thought-Rational Model

Like the Lens Model, Anderson's (1982) Adaptive Control of Thought - Rational (ACT-R) framework also indicates that 'experience' is a key component in formulating rapid and accurate assessments of uncertain situations. Using heuristics as the basic mechanism of association, the ACT-R model proposes that the development of a cognitive skill involves two stages: a declarative stage and a procedural stage.

During the declarative stage, factual knowledge about the skill domain is acquired. Declarative knowledge constitutes knowledge that describes a rule (i.e., a power transmission control room operator knowing *that* exceeded capacity in a grid of generators requires load shedding). The procedural stage involves the transformation of this declarative knowledge into procedural rules or productions in memory (e.g., knowing *how* to safely shed load to branches of the grid). Anderson (1982) proposes that these productions are necessary to undertake skilled performance, and that the transition from novice to expert performance is characterised as a transition from a reliance on declarative knowledge to a reliance on procedural knowledge.

Productions involve the integration of information into long-term memory in the form of 'IF-THEN' rules, that combine a condition statement (IF) with an action statement (THEN; Anderson, 1987). Production rules specify that 'IF' a particular pattern of information is encountered or satisfied, 'THEN' a particular response can be executed (Anderson, 1993). For example, a control room operator may develop the rule 'IF the goal is to respond to a system error, THEN first determine the location of the error'.

With repeated exposure and practice, multiple declarative facts and productions can be collapsed into a single production (Anderson, 1993). For example, a novice control room operator may rely on several productions to respond to a system error (i.e., IF the goal is to respond to a system error, THEN determine the location of the error', IF the goal is to respond to a system error, and the location of the error has been determined, THEN determine the nature of the error, and so on). However, over time, the same task will require less declarative knowledge, as information is subsumed into the response portion of the production (i.e., IF the goal is to respond to a system error, THEN determine the location of the error, determine the nature of the error, identify immediate safety concerns, contact field operators, and so on).

The processes of consolidating declarative knowledge into procedural responses are referred to as *compilation* and *proceduralisation* (Anderson, 1982; 1993). By reducing the reliance on declarative information, these mechanisms act to decrease the demands on working memory resources. Further, the proceduralisation of responses acts to reduce response latency and increase accuracy in response to critical signals (Anderson, 1993; Brown & Carr, 1989).

The ACT-R model is based on the proposition that individuals recall chunks of information from memory by deconstructing them into subgoals, and applying knowledge from working memory to meet the demands of the situation. According to the ACT-R production-based processing account (Anderson, 1993), skill acquisition relies on the construction of goal-driven rules. The rules for specific events consist of procedural knowledge stored in long-term memory. When an individual is faced with an event that requires a response, the goal-driven rules relating to that event are automatically activated, enabling rapid and accurate responses (Anderson, 1993). Therefore, an individual's level of expertise is predicted by the extent to which his or her performance is controlled by procedural, rather than declarative knowledge.

Consistent with the predictions of the ACT-R model, inexperienced and experienced operators have distinctly different, nuanced productions in memory, which are evident through their differential use of heuristics in problem solving. For instance, inexperienced operators frequently rely on weak heuristics that are founded on domain-general knowledge (Chi, Glaser, & Rees, 1982). Experienced operators, by comparison, are likely to apply more targeted heuristics based on their domain-specific knowledge, which is associated with more efficient performance (Chi et al., 1982). Therefore, according to the ACT-R, experienced operators rely on highly developed and refined productions in memory to enable rapid assessments and responses to critical situations.

The Recognition-Primed Decision Model.

The Recognition-Primed Decision (RPD) model (Klein, 1993, 1997) is a theoretical, and largely descriptive account of how operators in highly technical, naturalistic settings rely on their experience to make rapid and effective decisions when faced with complex situations. The RPD model is derived from Klein's (1989) descriptive inquiry into experienced and novice fire commanders' decisions under ambiguity and time pressure. These fire commanders were asked to recount a specific critical incident during which they had made command decisions (Klein, Calderwood, & Clinton-Cirocco, 2010). Klein noted that, during time-critical situations, experienced commanders do not generate and evaluate a *list* of alternative decisions. Rather, they rely on their experience to identify critical patterns in the situation, and match these indicators to a previous course of action (Kaempf & Klein, 1994; Klein, 1993).

Klein posits that, through experience, operators develop a repertoire of cues that describe the causal factors in a situation (Klein, 1993, 2003). These patterns are stored in long-term memory, and provide information regarding the type of situation, including plausible goals, cues to monitor, expectancies about the situation, and typical reactions (Coderre, Mandin, Harasym, & Fick, 2003; Croskerry, 2009; Ericsson & Kintsch, 1995; Klein, 1993, 2003). Under time constraints, operators are thought to match these patterns stored in memory to the situation, allowing them to rapidly identify a course of action that has proven successful in past situations (Kaempf & Klein, 1994; Klein, 1993).

The RPD model describes how recognition-primed decision-making can occur in both simple and complex forms (Klein, 1993; Klein & Klinger, 1991). In the simplest version of the RPD model, operators will recognise a new situation as familiar by matching the pattern of cues observed in the current operational environment to an existing pattern that is stored in long-term memory (Coderre et al., 2003; Kaempf, Klein, Gary, Thordsen, & Wolf, 1996; Klein, 2003; Klein et al., 1986; Schmidt & Boshuizen, 1993; Simon, 1992; Wickens & Flach, 1988; Wickens & Hollands, 2000). For experienced operators, this recognition of a pattern of cues will activate an action script, which will include routines for responding (Klein, 1989, 2003; Lipshitz, 1993). For example, an experienced fire commander may witness smoke escaping from under the eaves of a building. From past experience with similar, critical incidences, the fire commander may match the cue 'smoke-under-eaves' with the outcome of an engulfed building collapsing, leading to the rapid decision to cease attempts to extinguish the fire, commence search and rescue operations, and seek additional resources (Klein, 1993; Klein, Orasanu, Calderwood, & Zsambok, 1993).

Klein's (1989) descriptive enquiry into fire commanders' decisions under time pressure also revealed that, while experienced commanders were able to make rapid and sound decisions, they were often unable to articulate retrospectively why they had made a particular decision. Consequently, researchers have posited that the recognition and matching of patterns of cues is a largely non-conscious process (Kahneman & Klein, Gary, 2009; Klein, 1993, 1998). This intuitive response to time critical situations, which obviates the need for extensive deliberation of the observed cues and courses of action, is generally consistent with the heuristic account of human reasoning (Gigerenzer & Gaissmaier, 2011; Gigerenzer & Goldstein, 1996; Gigerenzer, Todd, & Group, 1999; Goldstein & Gigerenzer, 2002).

The more complex form of the RPD describes recognition-primed decision-making in situations where there is not a clear match with a pattern of cues stored in memory (Klein, 2008). In these ambiguous situations, operators will not be primed to an obvious course of action, and must rely on a combination of pattern matching and more deliberative and analytic strategies to evaluate the validity of response options. Klein (2008) proposes that operators must mentally simulate a course of action to determine its potential success. The response options are then evaluated in sequence, and may be considered either: (a) appropriate (and consequently implemented); (b) inadequate and rejected; or (c) potentially adequate but requiring modification. The RPD model is based on the proposition that, even in more complex situations, operators do not have to rely on demanding analytical strategies to determine an appropriate response (Klein, 2008). Rather, they can determine the optimal response by selecting the first workable alternative, rather than considering all of the alternatives available (Klein, Gary, 1999; Simon, 1978).

The RPD model describes how, through experience in naturalistic settings, operators develop a repertoire of patterns in memory, which enables situations to be recognised as familiar, thereby facilitating sound and rapid decisions (Klein, 1993). Further, in more complex situations, operators use recognition-based processes in combination with more considered, analytical strategies to assess the validity of response options in sequence (Klein, 2008). Empirical evidence to support the RPD model has been drawn from a variety of contexts where recognition-primed judgments may be required, including commercial aviation, offshore drilling, and military control (Kaempf et al., 1996; Lipshitz, Klein, Gary, &

Orasanu, 2001). For instance, Kaempf et al. (1996) demonstrated that, during uncertain, complex and time-pressured situations, experienced naval officers relied predominantly on the process of pattern matching to classify a situation as typical, which enabled them to recognise an appropriate action based on procedural knowledge, rather than comparing multiple options.

The multiple-cue judgement framework versus cue utilisation

The Lens Model (Brunswik, 1955), the ACT-R model (Anderson, 1982), and the RPD model (Klein, 1993, 1997) all posit that pattern recognition, through the utilisation of associations between features and events, enables rapid and sound responses to critical situations. This proposition forms the basis of multiple-cue judgement theories (e.g., Brehmer, 1994; Cooksey, 1996; Hammond, Stewart, Brehmer, & Steinmann, 1975), and contemporary cue utilisation theory, as applied in diagnostic performance contexts (e.g., Loveday, Wiggins, Festa, Schell, & Twigg, 2013; Wiggins, 2012; Wiggins & O'Hare, 2003). Both the multiple-cue judgement framework and cue utilisation define cues as associations between situation-specific environmental features and objects or events, and propose that the identification of predictive features in the environment enables the automatic activation of cues stored in long-term memory (Brunswik, 1955; Klein et al., 2010; Wiggins, 2012, 2015b). However, there are also important differences in the focus and research applications of cue utilisation and the multiple-cue judgement framework.

Multiple-cue judgement framework

Multiple cue frameworks such as Social Judgement Theory (Brehmer, 1994; Cooksey, 1996; Hammond et al., 1975) investigate the processes involved in making judgments that estimate, infer, and predict the nature of unknown events. Multiple-cue approaches are based on the Lens Model (Brunswik, 1955), which provided a conceptual template for judgement processes. According to the multiple-cue judgement perspective, individuals rely on approximations of a situation (cues) to estimate the true nature of the situation (criterion). The perceived importance of these cues varies depending on the specific situation and the individual making the judgement.

The multiple-cue judgement approach typically relies on linear, additive regression models to determine the cues that individuals use to make a judgement in a specific situation, and the weights that correspond to each cues' contribution to the judgement (Stewart, 1988). This produces a 'cue-weight profile', which rates the relative importance of particular cues to the individual (Harries & Harries, 2001; Stewart, Roebber, & Bosart, 1997). For example, a cue-weight profile might describe which cues a rail controller perceives as being most important, and those cues perceived as being least important, when judging whether a train needs to be rerouted.

Multiple-cue judgement frameworks have been used to describe how cues are weighed and combined in human judgements in a variety of domains, including medicine and clinical decision-making (Harries & Harries, 2001; Smith, Gilhooly, & Walker, 2002; Wigton, 1988, 1996), educational decision-making (Cooksey, 1988; Heald, 1991), weather forecasting (Stewart et al., 1997), and sports coaching (Plessner, Schweizer, Brand, & O'Hare, 2009). Further, multiple-cue judgement approaches have been used to aid human learning. For instance, in situations where the criterion is known, feedback relating to cue validities promotes learning and improves judgement (Balzer et al., 1989; Doherty & Balzer, 1988; Gattie & Bisantz, 2006; Lagnado, Newell, Kahan, & Shanks, 2006; Plessner et al., 2009).

Despite the opportunities afforded by multiple-cue judgement research, the approach embodies a number of practical limitations. Notably, due to the automatic processes associated with cue utilisation, individuals often lack insight into the cues that are utilised during judgements, making it difficult to articulate the relative importance of specific cues (Einhorn & Hogarth, 1981; Wigton, 1996). For instance, during multiple cue tasks, participants are often inaccurate in estimating their own cue weights (Brehmer & Brehmer, 1988; Lagnado et al., 2006; Smith, Brody, & Wigton, 1986). Similarly, experts themselves are often unable to report the cues on which they are relying to make a judgement, leading to challenges in capturing the judgement policies of experts (Evans, Clibbens, Cattani, Harris, & Dennis, 2003; Klein, 1993, 1998).

The specific cues upon which individuals rely vary from person to person (Brehmer, 1994; Cooksey, 1996; Hammond & Stewart, 2001; Karlsson, Juslin, & Olsson, 2004). Further, the number of cues on which individuals rely to make judgements in specific situations also varies between individuals (Dhami & Harries, 2001; Evans et al., 2003; Feldman, 1995; Gluck, Shohamy, & Myers, 2002; Walker & Catrambone, 1993). The inconsistencies evident in the cue-weight profiles and judgements between individuals indicates that there are individual differences in the acquisition and utilisation of cues.

Individual differences in the utilisation of cues between novice and expert operators derives from experts' repeated exposure to similar situations, and their opportunity to develop a repertoire of patterns of cues that describe the causal factors in that situation (Klein, 1993, 2003). Consequently, experts, to a greater degree than novices, demonstrate the automatic, rapid and implicitly triggered decision making, which is characteristic of cue utilisation (Anderson, 1982, 1993; Dreyfus & Dreyfus, 1986; Gigerenzer & Goldstein, 1996; Klein, 1993, 1998). However, individual differences in the utilisation of cues amongst multiple experts is also evident such that the predictive validity of individual cues amongst experts tends to be idiosyncratic (Wiggins, 2012; Wigton, 1996; Yuki, Maddux, & Masuda, 2007). For instance, two experienced physicians may rely on different patterns of cues to form the same correct diagnosis (Wigton, 1996).

These individual differences limit the extent to which multiple-cue methodologies can aggregate data. The aggregation of data typically enables researchers to draw meaningful inferences or generalisations using data from multiple participants (Epstein, 1980; Horowitz, 1969). For instance, statistical models enable commonalities amongst participants to be identified, which can be used to infer descriptions and trends that can be replicated later by other researchers (King, Keohane, & Verba, 1994). However, due to qualitative differences in cue-weight profiles amongst experts, the aggregation of cue-content (i.e., the specific cues being used) can lead to discrepant or meaningless aggregated data.

An alternative means of aggregating cue-related data involves the examination of participants' behavioural patterns of information acquisition (Loveday, Wiggins, Harris, O'Hare, & Smith, 2013; Morrison, Wiggins, Bond, & Tyler, 2013; Wiggins & O'Hare, 2003a). This approach does not require direct examination of the specific cues being used, and has therefore led to a shift in cue utilisation research, with less focus on cue-content, and a greater focus on the patterns of behaviour that are indicative of the utilisation of cues.

Contemporary cue utilisation theory

Cue utilisation refers specifically to an individual's application of cue-based processing as assessed by patterns of behaviour that are indicative of the utilisation of cues (Newell & Simon, 1972; Wiggins, 2012). Cue utilisation circumvents the subjective nature of classifying features as having greater or lesser predictive validity, with the distinguishing feature of focusing on the *way* that humans acquire and utilise cues, rather than focusing on specific feature or cues themselves (Loveday, Wiggins, Harris, et al., 2013; Weiss & Shanteau, 2003; Wiggins, Azar, Hawken, Loveday, & Newman, 2014).

Cue utilisation is based on the proposition that the effective use of cues, irrespective of the specific cues used, will result in predictable patterns of behaviour (Wiggins, 2012, 2015a). For example, a skilled power distribution network controller may respond appropriately to a loss of supply in the system based on the number of substations and the type of fault on the feeder. Another operator, just as skilled, may respond in the same manner based on different features, such as the type of distribution feeder and the current temperature. While the operators are relying on different cue-based associations, they both respond in an adaptive and predictable manner.

Assessing cue utilisation

Assessments of expertise that focus on patterns of behavioural responses to cues, rather than the cues themselves, include the Cochran-Weiss-Shanteau (CWS) index of expertise (Weiss & Shanteau, 2003), and the EXPERT Intensive Skills Evaluation (EXPERTise) situation judgment test (Wiggins, Loveday, & Auton, 2015). These tools are designed to differentiate operators' level of expertise on the basis of decision-making *behaviour* in response to cues, rather than using cue content or successful judgement outcomes.

The CWS protocol requires operators to provide a judgement rating to a range of domain-relevant stimuli (Weiss & Shanteau, 2003). For example, general practitioners (GPs) may be asked to judge the probability of heart failure for various cases based on real patients, with a number of cases presented twice (Skånér, Strender, & Bring, 1998). The CWS index of expertise is calculated by dividing the participant's discrimination score (e.g., variance of ratings between patients) by their inconsistency score (e.g., variance of ratings for the same patients). Typically, individuals with higher expertise should demonstrate high discrimination and low inconsistency in their evaluations (Weiss & Shanteau, 2003).

Consistent with the CWS, the EXPERTise situation judgment test (EXPERTise 2.0; Wiggins, Loveday, & Auton, 2015) is a composite assessment of cue utilisation, incorporating a series of experimental tasks that assess distinct behaviours that characterise aspects of the utilisation of cues. The behaviours examined include the response latency to key features, the accuracy with which key features are recognised, the specificity of feature-event relationships in memory, the capacity to discriminate relevant from less relevant features during problem-solving, and the prioritisation of features during problem identification (Loveday, Wiggins, Searle, Festa, & Schell, 2013; Morrison et al., 2013; Pauley, O'Hare, & Wiggins, 2009).

The series of EXPERTise tasks can be customised and adapted to reflect the operational domain of interest (Wiggins et al., 2015). For example, the driving version of EXPERTise (Wiggins, Brouwers, Davies, & Loveday, 2014) exposes drivers to a range of driving related tasks, while the power transmission control version (Small et al., 2014) utilises the same tasks incorporating features and events relevant to power transmission. To develop domain-specific stimuli, cognitive interviews and critical incident techniques are conducted with subject matter experts to help identify domain-relevant features, event timing and sequencing. To ensure assessments of cue utilisation capture representative behavioural responses, task stimuli are then sampled from the specific operational environment of interest.

EXPERTise task scores are typically used to classify participants into cue utilisation clusters or typologies (Brouwers, Wiggins, Griffin, Helton, & O'Hare, 2017; Brouwers, Wiggins, Helton, O'Hare, & Griffin, 2016; Falkland & Wiggins, 2018; Loveday, Wiggins, Harris, et al., 2013; Small et al., 2014). Using standardised task scores, a profile is generated that differentiates participants on the basis of their behaviour in response to the task-related features (Loveday, Wiggins, Harris, et al., 2013; Wiggins, Whincup, & Auton, 2018). Where some participants may demonstrate faster response latencies to key features, they may be less capable of differentiating relevant from less relevant features during problem-solving. Nevertheless, on balance, their profile may place them in the typology that demonstrates relatively higher cue utilisation in comparison to other participants. This type of profiling reflects the fact that the acquisition of cues, as a cognitive strategy, is non-linear, and that different cognitive strategies may be engaged at different stages of skill acquisition, to facilitate cue utilisation. This method of evaluating levels of performance relative to other operators in the same domain of practice, enables assessments of cue utilisation to capture the variation within and between individual operators in judgement ratings, rather than absolute 'correctness' in decisions or feature ratings.

Eye behaviour metrics and cue utilisation

An alternative measure employed for capturing behavioural patterns in response to specific visual information is eye tracking. Using eye tracking equipment to record eye behaviour metrics, including fixations, saccades, and scan paths, researchers have investigated how individuals acquire, identify and respond to environmental information (Hoffman & Subramaniam, 1995; Mann, Williams, Ward, & Janelle, 2007; Singer, Cauraugh, Chen, Steinberg, & Frehlich, 1996; Tai, Loehr, & Brigham, 2006). For example, Tai and colleagues (2006) used gaze patterns to differentiate teachers' level of specialised expertise. Secondary school science teachers were classified as either expert or non-expert within specific branches of science based on their level of education, and afterwards completed a number of visually-aided science problems. In addition to responding more rapidly and accurately to problems within their domain of expertise, teachers also recorded fewer fixations and saccades when solving problems within their specialty. These results indicate that ‘expert’ teachers attend to fewer visual elements when solving problems within their domain of expertise, suggesting that physiological responses to cue-features may provide additional insight into elements of cue utilisation beyond accuracy-based, single-test measures.

Cue utilisation and cognitive load

The activation and retrieval of cues from long-term memory is an automatic and non-conscious process (Kahneman & Klein, Gary, 2009; Klein, 1993, 1998). Consequently, cue utilisation has the advantage of imposing relatively fewer demands on working memory resources (Chung & Byrne, 2008; Evans & Fendley, 2017). Operators with higher cue utilisation should, therefore, consume cognitive resources at a slower rate during operational tasks, compared to operators with lower cue utilisation (Brouwers et al., 2016, 2017; Small et al., 2014). Given a period of time, a lower consumption of cognitive resources will result in greater residual cognitive resources, better enabling the management of changes in the system state (Hockey, 1997; Kahneman, 1973; Wickens, 2008). Consequently, based on resource depletion theory (Helton & Warm, 2008; Parasuraman et al., 2009), operators with higher cue utilisation should be able to sustain attention for longer periods before reaching overload, resulting in greater sustained attention, compared to operators with lower cue utilisation.

The relationship between cue utilisation and the consumption of cognitive resources has typically been assessed based on inferences derived from mean response latency and accuracy during process control sustained attention tasks (Brouwers et al., 2016, 2017; Small et al., 2014). For example, Brouwers et al. (2016) observed that, during a semi-automated process control task containing an implicit pattern, the addition of a concurrent, secondary cognitive task was associated with increased response latency over time for participants with lower cue utilisation. However, there was no impact on the response latency of participants with higher cue utilisation. Brouwers et al. (2016) suggest that participants with higher cue utilisation recognised the implicit pattern and used the pattern to adopt a strategy that reduced their cognitive load during the sustained attention task, thereby conserving cognitive resources that minimised the impact of the secondary task.

Brouwers et al. (2017) provide evidence to support differences in the recognition and utilisation of predictive features based on cue utilisation, demonstrating that participants with higher cue utilisation were significantly more likely to recognise an implicit pattern in a sustained attention rail control task, which in turn was associated with significantly lower response latencies to critical signals. Presumably, by recognising the implicit pattern, individuals with higher cue utilisation were able to anticipate events by attending to the most predictive features, thereby reducing the cognitive resources required to maintain performance (Brouwers et al., 2017).

It is noteworthy that Brouwers and colleagues (2016, 2017) implemented cross-task cue utilisation, whereby cue utilisation evaluated in one context (car driving) was used to predict performance in another context (rail control). These findings are consistent with the proposition that experience alone does not account for performance expertise (Campitelli &

Gobet, 2011; Hambrick et al., 2014; Meinz & Hambrick, 2010). Further, Brouwers and colleagues' findings indicate that the development of performance expertise may be partially influenced by an individuals' general capacity for cue utilisation. However, a limitation of these studies is the use of naïve practitioners engaging in simplified representations of rail control, as differences in response latency and resource allocation based on cue utilisation likely reflected differences in the rate with which naïve participants acquired new cue-based relationships during the novel tasks. Therefore, it remains unclear whether the differences in cognitive load based on cue utilisation would also be evident amongst more experienced system operators, who differ from novice operators insofar as they have had prior opportunities to acquired cue-based associations through exposure to the operational environment.

Context-based assessments of cue utilisation, where cue utilisation and task performance are assessed in the same domain, has also differentiated performance and decision making amongst experienced personnel (Loveday, Wiggins, & Searle, 2013; Small et al., 2014; Wiggins, Azar, et al., 2014). For instance, Loveday et al. (2013) observed that higher cue utilisation amongst software engineers was associated with superior error management when developing solutions to software engineering problems. Further, engineers who exhibited relatively higher levels of cue utilisation were significantly more likely to self-report engaging in behaviours associated with expert decision making, and were more likely to be nominated as an expert by their peers, compared to those software engineers who recorded relatively lower levels of cue utilisation (Loveday, Wiggins, & Searle, 2013). Similarly, context-based assessments of cue utilisation have differentiated performance during process control tasks (Small et al., 2014; Wiggins, Azar, et al., 2014). For example, Small et al. (2014) demonstrated that, amongst qualified power control room operators, those participants with higher cue utilisation became familiar with a novel, domain-related short vigilance task more rapidly than those participants with lower cue utilisation. Further, Wiggins et al. (2014) noted that levels of cue utilisation differentiated pilots' decisions to divert or continue a flight as planned during deteriorating weather conditions.

Alternative measures of cognitive load

Evidence supporting the association between cue utilisation and the rate at which cognitive resources are consumed relies primarily on self-reports and inferences derived from mean response latencies (Brouwers et al., 2016, 2017; Small et al., 2014). Based on a resource depletion theory, a decrease in accuracy and/or an increase in response latency is assumed indicative of insufficient attentional resources necessary to maintain performance (Grier et al., 2003; Helton & Russell, 2011). Consequently, the consumption of cognitive resources can be inferred by the rate of decline in performance, with a more rapid decline indicative of greater consumption of cognitive resources. However, to accurately measure cognitive resource consumption through changes in response latency, the task must have sufficient duration and/or workload to elicit a decline in performance. Further, a decline in performance may reflect factors other than resource depletion, such as a change in the participants' response strategy, engagement or motivation (Dember, Galinsky, & Warm, 1992; Helton, Kern, & Walker, 2009).

A further limitation apparent in the cue utilisation and sustained attention literature is the use of common methodologies. Assessments of both cue utilisation and cognitive resource allocation in these studies typically rely on measures of response latency and accuracy (Brouwers et al., 2016, 2017; Small et al., 2014). Further, the data for each participant are typically collected in a single experimental session. Consequently, the relationship between cue utilisation and cognitive resource consumption may be partially attributable to common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). To overcome these potential methodological issues, complementary evidence using alternative measures of cognitive resource consumption is required.

In experimental settings, physiological measures can provide additional insight regarding the consumption of cognitive resources. Potential physiological measures of cognitive resource consumption include measures of cerebral blood flow and eye behaviour metrics (Evans & Fendley, 2017; Langner & Eickhoff, 2013; Shaw, Satterfield, Ramirez, & Finomore, 2013). Eye behaviour metrics can be used to assess participants' acquisition of visual information (Ikehara & Crosby, 2005; Poole & Ball, 2006), with patterns of eye movements providing insight into the efficiency of participants' search patterns (Henderson, 2003; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003). For instance, fixation dispersion (the extent to which fixations are distributed while completing a task) demonstrates concurrent validity with alternative measures of workload (Camilli, Terenzi, & Di Nocera, 2007; Di Nocera, Camilli, & Terenzi, 2007). Similarly, changes in eye behaviour can be used to infer changes in the consumption of cognitive resources (Ikehara & Crosby, 2005; Poole & Ball, 2006). For example, a relative increase in the frequency of visual fixations is indicative of less efficient search patterns and consequent increased cognitive demands (Ikehara & Crosby, 2005; Poole & Ball, 2006).

Functional near infrared spectroscopy (fNIRS), which measures cerebral oxygenation levels, has also been used to assess cognitive load during sustained attention tasks (Durantin, Dehais, & Delorme, 2015; Helton et al., 2010). Sustained attention tasks are associated with increased cerebral oxygenation in the right frontal lobe (Helton, Hollander, Tripp, et al., 2007; Warm, Matthews, & Parasuraman, 2009), with larger increases assumed indicative of a greater consumption of cognitive resources (Fallgatter & Strik, 1997, 1998). For instance, as the utilisation of cues is posited to reduce the demands on working memory resources, operators with higher cue utilisation should demonstrate relatively smaller increases in cerebral oxygenation during a sustained attention task in comparison to participants with lower cue utilisation.

Individual differences in cue utilisation are based on the ability of individuals to recognise and utilise patterns of cues, allowing them to attend to features with greater predictive validity, while disregarding features which have less predictive validity (Brouwers et al., 2017; Wiggins, 2015a). As operators with higher cue utilisation rapidly acquire implicit cue-based patterns (Brouwers et al., 2016, 2017; Wiggins, Brouwers, et al., 2014), during novel operational tasks these operators should demonstrate a shift in eye behaviour reflecting more efficient search patterns. Similarly, during familiar operational tasks, operators with higher levels of cue utilisation should draw on existing patterns of cues in long-term memory to facilitate more efficient search patterns, compared to operators with lower cue utilisation. Further, these more efficient search patterns should reduce cognitive load, thereby reducing cerebral oxygenation in the prefrontal cortex.

Chapter Three

Overview of Research Questions

System controllers in high risk environments, such as power control, rail control and driving, are required to sustain visual searches to monitor and control their operating systems, often for extended periods (Parasuraman & Riley, 1997; Reinerman-Jones et al., 2011). During these periods, it is critical that operators are able to recognise, anticipate, and respond rapidly and accurately to changes in the system state. However, in monitoring environments, where responses to critical signals are required infrequently, an operator's capacity to respond quickly and accurately typically declines over time (Mackworth, 1948; Parasuraman, 1979). This performance decrement is described by the resource depletion theory of vigilance decrement, which posits that sustaining attention over extended periods results in the consumption of cognitive resources, thereby reducing the residual resources available to manage changes in the system state (Grier et al., 2003; Helton & Russell, 2012; Kahneman, 1973). As the activation and retrieval of cues from long-term memory has the advantage of imposing relatively fewer demands on working memory resources (Chung & Byrne, 2008; Evans & Fendley, 2017), operators with higher cue utilisation should consume fewer resources per unit time during sustained attention tasks, resulting in greater residual cognitive resources, compared to those operators with lower cue utilisation.

The overall aim of the present programme of research was to examine whether differences in cue utilisation are associated with differences in the rate at which cognitive resources are consumed across a range of operational settings, and under a range of operational conditions, using operators of varying levels of operator experience. The following chapters of this dissertation describe six studies, each of which was designed to investigate a specific research question.

As summarised in Figure 2, Study 1 was designed to examine whether differences in cue utilisation were associated with differences in performance decrements during novel monitoring and process control sustained attention tasks. Study 2 utilised physiological measures of cognitive load to examine whether differences in performance based on cue utilisation reflected the rate at which cognitive resources were consumed. Study 3 was intended to validate a power control sustained visual search task for process control environments. Study 4 examined whether qualified operators' cue utilisation differentiated performance during a domain-relevant sustained visual search task. In Study 5, a simulated driving task was used to ascertain whether differences in cognitive load based on cue utilisation are also evident with qualified operators in more dynamic operational environments. Finally, Study 6 was designed to examine whether differences in physiological measures of cognitive load based on cue utilisation were evident during power distribution operators' regular workdays.

Study 1: Process Control versus Monitoring

- **Research Question:** Do differences in sustained attention performance based on cue utilisation differ in process control environments compared to monitoring environments?
- **How was this tested?** Participants undertook an assessment of cue utilisation and completed a novel rail control task, which required participants to reroute trains either infrequently (monitoring task) or periodically (process control task).

Study 2: Physiological Measures of Cognitive Load

- **Research Question:** Are differences in cue utilisation associated with differences in physiological measures of cognitive load during a novel task?
- **How was this tested?** Participants' frequency of visual fixations and changes in cerebral oxygenation in the prefrontal cortex were recorded during a novel rail control task.

Study 3: Sustained Visual Search

- **Research Question:** Is a short duration sustained visual search task a valid alternative to a longer-duration process control task for experimental studies?
- **How was this tested?** Performance during a newly developed sustained visual search task was compared to performance during longer-duration rail control tasks.

Study 4: Context-based Cue Utilisation

- **Research Question:** Does qualified operators' cue utilisation differentiate performance during a domain relevant sustained visual search task?
- **How was this tested?** In two experiments, power distribution operators completed an assessment of cue utilisation and completed a power control sustained visual search task.

Study 5: Dynamic Environments

- **Research Question:** Are differences in cognitive load based on cue utilisation also evident in more dynamic environments?
- **How was this tested?** Indicators of qualified drivers' cognitive resource consumption was assessed during a simulated driving task.

Study 6: In Situ

- **Research Question:** Are differences in qualified operators' cue utilisation associated with differences in physiological measures of cognitive load during familiar operational tasks?
- **How was this tested?** Across two testing sessions, eye behaviour metrics and changes in cerebral oxygenation were measured during periods of power distribution operators' regular workdays.

Figure 2. A summary of the research questions in the present programme of research

STUDY 1 & 2

Publication history

Studies 1 and 2 were published as a single paper in the *Journal of Experimental Psychology: Applied* in 2019. This paper was entitled "Cue Utilisation Differentiates Resource Allocation during Sustained Attention Simulated Rail Control Tasks". The Journal of Experimental Psychology: Applied has has an impact factor of 1.718. The author of the present dissertation wrote approximately 80% of this paper.

Sturman, D., Wiggins, M. W., Auton, J. C., & Loft, S. (2019). Cue utilisation differentiates resource allocation during sustained attention simulated rail control tasks. *Journal of Experimental Psychology: Applied*, Advance online publication. doi:10.1037/xap0000204

Cue Utilisation Differentiates Resource Allocation during Sustained Attention Simulated Rail Control Tasks

Daniel Sturman[1], Mark W. Wiggins[1], Jaime C. Auton[1], and Shayne Loft[2]

[1] *Macquarie University, NSW, Australia*

[2] *The University of Western Australia, WA, Australia*

Acknowledgements

Support for this research was provided by the Australian Research Council [ARC LP160100211].

Abstract

This study was designed to examine whether cue utilisation differentiates performance and resource allocation during simulated rail control tasks that contain implicit patterns of train movement. Two experiments were conducted, the first of which involved the completion of a 30-minute rail control simulation that required participants to reroute trains either infrequently (monitoring task) or periodically (process control task). In the monitoring condition, participants with lower cue utilisation recorded a greater increase in response latency over time. However, in the process control condition, cue utilisation failed to differentiate performance. In the second experiment, the duration of the rail control task was increased, and measures of participant fixation rates and cerebral blood flow were taken. Participants with higher cue utilisation demonstrated greater decreases in fixation rates, smaller changes in cerebral oxygenation in the prefrontal cortex, and smaller increases in response latencies, compared to participants with lower cue utilisation. The results of the study provide support for the assertion that a relatively greater capacity for cue utilisation is associated with the consumption of fewer cognitive resources during sustained attention tasks that embody an implicit pattern of activity.

Cue Utilisation Differentiates Resource Allocation during Sustained Attention Simulated Rail Control Tasks

Across a range of industrial environments, including rail control, power control, and aviation, it is critical that operators sustain attention to enable rapid and accurate responses to non-routine deviations in system states (Edkins & Pollock, 1997; Mackworth, 1948). However, sustaining attention is challenging, and the speed and accuracy of operators' responses to infrequent events typically decreases as a function of the time on a task (Helton et al., 2007; Parasuraman, 1979; Teichner, 1974). This effect is referred to as the *vigilance decrement* (Grier et al., 2003; Parasuraman, 1979).

The vigilance decrement is often observed when operators monitor automated systems (Endsley & Kaber, 1999; Molloy & Parasuraman, 1996). In fully automated environments, system operators undertake monitoring tasks where overt responses to critical signals are required infrequently (Davies & Parasuraman, 1982). Here, operators act mainly in a fail-safe capacity, monitoring the system and only intervening when problems or system failures occur (Reinerman-Jones, Matthews, Langheim, & Warm, 2011; Sheridan, 1970). However, the vigilance decrement is also evident in semi-automated environments where operators engage in process control tasks. These tasks require operators to make routine adjustments in response to more regular system deviations (Cuny, 1979; Nachreiner, Nickel, & Meyer, 2006). For example, in industrial environments such as electricity control, rail control and air traffic control, there remains a requirement for frequent human intervention to ensure that the system operates at an optimal level (Balfe, Wilson, Sharples, & Clarke, 2012; Hopkin, 2017; Navon, 2005).

Competing theories have been proposed to explain the vigilance decrement, including the overload account, the underload account, and the Malleable Attentional Resources Theory (MART). These theories are based on the proposition that the proportion of cognitive resources allocated to a stimulus determines the quantity and quality of information processing (Gopher & Donchin, 1986; Kahneman, 1973; Navon & Gopher, 1979). The underload account of vigilance decrement is based on the premise that monitoring tasks are monotonous and boring (Manly, Robertson, Galloway, & Hawkins, 1999). As a result, operators are under-stimulated (Smallwood et al., 2004). Over time, this leads to the withdrawal of cognitive resources from the primary task, resulting in an automatic, or 'mindless', approach to the task (Smallwood & Schooler, 2006). This withdrawal of cognitive resources is presumed to result in a vigilance decrement (Manly et al., 1999).

MART is also based on the proposition that less stimulating tasks will result in greater decrements in performance over time (Young & Stanton, 2002). However, MART is based on the assumption that, during underload situations, individuals' cognitive resource capacity reduces to accommodate the reduction in task demands (Young & Stanton, 2002). Therefore, MART is founded on the contention that the vigilance decrement is explained by the reduced availability of attentional resources.

Unlike the underload account and MART, the overload or 'resource depletion' account of the vigilance decrement is based on the proposition that cognitive resources are drawn from a limited source, and that the depletion of cognitive resources is the mechanism by which the vigilance decrement occurs (Grier et al., 2003; Helton & Russell, 2012; Kahneman, 1973). Cognitive resources, in the context of sustained attention, typically refers to attentional resources (Wickens, 1980; Young, Brookhuis, Wickens, & Hancock, 2015). Sustaining attention to boring or monotonous tasks requires operators to consciously allocate increased cognitive resources to the task, and the associated depletion in cognitive resources that occurs over time eventually results in fewer resources than are necessary to remain alert and respond quickly and accurately (Grier et al., 2003). In turn, this leads to a decline in performance efficiency, which becomes evident in increased response latency and/or decreased response accuracy (Helton et al., 2007; Parasuraman, 1979; Teichner, 1974).

Empirical support for the resource depletion account of vigilance decrement can be drawn from evidence to suggest that more cognitively demanding tasks are associated with a more rapid decline in performance (Helton & Russell, 2013; Shaw, Funke, et al., 2013). For instance, Helton and Russel (2011) noted that adding a working memory task to a standard vigilance task increased the rate at which the vigilance decrement occurred. By contrast, the rate of vigilance decrement is reduced when the cognitive demands associated with a task are lowered (Helton & Warm, 2008; Parasuraman et al., 2009). These findings are not consistent with the underload nor MART accounts of vigilance decrement, because the higher workload associated with adding the working memory task should have reduced mind-wandering (underload) and increased the availability of cognitive resources (MART).

Consistent with a resource depletion account of the vigilance decrement, increasing the availability of residual cognitive resources would also be expected to be associated with a greater capacity for sustained attention (Matthews, Warm, Shaw, & Finomore, 2010). Support for this proposition can be derived from MacLean et al. (2009) who employed pre-event signals, allowing participants to pre-empt responses, thereby reducing cognitive demands during a sustained attention task. This effectively reduced the rate at which cognitive resources were consumed, thereby allowing participants to sustain attention for extended periods. Similarly, Helton and Warm (2008) increased the salience of signals, effectively reducing the cognitive resources required to detect the signals, reducing the rate at which cognitive resources were consumed, and resulting in a greater capacity to sustain attention.

On the basis of Helton and Warm (2008), it might be inferred that there are two distinct components to sustained attention tasks which consume cognitive resources: the resources required to sustain attention to a task; and the resources required to engage in task-related activities (Caggiano & Parasuraman, 2004). Consequently, the rate at which cognitive resources are consumed appears to be a function of operators' effortful attention to the task, and the cognitive demands associated with task activities. Therefore, controlling for task complexity, more engaging environments should require fewer cognitive resources per unit time and be associated with a relatively slower decline in performance.

In process control and monitoring environments, operational demands can be relatively low and the tasks repetitive. Consequently, the environment tends to be perceived as 'boring' or monotonous (Dunn & Williamson, 2011; Thackray, Bailey, & Touchstone, 1977). In response, operators must exert effort to sustain their attention to the task (Scerbo, 1998). This effort is presumed to be associated with an increase in the consumption of cognitive resources (Smallwood & Schooler, 2006; Thomson, Besner, & Smilek, 2015).

The rate at which cognitive resources are consumed likely differs in process control and monitoring environments. Process control tasks require a higher level of routine interaction by system operators in comparison to monitoring tasks (Carvalho, dos Santos, Gomes, Borges, & Guerlain, 2008; Nachreiner et al., 2006). Therefore, as process control environments necessitate more frequent engagement with a system, operators should exert less effort to sustain their attention, reducing their consumption of cognitive resources. Consequently, sustained attention should be greater in process control environments compared to monitoring environments.

Reducing the rate at which cognitive resources are consumed can also be achieved in a number of other ways, including the use of decision support systems, the use of instruction, or through efficiencies in the way in which information is acquired and processed (de Jong, 2010; Kirschner, 2002; Power & Sharda, 2009). Such efficiencies are evident especially in cue-based processing, where a limited number of key features are engaged in situation assessment, thereby obviating the requirement for extensive cognitive processing (Brouwers, Wiggins, Helton, O'Hare, & Griffin, 2016; Perry, Wiggins, Childs, & Fogarty, 2012; Wiggins, 2011).

Cue utilisation and cognitive demands

Cues are recognition-driven associations between situation-specific environmental features, task events, and associated operator responses (Brunswik, 1955; Klein, Calderwood, & Clinton-Cirocco, 1986). The utilisation of cues in response to changes in the system state is associated with accurate and efficient responses, even in complex, dynamic settings (Lansdale, Underwood, & Davies, 2010; Loveday, Wiggins, Searle, Festa, & Schell, 2013). For instance, in air traffic control, the relative location and trajectory of aircraft (event) might be recognised by operators as associated with a potential loss of separation (feature) (Loft, Neal, & Humphreys, 2007). However, the capacity for cue utilisation is dependent upon the identification of predictive features in the environment, the association between features with events or objects in memory, the retention of these cue-based relationships, and the appropriate application of cues in response to environmental features (Wiggins, 2012).

The recognition of task features that underpins cue utilisation enables the direction of attention towards features of a task that embody greater relevance to required responses (Brouwers, Wiggins, Griffin, Helton, & O'Hare, 2017; Wiggins, Stevens, Howard, Henley, & O'Hare, 2002). Therefore, in response to a change in the system state, higher cue utilisation tends to be associated with the acquisition of information from a relatively limited number of specific features, together with earlier, more accurate, anticipatory responses in comparison to participants with lower cue utilisation (Weiss & Shanteau, 2003; Williams, Ward, Knowles, & Smeeton, 2002).

Domain-related cue utilisation is commonly assessed using the EXPERT Intensive Skills Evaluation (EXPERTise; Wiggins, Loveday, & Auton, 2015) situation judgment test (Brouwers et al., 2016, 2017; Small, Wiggins, & Loveday, 2014). EXPERTise is a composite assessment of cue utilisation, incorporating experimental tasks which assess distinct behaviours that characterise aspects of the utilisation of cues (Wiggins et al., 2015). The behaviours examined include the response latency to key features, the accuracy with which key features are recognised, the specificity of feature-event relationships in memory, the capacity to discriminate relevant from less relevant features during problem-solving, and the prioritisation of features during problem identification (Loveday, Wiggins, Searle, et al., 2013; Morrison, Wiggins, Bond, & Tyler, 2013; Pauley, O'Hare, & Wiggins, 2009).

Using standardised task scores, a profile is generated that differentiates participants on the basis of their behaviour in response to the task-related features (Loveday, Wiggins, Harris, O'Hare, & Smith, 2013; Wiggins, Whincup, & Auton, 2018). Where some participants may demonstrate faster response latencies to key features, they may be less capable of differentiating relevant from less relevant features during problem-solving. Nevertheless, on balance, their profile may place them in the typology that demonstrates relatively higher cue utilisation in comparison to other participants. This type of profiling reflects the fact that the acquisition of cues, as a cognitive strategy, is non-linear, and that different cognitive strategies may be engaged at different stages of skill acquisition, to facilitate cue utilisation.

Higher cue utilisation is associated with relatively lower perceived cognitive demands during sustained attention tasks (Brouwers et al., 2016; Perry et al., 2012; Wiggins, 2011). Brouwers et al. (2016) observed that, during a semi-automated process control task containing an implicit pattern, the addition of a concurrent, secondary cognitive task was associated with increased response latency over time for participants with lower cue utilisation, however had no impact on the response latency of participants with higher cue utilisation. Brouwers et al. (2016) suggest that participants with higher cue utilisation recognised the implicit pattern and used the pattern to adopt a strategy that reduced their cognitive load during the sustained attention task, thereby conserving cognitive resources that subsequently minimised the impact of the secondary task.

Based on the proposition: (a) that monitoring tasks demand greater effort than process control tasks to sustain attention; and (b) that participants with higher cue utilisation consume fewer cognitive resources than participants with lower cue utilisation during task-related interactions, an interaction should be evident whereby the difference in performance between individuals with higher versus lower cue utilisation should be greater during the more resource demanding monitoring task compared to the less resource demanding process control task.

Alternative measures of cognitive resource allocation

Experiment 1 employs participant response latency and response accuracy data to draw inferences about resource allocation as a function of cue utilisation and the type of automated task. However, the allocation of cognitive resources can also be measured in a number of alternative ways, including performance data from secondary tasks, self-rating scales and physiological measures. Researchers examining resource allocation as a function of cue utilisation have typically relied on performance measures from primary and secondary tasks (e.g., Brouwers et al., 2016, 2017; Small et al., 2014). Based on a resource depletion theory of vigilance decrement, a reduction in accuracy and/or an increase in response latency is assumed indicative of insufficient attentional resources necessary to maintain their performance (Grier et al., 2003; Helton & Russell, 2011). Consequently, the consumption of cognitive resources can be inferred by the rate of decline in performance, with a more rapid decline indicative of a greater consumption of cognitive resources during the sustained attention task.

While performance measures are an effective method to measure the consumption of cognitive resources, there are some notable limitations. First, a decline in performance may reflect factors other than resource depletion, such as a change in the participants' response strategy or motivation (Dember, Galinsky, & Warm, 1992; Helton, Kern, & Walker, 2009). Second, to accurately measure resource allocation, the task must have sufficient duration and/or workload to elicit a decline in performance.

Contemporary operational environments have typically been developed to reduce operator workload. Consequently, given the practical time constraints associated with most experimental designs, experiments replicating low workload operating environments may not have sufficient duration to elicit a decline in performance. In these experimental settings, alternative assessments of resource allocation, such as physiological measures, can provide additional insight regarding the allocation of cognitive resources. Potential physiological measures of cognitive resource allocation include measures of cerebral blood flow and eye behaviour metrics (Evans & Fendley, 2017; Langner & Eickhoff, 2013; Shaw, Satterfield, Ramirez, & Finomore, 2013), which were employed in Experiment 2 to make inferences about cognitive resource allocation.

Eye behaviour metrics can be used to assess participants' acquisition of visual information (Ikehara & Crosby, 2005; Poole & Ball, 2006). For example, larger fixation rates, defined as a greater number of brief pauses in saccadic movements per unit of time (Salvucci & Goldberg, 2000), are indicative of less efficient search patterns and consequent increased cognitive demands (Ikehara & Crosby, 2005; Poole & Ball, 2006). Pattern recognition enables the direction of attention to features of a task that embody greater relevance, reducing the number of features to which an operator needs to attend (Brouwers et al., 2017; Wiggins et al., 2002). Operating systems in industrial environments can contain both implicit and explicit patterns. For example, rail control consists of explicit patterns such as train timetables, together with implicit patterns such as the likelihood of delays based on the time of day and section of track (Jere et al., 2014). As operators with higher cue utilisation rapidly acquire implicit cue-based patterns (Brouwers et al., 2016, 2017; Wiggins, Brouwers, Davies, & Loveday, 2014), they should experience a more rapid reduction in fixation rates compared to operators with lower cue utilisation.

Functional near infrared spectroscopy (fNIRS), which measures cerebral oxygenation levels, has also been used to assess the allocation of cognitive resources during sustained attention tasks (Durantin, Dehais, & Delorme, 2015; Helton et al., 2010). Sustained attention tasks are associated with increased cerebral oxygenation in the right frontal lobe (Helton, Hollander, Tripp, et al., 2007; Warm, Matthews, & Parasuraman, 2009), with larger increases indicative of a greater allocation of cognitive resources (Fallgatter & Strik, 1997, 1998). Based on the proposition that higher cue utilisation is associated with the conservation of cognitive resources during sustained attention tasks (Brouwers et al., 2017), operators with higher cue utilisation should show relatively less cerebral oxygenation during a sustained attention task in comparison to participants with lower cue utilisation.

Experiment 1

Experiment 1 was designed to examine cognitive resource allocation as a function of cue utilisation and the type of automated task, based on measures of response latency and accuracy. A 30-minute rail control simulation that required participants to reroute trains either infrequently (monitoring task) or periodically (process control task) was employed to assess sustained attention. Rail control was selected as the domain of interest as rail control systems currently vary across the industry from process control to fully automated monitoring systems (Balfe, Sharples, & Wilson, 2015), and offers a level of ecological validity and experimental control, while allowing for cue acquisition through patterns of rail movement.

As the process control task necessitates greater engagement, participants completing the process control task should need to exert less effort to sustain their attention, reducing their consumption of cognitive resources, and resulting in greater sustained attention compared to participants completing the monitoring task. Further, as higher cue utilisation is associated with a reduction in cognitive load, and better performance during more demanding sustained attention tasks (Brouwers et al., 2016, 2017), relatively higher cue utilisation should be associated with a greater capacity to sustain performance, and this difference between levels of cue utilisation should be greater during the more resource-demanding monitoring task compared to the process control task. Ceiling effects for accuracy have been evident in previous research using low workload rail control tasks (Brouwers et al., 2017). Therefore, response latency was used as the primary outcome variable. It was hypothesised that differences in response latencies over time between participants with higher or lower cue utilisation would be greater in the monitoring task condition compared to the process control task condition.

Method

Design

Experiment 1 comprised a 2 x 2 x 4 mixed factorial design incorporating two levels of cue utilisation (higher and lower) as a between-subjects factor, two levels of condition (monitoring and process control) as a between-subjects factor, and four levels of time (Block 1, Block 2, Block 3, and Block 4) as a within-subjects factor. Participants were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of driving. Participants were randomly allocated to either the monitoring or process condition of the sustained attention task. Time constituted the four quartiles of the sustained attention task.

Participants

Sixty undergraduate university students (46 females and 14 males) were recruited from an Australian university and received course credit in return for participation. The participants ranged in age from 17 to 52 years ($M = 21.13$, $SD = 6.03$). The inclusion criteria comprised current motor vehicle drivers with a minimum of two years driving experience, and no previous exposure to rail control operations. These selection criteria helped to ensure that all participants had sufficient exposure to driving to enable comparative assessments of cue utilisation, and that assessments of resource allocation during the rail control task were not influenced by previous exposure to rail control.

Instruments

The participants were asked to indicate their age, sex, years of driving experience and hours of daily driving. Cue utilisation was assessed using the Expert Skills Evaluation (EXPERTise 2.0) software platform (Wiggins et al., 2015).

EXPERTise

EXPERTise 2.0 is designed to assess behaviour consistent with the utilisation of cues within a specific context. As participants had no previous exposure to rail control operations, the driving version of EXPERTise (Brouwers et al., 2016) was used as it assesses the acquisition of cues in a context with which participants would be familiar. Performance on the driving version of EXPERTise has been shown to predict performance in different domains, including rail control (Brouwers et al., 2017). Tasks in the EXPERTise driving battery include a Feature Identification Task (FIT), a Feature Recognition Task (FRT), a Feature Association Task (FAT), a Feature Discrimination Task (FDT), and a Feature Prioritisation Task (FPT). A summary of the EXPERTise parameters is displayed in Table 1.

The FIT requires participants to identify key features from driving scenes, and is based on the observation that experts are able to identify and utilise visual features in the environment that are more diagnostic of the system state compared to novices (Müller, Abernethy, & Farrow, 2006; Schriver, Morrow, Wickens, & Talleur, 2008). Participants are presented with 21 photographs of a road as viewed from the driver's seat of a car, and are instructed to identify the area of greatest concern in the scene (e.g., cyclists or a car pulling out from a side street) as quickly as possible using a mouse cursor. Response latency is measured as the time in milliseconds from the initial presentation of the image to the selection of an area of concern. Higher cue utilisation is associated with a lower mean response latency (Loveday, Wiggins, & Searle, 2013; Schriver et al., 2008).

The FRT consists of 18 trials, during which participants are presented with an image of a road for 500 ms. After each image has been displayed, the image is removed and participants are asked to estimate the speed limit of the road from one of four multiple choice options (50–60, 70–80, 90–100 or 110+ km/h). The FRT is designed to assess the capacity to rapidly extract key information from a driving-related scene and form an accurate judgement. Higher cue utilisation is associated with a greater number of correct judgements (Brouwers et al., 2017).

The FAT consists of 30 trials, during which participants are presented with driving related terms (e.g., ‘journey time’, ‘car speed’). For each trial, two terms are displayed, after which participants rate the extent to which they believe the two terms are related on a 6-point Likert scale ranging from 1 (*Extremely unrelated*) to 6 (*Extremely related*). As cue utilisation requires the identification of predictive feature-event relationships, higher cue utilisation is associated with greater variance in the perceived relatedness of terms (Morrison et al., 2013; Schvaneveldt, Beringer, & Lamonica, 2001).

In the FDT, participants are presented with a short, written driving scenario, and are asked to make a decision based on their typical response to the scenario. Participants are then asked to rate a list of 14 features based on their relevance to the decision using a 10-point Likert scale ranging from 1 (*Not important at all*) to 10 (*Extremely important*). Effective cue utilisation requires features to be identified as more or less relevant (Pauley et al., 2009; Weiss & Shanteau, 2003), and therefore, higher cue utilisation is associated with a greater variance in ratings (Brouwers et al., 2017).

The FPT consists of a way-finding scenario, which is accompanied by 17 drop-down lists of key features relating to the scenario. Each drop-down menu is feature-labelled (e.g., ‘Current Weather’, ‘Closest Uber Available’), and upon selection, provides participants with information pertaining to the scenario. The FPT assesses participants’ capacity to prioritise feature cue acquisition (Wiggins & O’Hare, 1995; Wiggins et al., 2002). Lower cue utilisation is more likely to be associated with the selection of information in the sequence in which it is presented (e.g., from top to bottom as they appear on the display screen), while higher cue utilisation is associated with a lower ratio of menus selected in the sequence in which they are presented (Wiggins et al., 2002).

Table 1. *Summary of EXPERTise tasks*

Task	Cognitive Skill Examined	Task Description	Measure	Validity/ Reliability
FIT	Identification of predictive features	Identify, as quickly as possible, the area of greatest concern.	Response Latency	Loveday et al. (2013) Wiggins et al. (2014)
FRT	Identification of predictive features	Estimate the speed limit in an image displayed for 500 ms.	Accuracy	Loveday et al. (2013)
FAT	Feature-event relationships in memory	Rate the comparative relationship between pairs of features and events.	Variance divided by response latency	Morrison et al. (2013)
FDT	Discrimination between predictive features	Rate the relative importance of features during a task-related problem-solving process.	Variance	Pauley et al. (2009)
FPT	Prioritisation of feature-event relationships	Acquire task-related information to solve a problem-solving process.	Ratio of sequential to non-sequential menus accessed	Wiggins & O'Hare (1995) Wiggins et al. (2002)

Rail Control Task

A simulated rail control task was used to assess participants' capacity to sustain attention (Howard, Chen, & Wiggins, 2003). The simplified rail control display depicts four separate tracks, two running from left to right and two running from right to left across a 122cm by 68cm monitor (see Figure 1). Each track contains an intersection, dividing it into two endpoints, one labeled 'odd' and the other labeled 'even'. Trains, depicted by a red line, are distributed equally across the four tracks, and progress at the same speed across the screen. Each train is labeled with a three-digit number, which is either odd (e.g., 555) or even (e.g., 444). As the trains emerge, their assigned route is indicated by a green line in advance of the train. Odd-numbered trains should be assigned to the 'odd' labeled endpoint, and even-numbered trains should be assigned to the 'even' labeled end-point.

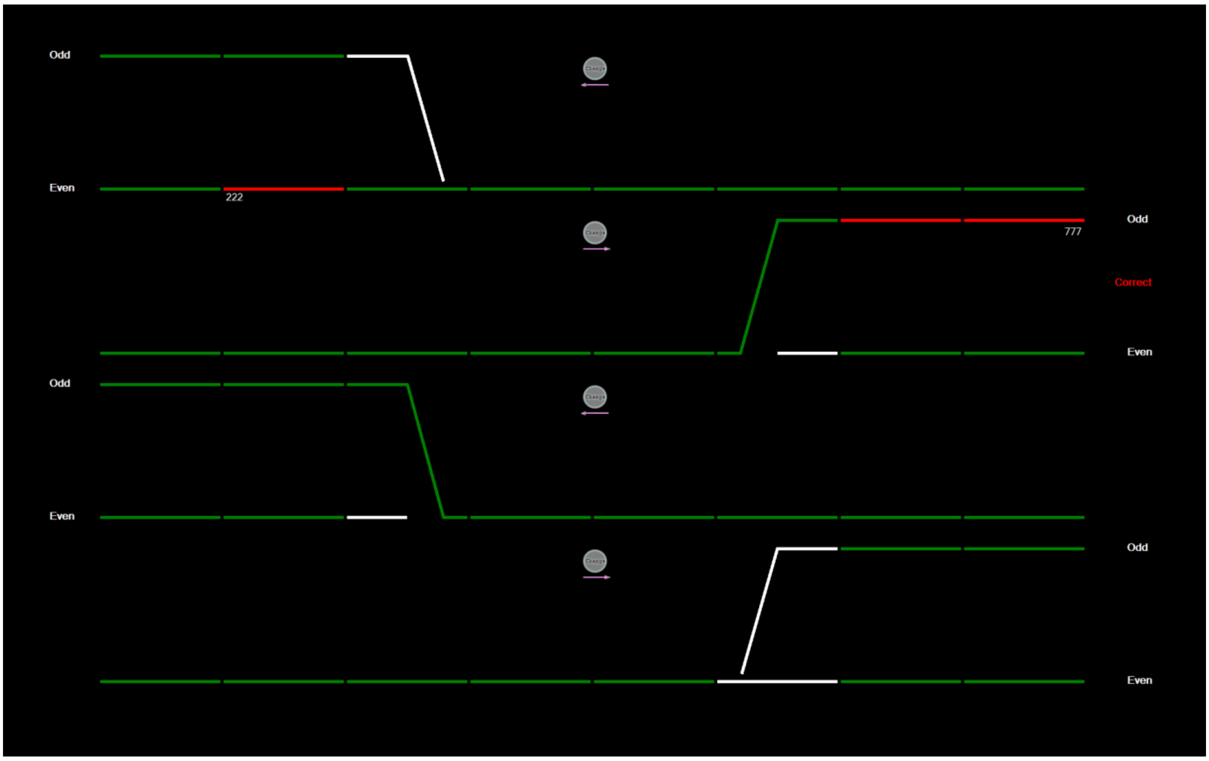

Figure 1. The graphical interface of the rail control task as it appeared to participants. A train is depicted by a red horizontal bar that appears at one end of a train line, and travels across the display. The white portions on each track are the intersection lines, which are controlled by an interlocking switch labelled, 'Change', which is depicted by a small circle icon, located above each track.

The participants are asked to ensure that the trains are assigned to the correct route. On the occasions when a train is assigned to the incorrect route, participants are instructed to use the mouse to select the grey 'change' button adjacent to the intersection as quickly as possible to reroute the train. Each train takes 11 seconds to move across the screen, and participants have seven seconds once the train appears on the screen to make a decision before the train reaches the intersection.

A total of 258 trains were presented over a 30-minute period in each condition. In the process control condition, 129 of the trains needed to be rerouted. In the monitoring condition, only four of the trains needed to be rerouted. Four trains in the process control condition were matched to the four trains that needed to be rerouted in the monitoring condition, to be used as target trains. One target train appeared in each time block, in both conditions appearing at 363, 839, 1260 and 1765 seconds. Accuracy and response latency for correct responses (time in milliseconds from appearance of the train to the selection of the 'change' icon) were recorded for each target train.

The rail task contained an implicit pattern in both the monitoring and process control conditions. In both conditions, the order in which the trains appeared was sequential, beginning with the uppermost track, followed by the second, third, fourth (lowermost), and then the first once again and so on. In the monitoring condition, the four trains that needed to be rerouted always appeared on the fourth (lowermost) track. In the process control condition, the pattern required users to divert trains on the uppermost and lowermost tracks but never the middle two tracks. In previous studies, the process control pattern was sufficiently complex to remain unrecognised by approximately two thirds of participants (Brouwers et al., 2017). Given the infrequent responses in the monitoring condition, it was necessary to create a pattern that would similarly be sufficiently discernible to a proportion of participants following one or two responses.

Cognitive ability

Set 1 of the Raven's Advanced Progressive Matrices (APM; Raven, Raven, & Court, 1998) was included as a measure of cognitive ability. The APM assesses general problem solving ability, encompassing constructs including processing speed and working memory capacity (Kaplan & Saccuzzo, 2017; Raven et al., 1998). The APM was included to establish whether cognitive ability was related to cue utilisation or performance on the rail control task.

Procedure

Following approval from the University Human Research Ethics Committee, participants were tested individually in 90-minute sessions. After reading and signing an information and consent form, participants were randomly allocated to either the process control or monitoring condition. Participants then completed an online demographic questionnaire, the five EXPERTise 2.0 tasks, the rail control task, and the APM in a counterbalanced order. Computer prompts directed participants through the demographic questionnaire and the EXPERTise tasks. Standardised instructions were provided verbally for the rail control task, with participants instructed to respond quickly, while prioritising accuracy. Following the verbal instructions, participants completed a two-minute practice trial before completing the 30-minute experimental trial. Participants were given verbal instructions for the APM before completing the paper-and-pencil version of the test.

Results

Rail task performance scores

Response latency for correct responses for the four target trains in the rail task comprised the primary dependent variable. Latencies were calculated as the number of milliseconds from the initial appearance of the train to selection of the 'change' icon. Response latencies were normally distributed for each condition and at each time-point. An error was recorded if a target train was not rerouted when required.

Cue utilisation typologies

Consistent with a standard approach to assessing cue utilisation, EXPERTise data were used to identify cue utilisation typologies that corresponded to higher or lower levels of cue utilisation (Brouwers et al., 2016; Small et al., 2014). Due to a software error, FAT scores were not recorded for 29 of the participants. Consequently, FAT scores were not included for the identification of cue utilisation typologies. Consistent with previous research (Brouwers et al., 2017; Loveday, Wiggins, Searle, et al., 2013; Wiggins et al., 2014), scores for each task were converted to *z* scores. A cluster analysis using the *z* scores was performed to identify two typologies. The first cluster, labelled the higher cue utilisation typology, consisted of participants who had a shorter response latency on the FIT, greater accuracy on the FRT, a greater variance in ratings on the FDT, and a lower ratio of sequential selections in the FPT. The second cluster, labelled the lower cue utilisation typology, consisted of participants who had a greater response latency on the FIT, lower accuracy on the FRT, a lower variance in ratings on the FDT, and a higher ratio of sequential selections in the FPT. Independent samples *t* tests demonstrated significant differences in FIT, FRT, FDT and FPT scores between the higher and lower cue utilisation typologies (see Table 2).

Table 2. *Cluster centroids for the EXPERTise task scores.*

EXPERTise	Cluster 1 (n=25)	Cluster 1 (n=35)			
Tasks	Higher cue utilisation	Lower cue utilisation	*t*	df	*p*
FIT	-.59	.42	4.44**	58	<.001
FRT	.63	-.45	4.87**	58	<.001
FDT	.34	-.24	2.30*	58	=.025
FPT	-.78	.56	6.74**	58	<.001

* Significant at the 0.05 level (two-tailed); **Significant at the 0.01 level (two-tailed)

Twenty-five participants were classified in the higher cue utilisation typology and 35 participants were classified in the lower cue utilisation typology. The ratio of participants with higher or lower levels of cue utilisation was approximately the same for each condition of the rail control task, $\chi^2(1) = 0.069$, $p = .793$ (see Table 3).

Table 3. *Number of participants with higher or lower levels of cue utilisation in the monitoring and process control conditions.*

Condition	Higher cue utilisation (N)	Lower cue utilisation (N)
Monitoring	12	18
Process control	13	17

Covariates

A series of ANCOVAs indicated that response latency during the rail control task was not associated with participants' age, scores on the APM, years of driving experience, hours of daily driving, or gender, in either the monitoring or process control conditions (*p*s > .05). Further, a series of logistic regressions indicated that accuracy during the rail control task was not associated with participants' age, scores on the APM, years of driving experience, hours of daily driving, or gender, in either the monitoring or process control conditions (*p*s > .05). Independent samples *t* tests indicated that cue utilisation was not related to participants' age, $t(58) = 1.24$, $p = .22$, scores on the APM, $t(58) = 1.40$, $p = .17$, years of driving experience, $t < 1$, or hours of daily driving, $t(58) = 1.36$, $p = .18$. A chi-square test indicated that cue utilisation was not related to participants' sex, $\chi^2(1) = 0.52$, $p = .47$. Consequently, it was not necessary to include covariates as potential explanatory variables in the main analyses.

Cue utilisation and rail task performance

A 2 x 2 x 4 Analysis of Variance (ANOVA) was conducted with condition (process control, monitoring) and cue utilisation typology (higher, lower) as between-groups variables, time (onset of target trains) as a within-groups variable (Block 1, Block 2, Block 3, and Block 4), and correct response latency as the dependent variable. A statistically significant main effect was evident for condition, $F(1,49) = 39.11$, $p < .001$, partial $\eta^2 = 0.44$. There was no statistically significant main effect for time, $F < 1$, or cue utilisation typology, $F < 1$.

There was a statistically significant interaction evident between cue utilisation and time, $F(3,147) = 3.53$, $p = .02$, partial $\eta^2 = .07$, and a statistically significant interaction between cue utilisation and condition, $F(1,49) = 7.88$, $p = .01$, partial $\eta^2 = .14$. There was no significant interaction between condition and time, $F<1$.

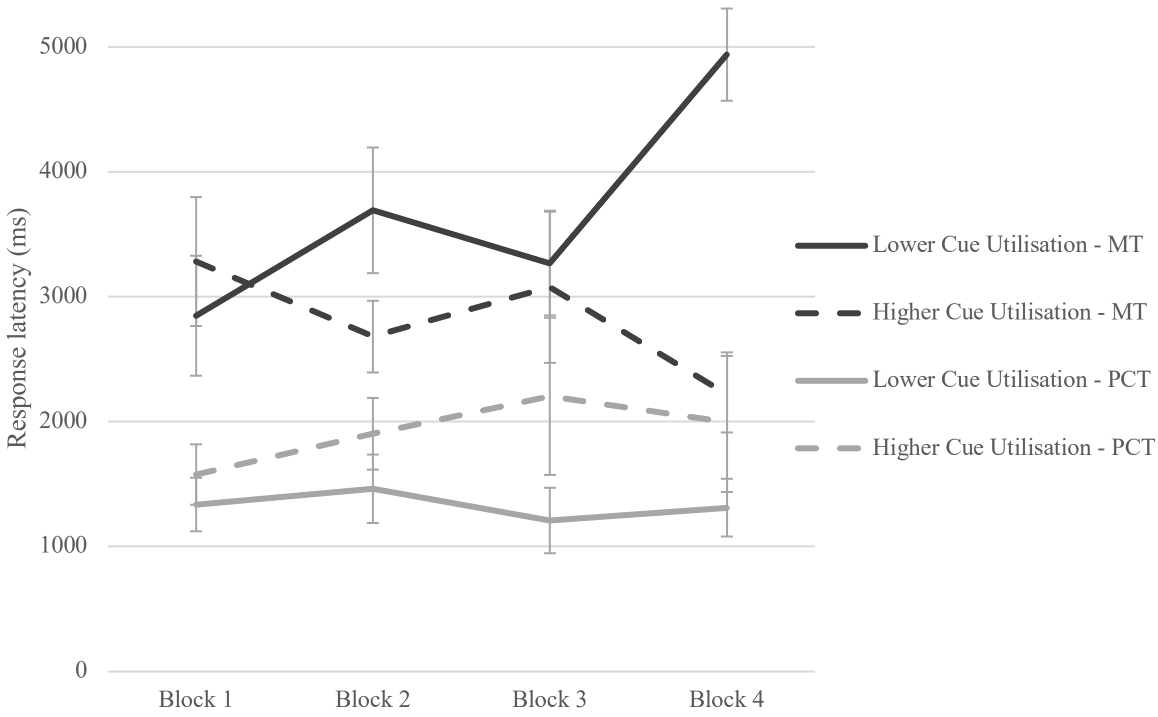

Figure 2. Mean response latencies by cue utilisation typology, condition and time. MT = monitoring task. PCT = process control task.

The main effects and the two-way interaction were qualified by a statistically significant three-way interaction between cue utilisation, condition and time, $F(3,147) = 4.42$, $p = .01$, partial $\eta^2 = .08$. Separate follow-up 2 (cue utilisation) x 4 (time) ANOVAs were conducted for the process control and monitoring conditions. For the process control condition, there was no significant interaction between cue utilisation typology and time, $F < 1$. However, for the monitoring condition, a statistically significant interaction was evident between cue utilisation condition and time, $F(3,69) = 5.54$, $p = .002$, partial $\eta^2 = .19$, with response latencies increasing more over time for the lower cue utilisation typology compared to the higher cue utilisation typology. This suggests that, compared to the higher cue utilisation typology, the lower cue utilisation typology did not experience a greater performance decrement in the process control condition, but did experience a greater performance decrement over time in the monitoring condition.

Accuracy

False alarm rates (proportion of trains that did not require re-routing, which were rerouted) were close to zero ($M = 0.01$, $SD = 0.02$). Consequently, signal detection theory was not used (Stanislaw & Todorov, 1999), and accuracy was assessed using 'correct hits' (correct re-routing of a target train). Chi-square tests compared correct hits made by each cue utilisation typology for each block of the rail control task separately for each condition (see Table 4). In both conditions, there was no significant difference in the number of hits between the higher and lower cue utilisation typologies at each of the four time points (smallest $p = .08$). This indicates that differences in mean response latencies between typologies were not the result of differences in speed-accuracy trade-off strategies.

Table 4. *Proportion of correct hits by cue utilisation typology, condition and time (standard errors in brackets)*

Condition	Cue Utilisation	Block 1	Block 2	Block 3	Block 4
Monitoring	Higher	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
	Lower	0.89 (0.08)	0.89 (0.08)	0.89 (0.08)	0.78 (0.08)
Process Control	Higher	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
	Lower	1.00 (0.00)	1.00 (0.00)	0.94 (0.06)	0.94 (0.06)

Discussion

The aim of Experiment 1 was to determine whether differences in performance over time between participants with higher or lower cue utilisation are greater in monitoring tasks compared to process control tasks. Rail control was selected as the domain of interest, and performance was assessed via response latencies at four time points. As anticipated, in the monitoring condition, participants with lower cue utilisation experienced a greater performance decrement over time than participants with higher levels of cue utilisation. However, no such differences were evident in the process control condition, suggesting that any reduction in cognitive load afforded by the utilisation of cues was not sufficiently large to differentially influence response latencies across cue utilisation typologies in the process control task.

The results from Experiment 1 indicate that decrements in performance differ as a function of the type of task and the level of cue utilisation. These findings contribute further evidence to the effect that cue utilisation allows operators to reduce cognitive demands during a sustained attention task, thereby improving response latencies to targets. However, it remains currently unclear precisely how participants with higher cue utilisation might be able to reduce their cognitive load during task-related activities.

Experiment 2

While the results from Experiment 1 provide further evidence to suggest that higher cue utilisation is associated with fewer cognitive demands during a sustained attention task (Brouwers et al., 2016; Perry et al., 2012; Wiggins, 2011), this evidence relies on inferences derived primarily from changes in response latency. To further establish that differences in response latencies, as a function of cue utilisation, reflect differences in cognitive demands, it would be useful to obtain complementary evidence using alternative measures of the manner in which individuals apply cognitive resources. If differences in response latencies, as a function of cue utilisation, reflect differences in cognitive demands, then there should be concomitant changes in physiological responses that purport to measure the allocation of cognitive resources.

Experiment 2 was designed to examine how participants with higher cue utilisation reduce the cognitive load associated with a task. Eye behaviour metrics and measures of cerebral blood flow have been used to measure resource allocation objectively during sustained attention tasks (Evans & Fendley, 2017; Langner & Eickhoff, 2013). We used these measures to establish whether participants differ in their interaction with task-related features based on their capacity for cue utilisation, and whether this in turn is associated with a reduction in cognitive demands in sustained attention tasks. Further, Experiment 2 was designed to replicate the response latency findings from Experiment 1, while extending the length of the sustained attention task to examine whether differences in performance decrements based on cue utilisation became more apparent during longer process control tasks.

Replicating Experiment 1, it was hypothesised that, during sustained attention tasks incorporating patterns of train movements, differences in response latencies over time as a function of cue utilisation would be greater in a monitoring condition compared to a process control condition. Based on the proposition that participants with higher cue utilisation rapidly identify feature-event relationships, those participants, when exposed to a novel sustained attention task with an embedded pattern, should restrict their attention to fewer relevant task features, thereby reducing the demands on cognitive resources. Therefore, it was predicted that participants with higher cue utilisation would record a relatively greater reduction in eye fixation rates over the period of the task. Further, given that the application of these more efficient search patterns should be associated with the allocation of fewer cognitive resources, it was hypothesised that participants with higher cue utilisation would show a relatively greater reduction in cerebral oxygenation over time.

Method

Design

Experiment 2 comprised a 2 x 2 x 5 mixed factorial design incorporating two levels of cue utilisation (higher and lower) as a between-subjects factor, two levels of condition (monitoring and process control) as a between-subjects factor, and five levels of time (Block 1, Block 2, Block 3, Block 4, and Block 5) as a within-subjects factor. As in Experiment 1, participants were classified post-hoc with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of driving. Participants were randomly allocated to either the monitoring or process control condition. The levels of time represent five, evenly spaced 5-minute blocks, with Block 1 running from zero to five minutes, Block 2 from 10 to 15 minutes, Block 3 from 20 to 25 minutes, Block 4 from 30 to 35 minutes, and Block 5 from 40 to 45 minutes.

Participants

Sixty-one university students (40 females and 21 males) were recruited and received course credit in return for participation. The participants ranged in age from 18 to 47 years (M = 21.2, SD = 5.5). The inclusion criteria were the same as Experiment 1.

Instruments

EXPERTise

The same five tasks from the EXPERTise 2.0 driving battery (Brouwers et al., 2016) utilised in Experiment 1 were used to assess participants' level of cue utilisation.

Rail Control Task

Participants in Experiment 2 completed the simulated train control task that was used in Experiment 1. However, in Experiment 2, a total of 390 trains were presented over a 45-minute period. In the process control condition, 195 of the trains needed to be rerouted. In the monitoring condition, only five of the trains needed to be rerouted. As in Experiment 1, five trains in the process control condition were matched to the five trains that need to be rerouted in the monitoring condition, to be used as target trains. One target train appeared in each time block, in both conditions appearing at 26, 727, 1400, 2031 and 2669 seconds. Accuracy and response latency (time in milliseconds from appearance of the train to the selection of the 'change' icon) were recorded for each target train.

The rail task incorporated an implicit pattern in both the monitoring and process control conditions. As in Experiment 1, the order in which the trains appeared was sequential, beginning with the uppermost track. In the monitoring condition, the five trains that needed to be rerouted always appeared on the fourth (lowermost) track. To ensure that the process control pattern was sufficiently complex to remain unrecognised by a proportion of participants, despite the extended running time of the rail control task in Experiment 2, a slightly more complex pattern was used in Experiment 2. In the process control condition, the trains that needed to be rerouted always appeared in the repeating pattern of [uppermost track, lowermost track, uppermost track, lowermost track, second track, third track] etc.

Eye tracking

Prior to the rail control task, participants were fitted with Tobii Eye Tracking Glasses (version 1) using the system's standard operating procedures, including a nine-point calibration. The sampling frequency of the system is 30 Hz with a 56° × 40° recording visual angle. Eye tracking data were recorded for the duration of the rail control task, and later analysed using Tobii Studio software. Fixation rates (average number of fixations made per minute) were calculated for each of the five, five-minute blocks throughout the rail control task. The Area of Interest (AOI) for fixation rates in this instance comprised the entire 122cm by 68cm rail control display on the basis that different operators may target slightly different display features to make their task decisions.

Near Infrared Spectroscopy (NIRS)

Immediately prior to the rail control task, participants were fitted with a Portalite Near Infrared Spectroscopy (NIRS) sensor (Portalite, Artinis Medical Solutions, Netherlands). The Portalite NIRS system uses light in the near-infrared spectrum (760 and 850nm) to measure cerebral activation. As oxyhemoglobin (O2Hb) and deoxyhemoglobin (HHb) have distinct optical absorption characteristics, NIRS can be used to determine the relative amounts of O2Hb and HHb in the cerebral tissue. The ratio of O2Hb to total haemoglobin (O2Hb + HHb) is used to calculate regional oxygen saturation (rSO2), which can be used as a measure of cerebral activation (Ekkekakis, 2009; Gratton & Fabiani, 2006).

The sensor was positioned approximately one centimetre above the participants' right eyebrow. The right frontal lobe was employed as vigilance tasks are associated with increased activity in this brain region (Helton, Hollander, Tripp, et al., 2007). During the five-minute baseline period, participants were asked to sit quietly, minimise body movements and to remain as relaxed as possible. rSO2 during the final minute of the baseline period was used as a baseline index. rSO2 for each of the five-minute blocks was calculated using the average O2Hb and HHb during the block.

Procedure

Participants were tested individually in 120-minute sessions. Other than the extended duration of the tasks and the use of eye tracking and NIRS during the rail control task, the procedure was identical to Experiment 1.

Results

Cue utilisation typologies

Consistent with Experiment 1, a cluster analysis was undertaken using *z* scores for all five EXPERTise tasks to identify the cue utilisation profiles that corresponded with relatively higher and lower cue utilisation. The first cluster, labelled the higher cue utilisation typology, comprised participants who had a shorter mean response latency on the FIT, greater mean accuracy on the FRT, a greater mean ratio of variance to reaction time on the FAT, a greater mean variance in ratings on the FDT, and a lower mean ratio of sequential selections in the FPT. The second cluster, labelled the lower cue utilisation typology, consisted of participants who had a greater mean response latency on the FIT, lower mean accuracy on the FRT, a lower mean ratio of variance to reaction time on the FAT, a lower mean variance in ratings on the FDT, and a higher mean ratio of sequential selections in the FPT. Independent samples *t* tests demonstrated statistically significant differences in FIT, FRT, FAT, FDT and FPT scores between the higher and lower cue utilisation typologies (see Table 5).

Table 5. *Cluster centroids for the EXPERTise task scores.*

EXPERTise	Cluster 1 (n=19)	Cluster 1 (n=42)			
Tasks	Higher cue utilisation	Lower cue utilisation	*t*	df	*p*
FIT	-0.73	0.33	-4.42**	59	<.001
FRT	0.87	-0.39	5.62**	59	<.001
FAT	0.63	-0.28	3.61**	59	.001
FDT	0.83	-0.38	5.25**	59	<.001
FPT	-0.37	0.17	-2.00*	59	.050

* Significant at the 0.05 level (two-tailed); **Significant at the 0.01 level (two-tailed)

Nineteen participants were classified in the higher cue utilisation typology and 42 participants were classified in the lower cue utilisation typology. The ratio of participants with higher or lower levels of cue utilisation was approximately the same for each condition of the rail control task, $\chi^2(1) = 0.036$, $p = .849$ (see Table 6).

Table 6. *Number of participants with higher or lower levels of cue utilisation in the monitoring and process control conditions.*

Condition	Higher cue utilisation (N)	Lower cue utilisation (N)
Monitoring	10	21
Process control	9	21

Covariates

A series of ANCOVAs indicated that response latency was not associated with participants' age, scores on the APM, years of driving experience, hours of daily driving, or gender, in either the monitoring or process control conditions (*p*s > .05). Further, a series of logistic regressions indicated that accuracy during the rail control task was not associated with participants' age, scores on the APM, years of driving experience, hours of daily driving, or gender, in either the monitoring or process control conditions (*p*s > .05). Independent samples *t* tests indicated that cue utilisation was not related to participants' age, years of driving, or hours of daily driving (all *t*s < 1). A chi-square test indicated that cue utilisation was not related to participants' sex, $\chi^2(1) = 0.07$, $p = .79$. Consequently, it was not necessary to include covariates in subsequent analyses.

Physiological measures

Relative measures of regional oxygen saturation (rSO2) compared to baseline were used to calculate rSO2 scores for each block. The mean rSO2 during each block was compared to the baseline measure taken prior to the rail control task, and scores represented the percentage change in rSO2 from baseline. Therefore, a positive score represented an increase in rSO2 compared to baseline, while a negative score represented a decrease in rSO2 compared to baseline.

Fixation rates for each of the five blocks were calculated as the mean number of eye fixations made per minute in any location on the rail control display. Due to calibration difficulties, fixation rates were unable to be accurately calculated for eight participants, six of whom were in the lower cue utilisation condition. Consequently, data for these eight participants were excluded from analyses involving fixation rates.

Cue utilisation and rail task performance

To investigate whether cue utilisation was associated with the rate of performance decrement during the sustained attention tasks, a 2 x 2 x 5 ANOVA was conducted with condition (process control vs monitoring) and cue utilisation typology (higher vs lower) as between-groups variables, time (onset of target trains) as a within-groups variable (Block 1, Block 2, Block 3, Block 4, Block 5), and correct response latency as the dependent variable. The assumption of sphericity was violated ($p < .001$). Therefore, a Greenhouse-Geisser adjustment was used for all within-subjects' effects. A statistically significant main effect was evident for condition, $F(1,49) = 20.45$, $p < .001$, partial $\eta^2 = .294$, with significantly greater response latencies in the monitoring condition ($M = 3179$, $SD = 993$), compared to the process control condition ($M = 1862$, $SD = 1141$). There was also a statistically significant main effect for time, with response latencies increasing significantly throughout the task, $F(3.1,150.8) = 5.77$, $p < .001$, $\eta^2 = .105$. There was no statistically significant main effect for cue utilisation, $F(1,49) = 2.11$, $p = .15$.

A statistically significant interaction was evident between cue utilisation and time, $F(3.1,150.8) = 2.65$, $p = .049$, partial $\eta^2 = .051$, with the lower cue utilisation typology recording a greater linear increase in response latency over time compared to the higher cue utilisation typology, $F(1,49) = 5.26$, $p = .026$, partial $\eta^2 = .097$. This indicates that, averaged over the process control and monitoring task conditions, higher cue utilisation was associated with a greater capacity to sustain attention as time on the tasks increased.

A statistically significant interaction was also evident between condition and time, $F(3.1,150.8) = 5.43$, $p = .001$, partial $\eta^2 = .100$, with the monitoring condition associated with a greater linear increase in response latency over time compared to process control condition, $F(1,49) = 7.85$, $p = .007$, partial $\eta^2 = .138$. These results are consistent with the outcomes of Experiment 1. However, in contrast to Experiment 1, there was no 3-way interaction between time, cue utilisation, and condition, $F < 1$, a point of difference between the results of the two experiments which we discuss later.

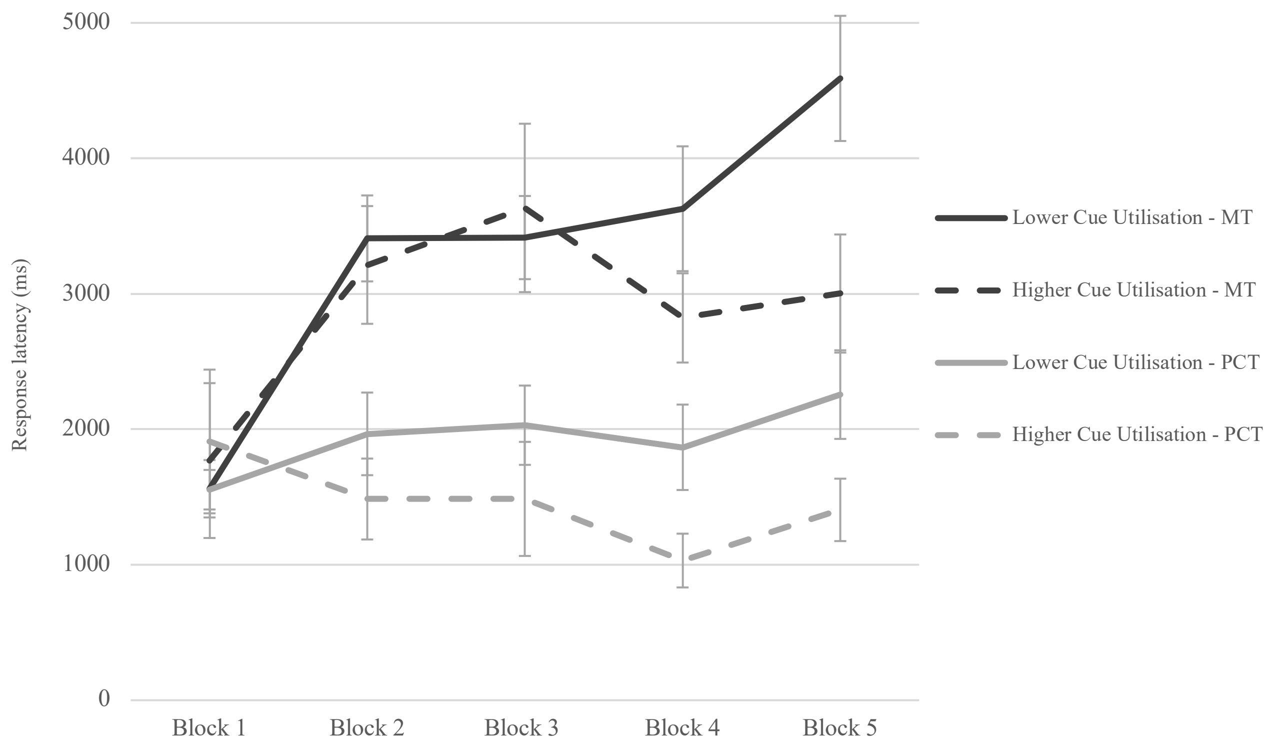

Figure 3. Mean response latencies by cue utilisation typology, condition and time. Error bars represent ± 1 SE.

Accuracy

As in Experiment 1, false alarm rates were close to zero ($M = 0.01$, $SD = 0.01$), and consequently, accuracy was assessed using 'correct hits'. Chi-square tests were employed to compare correct hits made by each cue utilisation typology for each block of the rail control task separately for each condition (see Table 7). In both conditions, there was no significant difference in the number of hits between the higher and lower cue utilisation typologies at each of the five time points (smallest $p = .09$). This indicates that differences in mean response latencies between typologies were not the result of differences in speed-accuracy trade-off strategies.

Table 7. *Proportion of correct hits by cue utilisation typology, condition and time (standard errors in brackets)*

Condition	Cue Utilisation	Block 1	Block 2	Block 3	Block 4	Block 5
Monitoring	Higher	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
	Lower	1.00 (0.00)	0.91 (0.07)	0.91 (0.07)	0.76 (0.10)	0.86 (0.08)
Process Control	Higher	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
	Lower	1.00 (0.00)	0.95 (0.05)	1.00 (0.00)	0.91 (0.07)	0.86 (0.08)

Cue utilisation and fixation rate

To investigate whether participants' information acquisition differed based on their level of cue utilisation, a 2 x 2 x 5 ANOVA was conducted with condition and cue utilisation typology as between-groups variables, time as a within-groups variable, and fixation rate as the dependent variable. The assumption of sphericity was violated (p < .001). Therefore, a Greenhouse-Geisser adjustment was used for all within-subjects' effects. A statistically significant main effect was evident for time, $F(2.4,119.2) = 16.50$, p < .001, partial $\eta^2 = .25$. There was also a statistically significant interaction between cue utilisation and time, $F(2.4,119.2) = 4.52$, $p = .01$, partial $\eta^2 = .08$. This suggests that participants with higher cue utilisation recorded a greater decrease in fixation rates over time during the sustained attention tasks compared to participants with lower cue utilisation (see Figure 4).

There was no main effect for condition, F<1, and no main effect for cue utilisation, $F(1,49) = 1.70$, $p = .20$. No statistically significant interaction was evident between time and condition, F<1, between cue utilisation and condition, F<1, or between time, cue utilisation and condition $F(2.4,119.2) = 1.64$, $p = .19$.

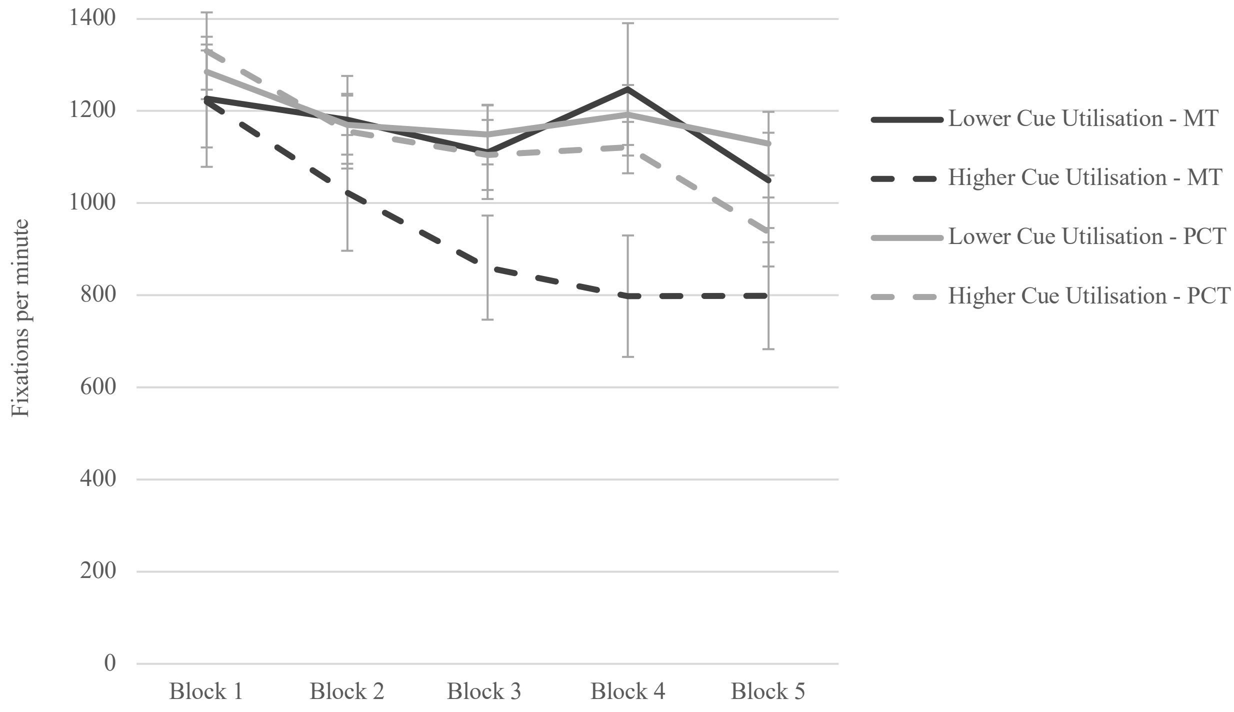

Figure 4. Mean number of eye fixations per minute by cue utilisation typology, condition and time. Error bars represent ± 1 SE.

Cue utilisation and rSO2

To evaluate further whether participants with higher cue utilisation allocated fewer cognitive resources to sustained attention tasks, a 2 x 2 x 5 ANOVA was conducted with condition and cue utilisation typology as between-groups variables, time as a within-groups variable, and rSO2 scores as the dependent variable. The assumption of sphericity was violated ($p < .001$). Therefore, a Greenhouse-Geisser adjustment was used for all within-subjects' effects. A statistically significant main effect was evident for cue utilisation, $F(1,56) = 4.23$, $p = .04$, partial $\eta^2 = .07$, with significantly greater increases in rSO2 compared to baseline in the lower cue utilisation typology ($M = 3.79$, $SD = 2.37$) compared to the higher cue utilisation typology ($M = 2.48$, $SD = 1.860$; see Figure 5).

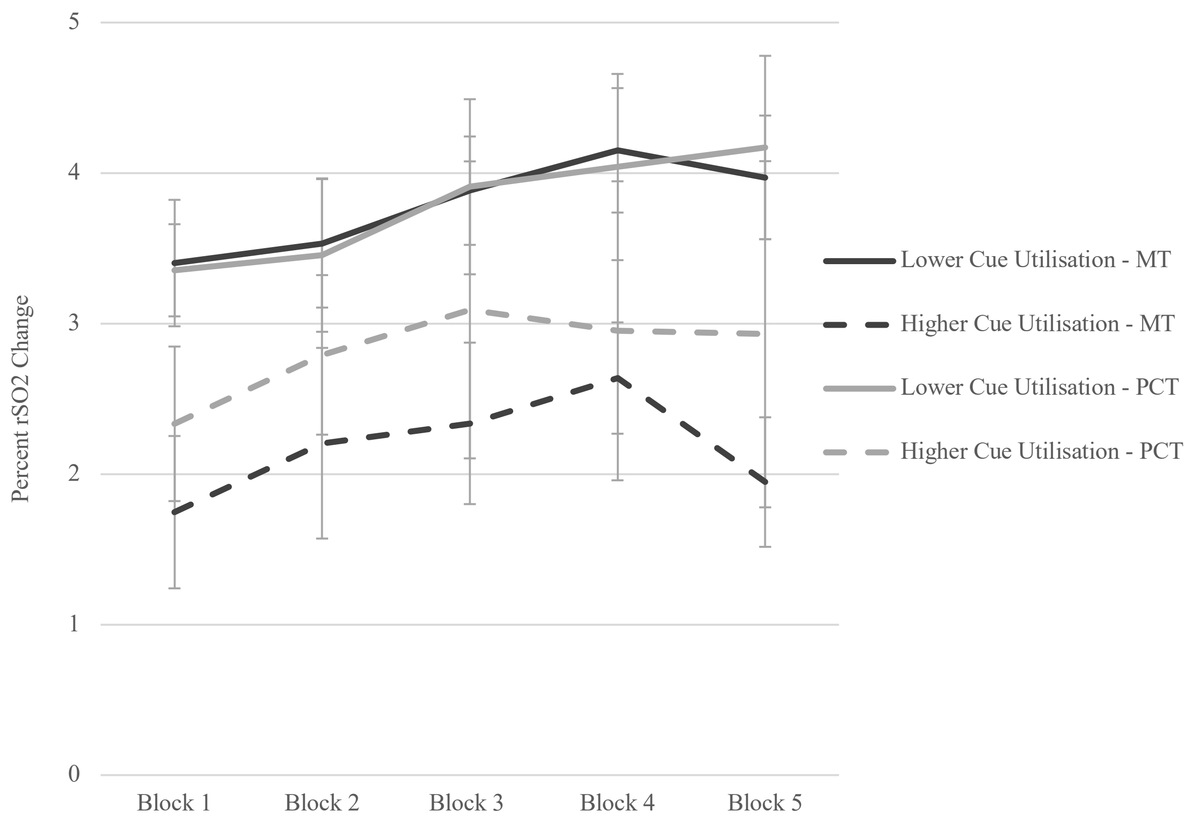

Figure 5. Mean oxygenation scores in the right hemisphere by cue utilisation typology, condition and time. Oxygenation scores are based on percent change relative to baseline. Error bars represent ± 1 SE.

There was no main effect for time, $F(2.2,121) = 2.34$, $p = .10$, nor condition, $F<1$. There was also no significant interaction between time and condition, between time and cue utilisation, between cue utilisation and condition, or between time, cue utilisation and condition (all $Fs < 1$).

Discussion

The primary aim of Experiment 2 was to determine whether participants differ in their physiological responses and rate of performance decrement during a sustained attention task, based on their level of cue utilisation. Using a 45-minute rail control task, physiological responses were assessed through changes in cerebral oxygenation compared to baseline, and changes in fixation rate throughout the task, while the capacity for sustained attention was assessed through changes in response latency over time.

Lower cue utilisation was associated with a greater increase in response latency over time compared to higher cue utilisation, but only when the data were averaged across the two rail control conditions. Therefore, in contrast to Experiment 1, the difference in response latency between participants with higher or lower cue utilisation was not greater in the monitoring condition compared to the process control condition. Inspection of Figure 2 and Figure 3 indicates that the pattern of response times for the monitoring task for higher versus lower cue utilisation typologies were quite similar across Experiments 1 and 2. In contrast, response latencies in the process control task tended to be shorter for participants with higher compared to lower cue utilisation in Experiment 2, whereas this was not the case in Experiment 1. It is not clear why this was the case, but one reason could be that participants with higher cue utilisation were better able to manage the more complex pattern of target trains presented in Experiment 2 compared to Experiment 1.

Consistent with expectations, participants with higher cue utilisation demonstrated a greater decrease in fixation rate throughout the rail control task compared to participants with lower cue utilisation. Participants with higher cue utilisation also demonstrated smaller changes in cerebral oxygenation in the prefrontal cortex, on average, compared to participants with lower cue utilisation. In combination, these results indicate that higher cue utilisation is associated with the allocation of fewer resources, and the capacity to sustain attention to a task for an extended period.

General Discussion

The empirical findings from Experiments 1 and 2 indicate that participants with higher cue utilisation allocate fewer resources to sustained attention tasks. In Experiment 1, participants with lower cue utilisation recorded greater increases in response latency during a

30-minute rail control task compared to participants with higher cue utilisation in the monitoring condition, but not in the process control condition. Using a 45-minute rail control task in Experiment 2, participants with lower cue utilisation recorded greater increases in response latency, but only when the data were averaged across both the monitoring and process control conditions. Experiment 2 extended Experiment 1 by demonstrating that participants with higher cue utilisation demonstrated smaller changes in cerebral oxygenation in the prefrontal cortex and a greater decrease in fixation rate throughout the rail control task.

Theoretical implications

The results of the present study provide support for the assertion that a greater capacity for cue utilisation is associated with the allocation of fewer cognitive resources during tasks that embody an implicit pattern of activity. The outcomes replicate previous research, which has inferred that higher cue utilisation is associated with fewer cognitive demands during a sustained attention task, by demonstrating that higher cue utilisation is associated with smaller changes in response latencies over time (Brouwers et al., 2016, 2017; Small et al., 2014). Further, Experiment 2 extended previous research, revealing that, throughout the sustained attention tasks, fixation rates were lower for participants with higher cue utilisation compared to participants with lower cue utilisation. This outcome is consistent with the proposition that higher cue utilisation is associated with the identification of implicit patterns, and that this enables attention to be directed towards task features of greater relevance, thereby reducing the overall number of features to which they attend within a given period of time (Brouwers et al., 2016; McCormack, Wiggins, Loveday, & Festa, 2014; Williams et al., 2002). Consistent with our interpretation that the application of more efficient search patterns by participants with higher cue utilisation is associated with the allocation of fewer cognitive resources, we found that higher cue utilisation was associated with a relatively greater reduction in cerebral oxygenation compared to lower cue utilisation.

At a theoretical level, the differences in response latencies between the monitoring and process control conditions are consistent with the premise that, for process control tasks, the routine interactions necessitated by the system tend to increase operator engagement. However, these results can be explained equally by the resource depletion account, the underload account, and MART. For example, the higher level of engagement in the process control condition may have lowered response latencies by reducing the likelihood of mind-wandering, as predicted by the underload account, or by reducing the cognitive resources typically required to sustain attention to a monotonous task, as predicted by the resource depletion account. Alternatively, the higher workload associated with the process control task may increase the availability of cognitive resources, as posited by MART.

Despite the explanations offered by the underload and MART accounts, the resource depletion account of the vigilance decrement offers the most complete explanation of the interactions observed between cue utilisation typologies and time in both Experiment 1 and Experiment 2. A resource theory account of vigilance decrement posits that the depletion of cognitive resources is the mechanism by which the vigilance decrement occurs (Helton & Warm, 2008; Parasuraman, Warm, & Dember, 1987). Consequently, participants who allocate greater cognitive resources to a task should also deplete their cognitive resources more rapidly, and therefore, experience a greater decline in performance (Helton & Warm, 2008; Matthews et al., 2010).

In the present study, higher cue utilisation was associated with smaller increases in response latencies over time and lower blood oxygenation, which is indicative of the allocation of fewer resources to the sustained attention tasks. This indicates that a conservation of cognitive resources (as indicated at a physiological level by the eye fixation and cerebral blood flow data) is associated with a greater capacity for sustained attention, providing support for a resource theory account of vigilance decrement.

The differences in the rate of vigilance decrement on the basis of cue utilisation typology cannot be as easily explained by the underload account nor MART. The physiological measures from Experiment 2 indicate that participants with higher levels of cue utilisation experienced lower cognitive stimulation during the sustained attention tasks. Consequently, if understimulation results in vigilance decrement, as predicted by the underload account, participants with higher cue utilisation should have experienced a greater increase in response latency over time. Similarly, if greater perceived effort is associated with greater availability of cognitive resources, as posited by MART, participants with lower cue utilisation should have experienced smaller decrements in response latency over time. As participants with lower cue utilisation experienced greater decrements in performance over time, the results of the present study tend to provide support for the resource depletion account of vigilance decrement.

Practical implications

The findings of the present studies have applied implications for the selection, training and management of operators in high risk industrial environments, such as rail control, power control and aviation, which embody implicit patterns of activity in their task environments. The outcomes extend previous research, indicating that, in addition to improving performance, higher cue utilisation is associated with the allocation of fewer resources and lower cognitive demands during sustained attention tasks. The allocation of fewer cognitive resources to sustained attention tasks should result in greater residual cognitive resources for operators with higher cue utilisation. The availability of residual cognitive resources can aid in performance, learning and the management of other (secondary) tasks (Wickens, 2002). Consequently, operators with higher cue utilisation may demonstrate greater performance during tasks involving greater cognitive demands, secondary tasks or longer vigils in high risk industrial environments.

The finding that cue utilisation moderates participants' interactions with sustained attention tasks also suggests that cue-based training may improve the performance of operators. Cue-based training could be used to change the way that information is presented to operators at particular stages of training to help them acquire, interpret, and integrate cues associated with implicit patterns in their operational environment (Scherer et al., 2008; Wiggins & O'Hare, 2003). For instance, McCammon and Hägeli (2007) developed a cue-based training system for experienced recreationists in avalanche terrain, which aids in the identification of features associated with increased risk of avalanche accidents.

Cue discovery, which involves the ad-hoc acquisition of cues through engagement with simulated operating systems (Klayman, 1988), could also be used to facilitate the acquisition of cues that can be generalised to the broader operating environment. For example, Klayman (1988) demonstrated that enabling participants to discover cues associated with a task led to subsequent improvements in their ability to predict outcomes in a computer-controlled graphic display. Finally, selecting operators using measures of cue utilisation that demonstrate predictive validity for technical training may reduce training failure.

Limitations and future direction

While the present study demonstrated that cue utilisation is differentially associated with performance decrements, cerebral oxygenation, and fixation rates during sustained attention tasks, these outcomes are restricted to naïve operators. It remains unclear whether similar patterns of results would be equally as evident amongst skilled personnel in industrial environments. Establishing whether cue utilisation predicts search patterns, cognitive load and performance amongst experienced operators is a necessary step before measures of cue utilisation can be applied more broadly. Further, as artificial patterns were incorporated within the present study, it is necessary to establish that the results from the present study can be replicated using naturally occurring, implicit patterns in industrial environments.

As cue utilisation typologies are quasi-experimental in nature, causal relationships could not be established in the present study. Consequently, the direct effect of cue utilisation on search patterns, and the direct effect of search patterns on cognitive load remain unclear. Experimental studies could be run, in which participants with lower cue utilisation are trained to identify patterns and interact with task-related features in a similar way to participants with higher cue utilisation, to determine whether this reduces cognitive load and improves performance. Establishing a causal link would aid in the development of strategies for improving the performance of operators in high risk industrial environments.

Conclusion

The current study was designed to determine whether behavioural and physiological responding during sustained attention tasks differs based on participants' level of cue utilisation. Participants with higher cue utilisation demonstrated a greater decrease in fixation rates, a smaller increase in response latencies, and a smaller increase in cerebral oxygenation throughout a sustained attention task, compared to participants with lower cue utilisation. These results indicate that, for tasks containing implicit patterns, participants with higher cue utilisation may engage more efficient search patterns, thereby lowering the cognitive demands of the task and decreasing the rate of vigilance decrement. From an applied perspective, this suggests that measures of cue utilisation may assist in the selection, training and management of operators in both monitoring and process control environments.

References (Studies 1 and 2)

Balfe, N., Sharples, S., & Wilson, J. R. (2015). Impact of automation: Measurement of performance, workload and behaviour in a complex control environment. *Applied Ergonomics*, *47*, 52–64. doi:10.1016/j.apergo.2014.08.002

Balfe, N., Wilson, J. R., Sharples, S., & Clarke, T. (2012). Development of design principles for automated systems in transport control. *Ergonomics*, *55*, 37–54. doi:10.1080/00140139.2011.636456

Brouwers, S., Wiggins, M. W., Griffin, B., Helton, W. S., & O'Hare, D. (2017). The role of cue utilisation in reducing the workload in a train control task. *Ergonomics*, *60*, 1–16. doi:10.1080/00140139.2017.1330494

Brouwers, S., Wiggins, M. W., Helton, W., O'Hare, D., & Griffin, B. (2016). Cue utilization and cognitive load in novel task performance. *Frontiers in Psychology*, *7*, 1–12. doi:10.3389/fpsyg.2016.00435

Brunswik, E. (1955). Representative Design and Probabilistic Theory in a Functional Psychology. *Psychological Review*, *62*, 193–217. doi:10.1037/h0047470

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement, *11*, 932–937. doi:10.3758/bf03196724

Carvalho, P. V. R., dos Santos, I. L., Gomes, J. O., Borges, M. R. S., & Guerlain, S. (2008). Human factors approach for evaluation and redesign of human-system interfaces of a nuclear power plant simulator. *Displays*, *29*, 273–284. doi:10.1016/j.displa.2007.08.010

Cuny, X. (1979). Different levels of analysing process control tasks. *Ergonomics*, *22*, 415–425. doi:10.1080/00140137908924626

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*, 105–134. doi:10.1007/s11251-009-9110-0

Dember, W. N., Galinsky, T. L., & Warm, J. S. (1992). The role of choice in vigilance performance. *Bulletin of the Psychonomic Society*, *30*, 201–204. doi:10.3758/BF03330441

Dunn, N., & Williamson, A. (2011). Monotony in the rail industry: The role of task demand in mitigating monotony-related effects on performance. *Ergonomics Australia–HFESA*, 1–5. doi:10.1201/b12742-87

Durantin, G., Dehais, F., & Delorme, A. (2015). Characterization of mind wandering using fNIRS. *Frontiers in Systems Neuroscience*, *9*, 1–7. doi:10.3389/fnsys.2015.00045

Edkins, G. D., & Pollock, C. M. (1997). The influence of sustained attention on railway accidents. *Accident Analysis & Prevention*, *29*, 533–539. doi:10.1016/s0001-4575(97)00033-x

Ekkekakis, P. (2009). Illuminating the Black Box: Investigating Prefrontal Cortical Hemodynamics during Exercise with Near-Infrared Spectroscopy. *Journal of Sport and Exercise Psychology*, *31*, 505–553. doi:10.1123/jsep.31.4.505

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*, 462–492. doi:10.1080/001401399185595

Evans, D. C., & Fendley, M. (2017). A multi-measure approach for connecting cognitive workload and automation. *Journal of Human Computer Studies*, *97*, 182–189. doi:10.1016/j.ijhcs.2016.05.008

Fallgatter, A. J., & Strik, W. K. (1997). Right frontal activation during the continuous performance test assessed with near-infrared spectroscopy in healthy subjects. *Neuroscience Letters*, *223*, 89–92. doi:10.1016/s0304-3940(97)13416-4

Fallgatter, A. J., & Strik, W. K. (1998). Frontal brain activation during the Wisconsin Card Sorting Test assessed with two-channel near-infrared spectroscopy. *European Archives of Psychiatry and Clinical Neuroscience*, *248*, 245–249. doi:10.1007/s004060050045

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In KR Boff, L. Kaufman, & JP Thomas (Eds.), Handbook of perception and human performance, Vol. 2. Cognitive processes and performance (pp. 1–49). John Wiley & Sons.

Gratton, G., & Fabiani, M. (2006). Optical imaging of brain function in R. Parasuraman, M. Rizzo (Eds.), Neuroergonomics: The Brain At Work, Oxford University. In *Oxford Series in Human-Technology Interactions* (pp. 65–81). Cambridge, MA.

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

Helton, W. S., Hollander, T. D., Tripp, L. D., Parsons, K., Warm, J. S., Matthews, G., & Dember, W. N. (2007). Cerebral hemodynamics and vigilance performance. *Journal of Clinical and Experimental Neuropsychology*, *29*, 545–552. doi:10.1080/13803390600814757

Helton, W. S., Hollander, T. D., Warm, J. S., Tripp, L. D., Parsons, K., Matthews, G., … Hancock, P. a. (2007). The abbreviated vigilance task and cerebral hemodynamics. *Journal of Clinical and Experimental Neuropsychology*, *29*, 545–552. doi:10.1080/13803390600814757

Helton, W. S., Kern, R. P., & Walker, D. R. (2009). Conscious thought and the sustained attention to response task. *Consciousness and Cognition*, *18*, 600–607. doi:10.1016/j.concog.2009.06.002

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*, 429–437. doi:10.1007/s00221-011-2749-1

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

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*, 429–436. doi:10.1007/s00221-012-3322-2

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*, 18–25. doi:10.1016/j.actpsy.2008.04.002

Helton, W. S., Warm, J. S., Tripp, L. D., Matthews, G., Parasuraman, R., & Hancock, P. A. (2010). Cerebral lateralization of vigilance: A function of task difficulty. *Neuropsychologia*, *48*, 1683–1688. doi:10.1016/j.neuropsychologia.2010.02.014

Hopkin, V. D. (2017). *Human factors in air traffic control*. CRC Press.

Howard, G., Chen, J., & Wiggins, M. W. (2003). Simulated Rail Control. Sydney: Unpublished.

Ikehara, C. S., & Crosby, M. E. (2005). Assessing Cognitive Load with Physiological Sensors. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, *00*, 295a–295a. doi:10.1109/HICSS.2005.103

Jere, S., Dauwels, J., Asif, M. T., Vie, N. M., Cichocki, A., & Jaillet, P. (2014). Extracting commuting patterns in railway networks through matrix decompositions. In *13th International Conference on Control Automation Robotics and Vision, ICARCV 2014* (pp. 541–546). doi:10.1109/ICARCV.2014.7064362

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ.

Kaplan, R. M., & Saccuzzo, D. P. (2017). *Psychological testing: Principles, applications, and issues*. Nelson Education.

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*, 1–10. doi:10.1016/s0959-4752(01)00014-7

Klayman, J. (1988). Cue discovery in probabilistic environments: Uncertainty and experimentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 317. doi:10.1037//0278-7393.14.2.317

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 30, pp. 576–580). doi:10.1177/154193128603000616

Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin*, *139*, 870–900. doi:10.1037/a0030694

Lansdale, M., Underwood, G., & Davies, C. (2010). Something Overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, *24*, 213–225. doi:10.1002/acp.1552

Loft, S., Neal, A., & Humphreys, M. S. (2007). The development of a general associative learning account of skill acquisition in a relative arrival-time judgment task. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 938–959. doi:10.1037/0096-1523.33.4.938

Loveday, T., Wiggins, M. W., Harris, J. M., O'Hare, D., & Smith, N. (2013). An objective approach to identifying diagnostic expertise among power system controllers. *Human Factors*, *55*, 90–107. doi:10.1177/0018720812450911

Loveday, T., Wiggins, M. W., & Searle, B. J. (2013). Cue utilization and broad indicators of workplace expertise. *Journal of Cognitive Engineering and Decision Making*, *8*, 98–113. doi:10.1177/1555343413497019

Loveday, T., Wiggins, M. W., Searle, B. J., Festa, M., & Schell, D. (2013). The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*, 125–137. doi:10.1177/0018720812448475

Mackworth, N. H. (1948). The breakdown of vigilance durning prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*, 6–21. doi:10.1080/17470214808416738

MacLean, K. A., Aichele, S. R., Bridwell, D. A., Mangun, G. R., Wojciulik, E., & Saron, C. D. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception, & Psychophysics*, *71*, 1042–1058. doi:10.3758/app.71.5.1042

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*, 661–670. doi:10.1016/S0028-3932(98)00127-4

Matthews, G., Warm, J. S., Shaw, T. H., & Finomore, V. S. (2010). A multivariate test battery for predicting vigilance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*, 1072–1076. doi:10.1177/154193121005401405

McCammon, I., & Hägeli, P. (2007). An evaluation of rule-based decision tools for travel in avalanche terrain. *Cold Regions Science and Technology*, *47*, 193–206. doi:10.1016/j.coldregions.2006.08.007

McCormack, C., Wiggins, M. W., Loveday, T., & Festa, M. (2014). Expert and competent non-expert visual cues during simulated diagnosis in intensive care. *Frontiers in Psychology*, *5*, 1–7. doi:10.3389/fpsyg.2014.00949

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *38*, 311–322. doi:10.1177/001872089606380211

Morrison, B. W., Wiggins, M. W., Bond, N. W., & Tyler, M. D. (2013). Measuring relative cue strength as a means of validating an inventory of expert offender profiling cues. *Journal of Cognitive Engineering and Decision Making*, *7*, 211–226. doi:10.1177/1555343412459192

Müller, S., Abernethy, B., & Farrow, D. (2006). How do world-class cricket batsmen anticipate a bowler's intention? *The Quarterly Journal of Experimental Psychology*, *59*, 2162–2186. doi:10.1080/02643290600576595

Nachreiner, F., Nickel, P., & Meyer, I. (2006). Human factors in process control systems: The design of human-machine interfaces. *Safety Science*, *44*, 5–26. doi:10.1016/j.ssci.2005.09.003

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*, 214–255. doi:10.1037//0033-295x.86.3.214

Navon, R. (2005). Automated project performance control of construction projects. *Automation in Construction*, *14*, 467–476. doi:10.1016/j.autcon.2004.09.006

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*, 924–927. doi:10.1126/science.472714

Parasuraman, R., de Visser, E., Clarke, E., McGarry, W. R., Hussey, E., Shaw, T. H., & Thompson, J. C. (2009). Detecting threat-related intentional actions of others: Effects of image quality, response mode, and target cuing on vigilance. *Journal of Experimental Psychology: Applied*, *15*, 275–290. doi:10.1037/a0017132

Parasuraman, R., Warm, J. S., & Dember, W. N. (1987). Vigilance: Taxonomy and utility. In

Ergonomics and Human Factors: Recent Research (pp. 11–32). doi:10.1007/978-1-4612-4756-2_2

Pauley, K., O'Hare, D., & Wiggins, M. W. (2009). Measuring expertise in weather-related aeronautical risk perception: The validity of the Cochran–Weiss–Shanteau (CWS) Index. *The International Journal of Aviation Psychology*, *19*, 201–216. doi:10.1080/10508410902979993

Perry, N. C., Wiggins, M. W., Childs, M., & Fogarty, G. (2012). Can reduced processing decision support interfaces improve the decision-making of less-experienced incident commanders? *Decision Support Systems*, *52*, 497–504. doi:10.1016/j.dss.2011.10.010

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211–219). Hershey, PA: IGI Publishing. doi:10.4018/978-1-59140-562-7.ch034

Power, D. J., & Sharda, R. (2009). Decision support systems. In *Springer handbook of automation* (pp. 1539–1548). Springer.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Section 4: The Advanced Progressive Matrices. Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Reinerman-Jones, L. E., Matthews, G., Langheim, L. K., & Warm, J. S. (2011). Selection for vigilance assignments: A review and proposed new direction. *Theoretical Issues in Ergonomics Science*, *12*, 273–296. doi:10.1080/14639221003622620

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Eye Tracking Research & Applications (ETRA) Symposium* (pp. 71–78). Palm Beach Gardens, FL: ACM.

Scerbo, M. W. (1998). What's so boring about vigilance? In R. R. Hoffman, M. F. Sherrick,

& J. S. Warm (Eds.), Viewing psychology as a whole: The integrative science of William N. Dember (pp. 145–166). Washington DC.

Scherer, R., Lee, F., Schlogl, A., Leeb, R., Bischof, H., & Pfurtscheller, G. (2008). Toward self-paced brain–computer communication: navigation through virtual worlds. *IEEE Transactions on Biomedical Engineering*, *55*, 675–682. doi:10.1109/tbme.2007.903709

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50*, 864–878. doi:10.1037/e578082012-006

Schvaneveldt, R. W., Beringer, D. B., & Lamonica, J. A. (2001). Priority and organization of information accessed by pilots in various phases of flight. *The International Journal of Aviation Psychology*, *11*, 253–280. doi:10.1037/e406222004-001

Shaw, T. H., Funke, M. E., Dillard, M., Funke, G. J., Warm, J. S., & Parasuraman, R. (2013). Event-related cerebral hemodynamics reveal target-specific resource allocation for both “go” and “no-go” response-based vigilance tasks. *Brain and Cognition*, *82*, 265–273. doi:10.1016/j.bandc.2013.05.003

Shaw, T. H., Satterfield, K., Ramirez, R., & Finomore, V. (2013). Using cerebral hemovelocity to measure workload during a spatialised auditory vigilance task in novice and experienced observers. *Ergonomics*, *56*, 1251–1263. doi:10.1080/00140139.2013.809154

Sheridan, T. B. (1970). On how often the supervisor should sample. *IEEE Transactions on Systems Science and Cybernetics*, *6*, 140–145. doi:10.1109/TSSC.1970.300289

Small, A. J., Wiggins, M. W., & Loveday, T. (2014). Cue-based processing capacity, cognitive load and the completion of simulated short-duration vigilance tasks in power transmission control. *Applied Cognitive Psychology*, *28*, 481–487. doi:10.1002/acp.3016

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., Connor, R. O., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, *13*, 657–690. doi:10.1016/j.concog.2004.06.003

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*, 946–958. doi:10.1037/0033-2909.132.6.946

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*, 137–149. doi:10.3758/bf03207704

Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors*, *16*, 339–352. doi:10.1177/001872087401600402

Thackray, R. I., Bailey, J. P., & Touchstone, R. M. (1977). Physiological, subjective, and performance correlates of reported boredom and monotony while performing a simulated radar control task. In *Vigilance* (pp. 203–215). Springer. doi:10.1007/978-1-4684-2529-1_12

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*, 82–96. doi:10.1177/1745691614556681

Warm, J. S., Matthews, G., & Parasuraman, R. (2009). Cerebral Hemodynamics and Vigilance Performance. *Military Psychology*, *21*, 75–100. doi:10.1080/08995600802554706

Weiss, D. J., & Shanteau, J. (2003). Empirical assessment of expertise. *Human Factors*, *45*, 104–116. doi:10.1518/hfes.45.1.104.27233

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickers (Ed.), *Attention* *and Performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159–177. doi:10.1080/14639220210123806

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology*, *25*, 229–235. doi:10.1002/acp.1668

Wiggins, M. W. (2012). The role of cue utilisation and adaptive interface design in the management of skilled performance in operations control. *Theoretical Issues in Ergonomics Science*, *15*, 283–292. doi:10.1080/1463922X.2012.724725

Wiggins, M. W., Brouwers, S., Davies, J., & Loveday, T. (2014). Trait-based cue utilization and initial skill acquisition: Implications for models of the progression to expertise. *Frontiers in Psychology*, *5*, 1–8. doi:10.3389/fpsyg.2014.00541

Wiggins, M. W., Loveday, T., & Auton, J. C. (2015). *EXPERT Intensive Skills Evaluation (EXPERTise) Test*. Sydney: Macquarie University.

Wiggins, M. W., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, *1*, 305–320. doi:10.1037/1076-898X.1.4.305

Wiggins, M. W., & O'Hare, D. (2003). Weatherwise: Evaluation of a Cue-Based Training Approach for the Recognition of Deteriorating Weather Conditions during Flight. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*, 337–345. doi:10.1518/hfes.45.2.337.27246

Wiggins, M. W., Stevens, C., Howard, A., Henley, I., & O'Hare, D. (2002). Expert, intermediate and novice performance during simulated pre-flight decision-making. *Australian Journal of Psychology*, *54*, 162–167. doi:10.1080/00049530412331312744

Wiggins, M. W., Whincup, E., & Auton, J. C. (2018). Cue utilisation reduces effort but increases arousal during a process control task. *Applied Ergonomics*, *69*, 120–127. doi:10.1016/j.apergo.2018.01.012

Williams, A. M., Ward, P., Knowles, J. M., & Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied*, *8*, 259–270. doi:10.1037/1076-898X.8.4.259

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics*, *58*, 1–17. doi:10.1080/00140139.2014.956151

Young, M. S., & Stanton, N. A. (2002). Malleable Attentional Resources Theory: A new explanation for the effects of mental underload on performance. *Human Factors*, *44*, 365–375. doi:10.1518/0018720024497709

Bridging Section 1

Studies 1 and 2 were designed to establish whether a general capacity for cue utilisation predicts differences in the consumption of cognitive resources during novel sustained attention tasks. Across both the monitoring and process control conditions, participants with higher cue utilisation recorded smaller increases in response latencies and cerebral oxygenation, and smaller decreases in fixation rates, compared to those with lower cue utilisation. These findings suggest that individuals with a greater capacity for cue utilisation had adopted a strategy that reduced the rate at which cognitive resources were consumed. However, Studies 1 and 2 relied on cross-task cue utilisation, whereby cue utilisation in the context of driving was used to predict performance during a novel rail control task. Therefore, these studies involved naïve practitioners with no prior exposure to rail control. Consequently, the observed differences in response latency and cognitive load based on cue utilisation likely reflected differences in the speed with which participants acquired new cue-based relationships during the rail control task.

Amongst experienced practitioners, who differ from novice operators in that they have had prior opportunities to acquire cue-based associations, higher cue utilisation is purported to be associated with a greater frequency of previously acquired cues (McCormack et al., 2014; Morrison et al., 2013). Therefore, there is a need to establish whether similar effects to those observed in Studies 1 and 2 are demonstrated amongst more experienced practitioners in an operational context. However, due to time constraints and costs, the availability of experienced practitioners in an operational context is time restricted. Consequently, the utility of longer-duration sustained attention tasks for studies involving experienced operators is limited. Nevertheless, Posner (1978) has argued that the duration of sustained attention tasks is arbitrary, and that shorter-duration tasks should demonstrate effects similar to those achieved with longer-duration sustained attention tasks.

Study 3 was designed to validate an adapted short-duration sustained visual search task for process control environments. To achieve this, participants were recruited to complete the sustained visual search task, in addition to a longer-duration process control or monitoring task. The sustained visual search task required participants to monitor simulated operating power transmission interfaces, and identify system faults. It was hypothesised that performance during the sustained visual search task would be positively associated with performance during the longer-duration process-control and monitoring tasks. The domain of power control was selected to enable future assessments of sustained attention amongst experienced power control room operators, where participants' availability is likely to be limited.

STUDY 3

Publication history

The paper arising from Study 3 was entitled, “The Development and Validation of a Short-Duration Sustained Visual Search Task for Process Control Environments”. A version of this paper has been submitted to the *Applied Ergonomics* Journal, and is currently under review. The author of the present dissertation wrote approximately 80% of this paper.

The Development and Validation of a Short-Duration Sustained Visual Search Task for Process Control Environments

Daniel Sturman[1], Mark W. Wiggins[1], William S. Helton[2], and Jaime C. Auton[1]

[1] *Macquarie University, NSW, Australia*

[2] George Mason University, VA, United States

Acknowledgements

Support for this research was provided by the Australian Research Council [ARC LP160100211].

Abstract

This study was designed to validate a newly adapted, short-duration sustained visual search task for process control environments. The sustained visual search task consists of 10 short dynamic scenarios, which require participants to monitor simulated operating power transmission interfaces, and identify system faults. A vigilance decrement was demonstrated throughout the sustained visual search task, as evident in increased response latencies and decreased accuracy over time. Further, changes in response latency throughout the sustained visual search task were positively associated with changes in response latency during a 30-minute, low signal probability rail control task, a 45-minute low signal probability rail control task, and a 45-minute high signal probability rail control task. The findings indicate that the sustained visual search task may be a valid alternative to a longer-duration process control task for experimental studies. The sustained visual search task is likely to be of value for research where there are time constraints, which may occur during field research, or when participants' availability is limited.

The Development and Validation of a Short-Duration Sustained Visual Search Task for Process Control Environments

Many industrial environments are transitioning to automated or semi-automated systems which are designed to reduce operators' workload, while increasing reliability and accuracy (Anders, Seijmonsbergen, & Bouten, 2011; Lin, Yenn, & Yang, 2010; Metzger & Polakow, 2011; Navon, 2005). In fully-automated environments, operators intervene rarely, with technological systems sufficiently mature to manage the vast majority of changes to the system state (Reinerman-Jones, Matthews, Langheim, & Warm, 2011; Sheridan, 1970). In semi-automated environments, operator intervention tends to be more frequent due to unreliable or less mature support systems (Dimeas & Hatziargyriou, 2005; Navon, 2005; Shaw et al., 2012). To ensure that they maintain control of an otherwise dynamic system, operators within semi-automated environments tend to engage in process-driven behaviour, making routine adjustments in response to regular system deviations (Grier et al., 2003; Gupta & Chow, 2010).

One of the challenges for operators involved in the management of either fully automated or process-control systems concerns the requirement to remain vigilant in an otherwise minimally engaging and/or repetitive environment. Vigilance, or sustained attention, refers to the ability of operators to remain alert and focused, and respond rapidly and accurately to deviations in the system state (Langner & Eickhoff, 2013). Over a period of watch, the capacity for operators to sustain attention typically declines (Davies & Parasuraman, 1982; Mackworth, 1948). This phenomenon is referred to as the vigilance decrement (Parasuraman, 1979).

The vigilance decrement has been observed in both novice and experienced operators in laboratory experiments using tasks that typically run for a minimum duration of 30 minutes (Grier et al., 2003; Shaw, Satterfield, Ramirez, & Finomore, 2013; Warm, Parasuraman, &

Matthews, 2008; Wiggins, 2011). During these tasks, participants are required to respond to infrequent target signals, while disregarding more frequent non-target or distractor signals (Davies & Parasuraman, 1982; Molloy & Parasuraman, 1996; Parasuraman, 1979). These vigilance tasks have enabled the assessment of factors that may be associated with rates of decline in sustained attention, including fatigue, workload, stress and individual differences (Brouwers, Wiggins, Helton, O'Hare, & Griffin, 2016; Matthews, Warm, Reinerman-Jones, et al., 2010; Matthews, Warm, Shaw, & Finomore, 2010; Temple et al., 2000).

Long duration vigilance tasks have been used to assess the sustained attention of a range of operators including pilots (Wiggins, 2011), certified x-ray inspectors (Ghylin, Drury, Batta, & Lin, 2007), and radar surveillance operators (Pigeau, Angus, O'Neill, & Mack, 1995). However, the utility of long duration vigilance tasks is limited in operational environments where time constraints and costs limit the availability of operators for extended periods.

Posner (1978) has argued that the duration of vigilance tasks is arbitrary, and that shorter vigilance tasks should demonstrate effects similar to those achieved with longer vigilance tasks. Evidence to support this proposition has been demonstrated using the continuous performance task (Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956), and the Temple task (Temple et al., 2000). The latter is a 12-minute task, during which Temple et al. (2000) observed vigilance decrements similar to those recorded during longer vigils. Furthermore, the Temple task reproduced the higher workload and stressful character associated with longer vigilance tasks (Temple et al., 2000).

Like long duration vigilance tasks, vigilance tasks over shorter periods are designed primarily for use within fully-automated systems, where critical signal probability is low (Davies & Parasuraman, 1982; Matthews, Warm, Shaw, et al., 2010; Nuechterlein, Parasuraman, & Jiang, 1983; Temple et al., 2000). However, they do not necessarily capture sustained attention in process control systems, which require routine operator intervention as a consequence of higher critical signal probability. Consequently, there is a need to develop and validate a short-duration vigilance task that is capable of assessing changes in sustained attention in process control environments.

Development of a short-duration sustained visual search task

The attentional resource theory of vigilance decrement posits that sustaining attention is cognitively demanding, and that the vigilance decrement results from the depletion of cognitive resources over the watch (Parasuraman, Warm, & Dember, 1987). Empirical support for attentional resource theory can be drawn from research demonstrating that more cognitively demanding tasks result in a more rapid decline in performance (Helton & Russell, 2013; Shaw, Funke, et al., 2013). Consequently, by increasing cognitive demands, previously developed, short-duration vigilance tasks have demonstrated decrements in performance of a similar magnitude to those observed during longer vigilance tasks (Matthews, Warm, Shaw, et al., 2010; Nuechterlein et al., 1983; Temple et al., 2000).

The rate of decline in performance during a process control task is typically lower than that observed during a traditional monitoring task (Brouwers et al., 2016; Sturman et al., 2019). Sturman et al. posited that the higher level of routine interaction required for process control tasks compared to monitoring tasks, necessitates more frequent engagement with a system, thereby reducing the resources required to sustain attention. Consequently, during process control tasks, cognitive resources are consumed at a reduced rate, thereby extending the period over which attention can be sustained (Sturman et al., 2019).

To assess the construct validity of a short-duration process control task, it is necessary for the task to demonstrate a decline in performance, consistent with a resource theory account of the vigilance decrement (Myers, Well, & Lorch, 2010; Parasuraman et al., 1987). To accelerate this decline, the imposition of cognitive demands is necessary to consume cognitive resources at a rate greater than would normally be the case (Matthews, Warm, Shaw, et al., 2010; Temple et al., 2000). However, this acceleration in the rate of decline should not occur at the expense of variability in performance, since an overly complex task risks confounding sustained attention with cognitive ability.

Increasing the variability in the rate at which cognitive resources are consumed can be achieved by providing participants with the opportunity to reduce their cognitive load during the task (Brouwers et al., 2016). Process control tasks that incorporate implicit patterns allow a proportion of participants to discern the pattern and use strategies to reduce the cognitive demands of a task (Brouwers et al., 2016; Sturman et al., 2019). A similar finding has been discovered in low signal probability vigilance tasks, where signal regularity, a temporal pattern in occurrence of critical signals, improves performance (Helton et al., 2005).

The aim of the present study was to establish the validity of a short-duration sustained visual search task for process control environments. The sustained visual search task consists of a power control simulation, during which participants are routinely required to identify failed circuit breakers. Performance is calculated as the change in mean response latency and accuracy during the sustained visual search task. The construct validity of the sustained visual search task was assessed by examining whether performance declined over the period of watch, consistent with the vigilance decrement. Convergent validity was assessed by examining whether changes in performance during the sustained visual search task correlated with changes in performance during a longer-duration monitoring task, and changes in performance during a longer-duration process control task. A rail control simulation, that required participants to reroute trains either infrequently (monitoring task) or periodically (process control task), was employed as the longer-duration vigilance task. It was hypothesised that: (a) mean response latencies would increase significantly over time during the sustained visual search task, (b) accuracy would decrease significantly over time during the sustained visual search task, (c) performance during the sustained visual search task would be positively associated with performance during the longer-duration monitoring task, and that (d) performance during the sustained visual search task would be positively associated with performance during the longer-duration process control task.

Method

Participants

One hundred twenty-one undergraduate university students (86 females and 35 males) were recruited and received course credit in return for participation. The participants ranged in age from 17 to 52 years (*M* = 21.18, *SD* = 5.73). The inclusion criteria required no previous exposure to power control or rail control operations.

Short-duration vigilance task

During the sustained visual search task, participants viewed a series of 10 short, dynamic scenarios. The stimuli were adapted from a vigilance task developed by Small et al. (2014). Each scenario displayed a Supervisory Control and Data Acquisition (SCADA) interface consisting of an operating power transmission grid (see Figure 1). Each scenario displayed a different transmission grid. No experience with power control operations was necessary for the completion of the task.

Written instructions explained to participants that a fault (a failed circuit breaker), represented by a flashing red square, would appear during each scenario. Participants were instructed to select the fault, as quickly as possible, using the mouse. A failed circuit breaker would always appear on a section of the grid 30 seconds from the onset of each scenario and remain visible until the operator selected the fault. Importantly, participants were not informed that faults would appear 30 seconds after the beginning of each scenario. Response latencies were calculated as the time, in milliseconds, from the initial appearance of a fault to the correct selection of the circuit breaker. The sustained visual search task continued for approximately seven minutes, with the total duration of the vigil dependent upon the response latency of the operator.

Response latency and accuracy were recorded for each individual across each of 10 experimental scenarios that comprised the sustained visual search task. Accuracies were recorded as correct if responses were less than 15 pixels from the fault location.

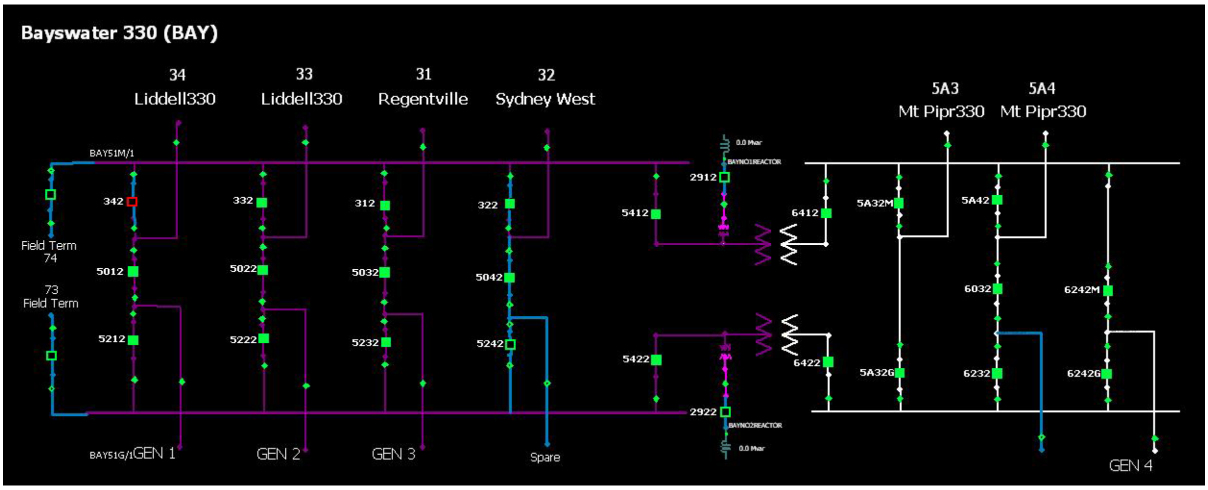

Figure 1. Simulated Supervisory Control and Data Acquisition screen for a power transmission network. A failed circuit appeared as a red square on one of the transmission lines located within the network

Rail control task

A simulated rail control task was used to assess participants' capacity to sustain attention during a longer vigilance task (Howard, Chen, & Wiggins, 2003). The simplified rail control display depicts four separate tracks, two running from left to right and two running from right to left across the screen (see Figure 2). Each track contains an intersection, dividing it into two endpoints, one labeled 'odd' and the other labeled 'even'. Trains, depicted by a red line, are distributed equally across the four tracks, and progress at the same speed across the screen. Each train is labeled with a three-digit number, which is either odd (e.g., 555) or even (e.g., 444). As the trains emerge, their assigned route is indicated by a green line in advance of the train. Odd-numbered trains should be assigned to the 'odd' labeled endpoint, and even-numbered trains should be assigned to the 'even' labeled end-point.

The participants are asked to ensure that the trains are assigned to the correct route. On the occasions when a train is assigned to the incorrect route, participants are instructed to use the mouse to select the grey 'change' button adjacent to the intersection as quickly as possible to reroute the train. Participants have seven seconds once the train appears on the screen to make a decision before the train reaches the intersection.

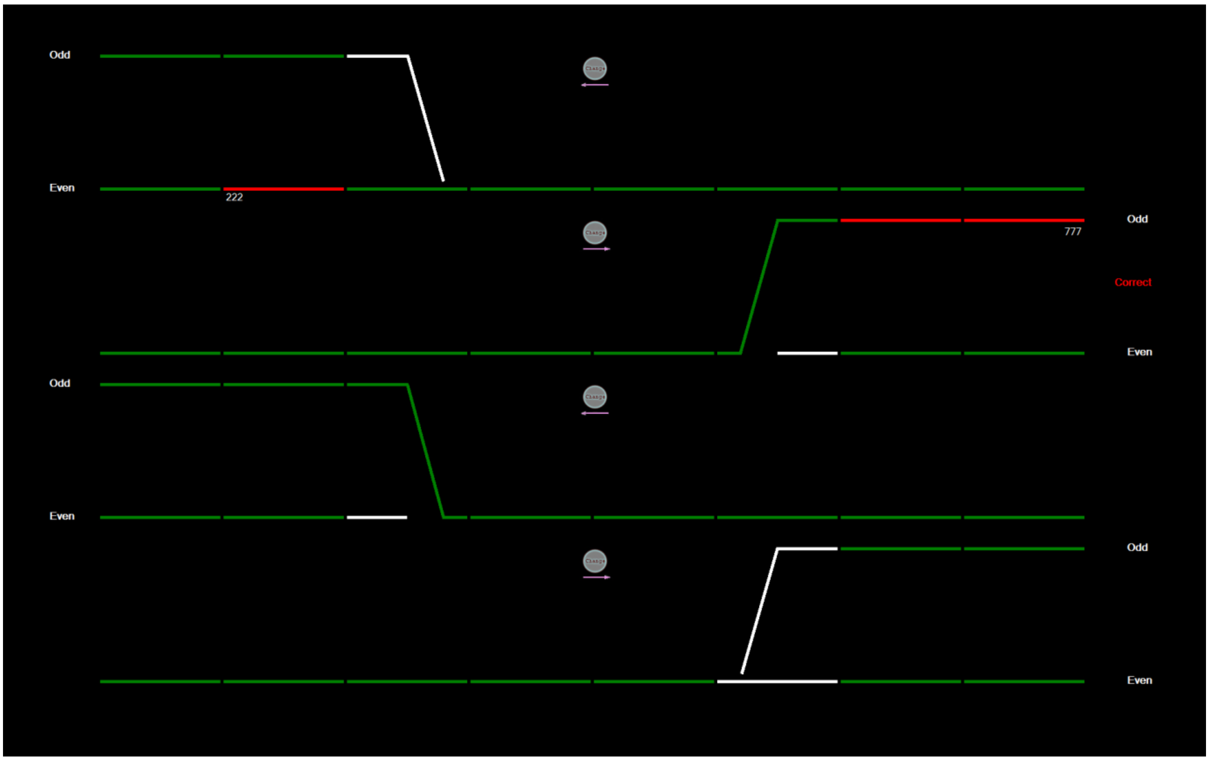

Figure 2. The simulated rail control display as as it appeared to participants. The horizontal green bars represent sections of railway track, while the white bars represent intersection lines. Trains, represented by red bars, move horizontally across the screen, and can be diverted onto intersecting lines by selecting the 'Change' icon located above each track.

In the present study, the rail control task consisted of four conditions, which included a 30-minute process control condition, a 30-minute monitoring condition, a 45-minute process control condition, and a 45-minute monitoring condition. During the 30-minute rail control conditions, a total of 258 trains were presented. In the process control condition, 129 of the trains needed to be rerouted. In the monitoring condition, only four of the trains needed to be rerouted. Four trains in the process control condition were matched to the four trains that needed to be rerouted in the monitoring condition, which were then used as the target trains. One target train appeared in each time block, in both conditions appearing 363, 839, 1260 and 1765 seconds from the beginning of the task. During the 45-minute rail control task, a total of 390 trains were presented. In the process control condition, 195 of the trains needed to be rerouted. In the monitoring condition, only five of the trains needed to be rerouted. One target train appeared in each time block, in both conditions appearing 26, 727, 1400, 2031 and 2669 seconds from the beginning of the task. Accuracy and response latency for correct responses (time in milliseconds from appearance of the train to the selection of the 'change' icon) were recorded for each target train.

Procedure

After reading and signing an information and consent form, participants were randomly allocated to one of the rail control conditions. Participants then completed an online demographic questionnaire, the sustained visual search task, and the rail control task in a counterbalanced order. Computer prompts directed participants through the demographic questionnaire and the sustained visual search task. Standardised instructions were provided verbally for the rail control task. Following the verbal instructions, participants completed a two-minute practice trial before completing the experimental trial.

Results

Response latencies for the sustained visual search task were calculated from the initial appearance of a failed circuit breaker to the correct selection of the circuit breaker. Response latencies for correct responses were divided into five blocks of two consecutive scenarios, and these five variables comprised the dependent variables in subsequent analyses. Accuracy scores were calculated as the percentage of correctly identified targets during each block of the sustained visual search task. Response latencies for correctly rerouted target trains in the rail control task were used as predictor variables.

Sustained visual search task performance

A one-way within-groups Analysis of Variance (ANOVA) was conducted with time (Block 1, Block 2, Block 3, Block 4, and Block 5) as the within-group variable, and mean response latency as the dependent variable (see Figure 3). The assumption of sphericity was violated ($p < .001$). Therefore, a Greenhouse-Geisser adjustment was used. A statistically significant effect was evident for Time, $F(3.1,371.7) = 22.41$, $p < .001$, partial $\eta^2 = .16$. Post-hoc analyses revealed that mean response latencies increased in a linear trend throughout the sustained visual search task, $F(1,120) = 13.66$, $p < .001$, partial $\eta^2 = .10$. There was also a significant quadratic trend for the sustained visual search task, $F(1,120) = 62.19$, $p < .001$, partial $\eta^2 = .34$, suggesting that, consistent with a learning effect, response latency decreased significantly following Block 1, before beginning to increase, consistent with a loss of performance.

A Friedman ANOVA was conducted with time as the within-groups variable, and proportion of correct detections as the dependent variable. A statistically significant effect of time was evident, with a decline in signal detection over the watch, $\chi^2(4) = 15.63$, $p = .004$.

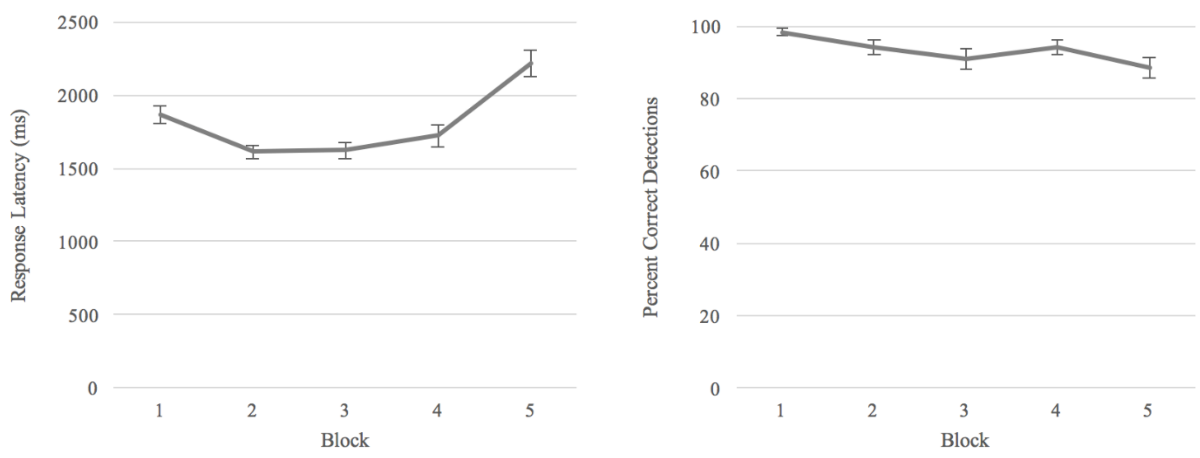

Figure 3. Mean response latency over time on the sustained visual search task (left panel) and mean percentages of correct detections over time on the sustained visual search task (right panel). Error bars represent ± 1 SE.

Correlations between sustained visual search task and rail task performance

Mixed effects modeling was used to calculate the average change in response latency between time blocks during the sustained visual search task. A line of best fit was calculated for each participant, representing the linear relationship between response latency and time throughout the sustained visual search task (Laird & Ware, 1982). The slope of each line reflects the mean change in response latency per time interval of the task. For instance, a slope of 200 indicates that a participant's response latency increased by an average of 200 ms between each time block. The same procedure was used to calculate the mean change in response latency throughout the rail control task. A line of best fit was not calculated for two participants in the 30-minute monitoring condition, as both rerouted fewer than two trains.

Correlations between the change in response latency during the sustained visual search task and change in response latency during the rail control tasks were used to determine whether performance during the sustained visual search task predicts performance during longer sustained attention tasks. There was no statistically significant correlation between the change in response latency on the sustained visual search task and the 30-minute process control condition, $r(30) = .167$, 95% CI [-.215, .549], $p = .377$. However, there was a statistically significant, positive correlation between change in response latency on the sustained visual search task and the 30-minute monitoring condition, $r(28) = .420$, 95% CI [.054, .780], $p = .026$. There was also a significant positive correlation between change in response latency on the sustained visual search task and the 45-minute process control condition, $r(30) = .479$, 95% CI [.139, .819], $p = .007$, and the 45-minute monitoring condition, $r(31) = .364$, 95% CI [.010, .718], $p = .044$.

Change in response latency on the sustained visual search task was not significantly correlated with mean response latency during Block 1 of the sustained visual search task, $r(121) = .063$, 95% CI [-.118, .244], $p = .494$. For each rail control condition, there were also no significant correlations between change in response latency and mean response latency during Block 1 (smallest $p = .201$). The number of correct detections during the sustained visual search task was not correlated to change in response latency during the sustained visual search task, $r(121) = .139$, 95% CI [-.041, .319], $p = .129$, indicating that variations in change in response latency were not the result of differences in speed-accuracy trade-off strategies.

Discussion

The primary aim of this study was to establish the validity of a short-duration vigilance task for process control environments. Changes in mean response latency and accuracy were examined throughout the sustained visual search task, and participants' change in response latency during the sustained visual search task was compared to their change in response latency during either a 30-minute monitoring task, a 45-minute monitoring task, a 30-minute process control task, or a 45-minute process control task. Consistent with previous research using shorter-duration vigilance tasks (Matthews, Warm, Shaw, et al., 2010; Nuechterlein et al., 1983; Temple et al., 2000), it was anticipated that participants would demonstrate a vigilance decrement throughout the sustained visual search task, as evident in increased response latencies and decreased accuracy over the watch. Further, it was hypothesised that changes in response latency over time during the sustained visual search task would be positively associated with changes in response latency over time during the longer-duration monitoring and process control tasks.

As anticipated, mean response latencies increased significantly throughout the sustained visual search task, while the proportion of correct detections decreased significantly. There was also a statistically significant correlation between change in response latency during the sustained visual search task and both the 30-minute and 45-minute monitoring tasks. The change in response latency during the sustained visual search task was not significantly correlated with the change in response latency during the 30-minute process control task, but was positively correlated with the change in performance during the 45-minute process control task. Together, these findings provide evidence to support the construct validity of the sustained visual search task, and support for the hypothesis that, for individual participants, shorter-duration vigilance tasks produce similar experimental outcomes to those observed using longer-duration monitoring and process control tasks.

It is interesting to note that while the sustained visual search task did not predict performance during the 30-minute process control task, it did predict performance during the 45-minute process control task. This is possibly due to the lack of variability in mean response latency during the initial 30-minutes of the task. This is consistent with evidence to suggest that sustained attention in process control environments is maintained for longer periods than it is in pure monitoring conditions (Sturman et al., 2019).

At a theoretical level, the statistically significant correlations evident between performance on the sustained visual search task and performance on the rail control tasks provides evidence to suggest that similar mechanisms influence the vigilance decrement during both shorter and longer vigils. The attentional resource theory of vigilance decrement is based on the proposition that performance declines due to the depletion of cognitive resources during sustained attention tasks (Grier et al., 2003; Parasuraman et al., 1987). Consequently, the correlations observed in the present study may reflect a general capacity on the part of individual participants to conserve their cognitive resources during both shorter-duration and longer-duration tasks.

The findings from the present study suggest that the sustained visual search task may be a valid alternative to a longer-duration process control task for experimental studies. This is important for studies examining individual differences in the rate of vigilance decrement in process control environments, as it suggests that reliable inferences can be made regarding a participant's performance during longer-duration tasks from the participant's performance during the sustained visual search task. At an applied level, the sustained visual search task is likely to be of value for research where there are time constraints, which may occur during field research, or when participants' availability is limited.

Conclusion

The present study was designed to establish the validity of a short-duration sustained visual search task for process control environments. Throughout the sustained visual search task, mean response latencies increased significantly while the proportion of correct detections decreased significantly, in patterns similar to those observed during longer-duration monitoring and process control tasks. Further, changes in response latency during the sustained visual search task were positively correlated with changes in performance during a 30-minute monitoring task, a 45-minute monitoring task, and a 45-minute process control task. These results indicate that the sustained visual search task offers a valid alternative to longer-duration process control tasks for experimental studies.

References (Study 3)

Anders, N. S., Seijmonsbergen, A. C., & Bouten, W. (2011). Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. *Remote Sensing of Environment*, *115*, 2976–2985. doi:10.1016/j.rse.2011.05.007

Brouwers, S., Wiggins, M. W., Helton, W., O'Hare, D., & Griffin, B. (2016). Cue utilization and cognitive load in novel task performance. *Frontiers in Psychology*, *7*, 1–12. doi:10.3389/fpsyg.2016.00435

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

Dimeas, A. L., & Hatziargyriou, N. D. (2005). Operation of a multiagent system for microgrid control. *IEEE Transactions on Power Systems*, *20*, 1447–1455. doi:10.1109/TPWRS.2005.852060

Ghylin, K. M., Drury, C. G., Batta, R., & Lin, L. (2007). Temporal effects in a security inspection task: Breakdown of performance components. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 51, pp. 93–97). Sage Publications Sage CA: Los Angeles, CA.

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

Gupta, R. A., & Chow, M.-Y. (2010). Networked Control System: Overview and Research Trends. *IEEE Transactions on Industrial Electronics*, *57*, 2527–2535. doi:10.1109/TIE.2009.2035462

Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., Wallaart, M., … Hancock, P. A. (2005). Signal regularity and the mindlessness model of vigilance.

British Journal of Psychology, *96*, 249–261. doi:10.1348/000712605x38369

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*, 429–436. doi:10.1007/s00221-012-3322-2

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*, 963–974. doi:10.2307/2529876

Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin*, *139*, 870–900. doi:10.1037/a0030694

Lin, C. J., Yenn, T.-C., & Yang, C.-W. (2010). Automation design in advanced control rooms of the modernized nuclear power plants. *Safety Science*, *48*, 63–71. doi:10.1016/j.ssci.2009.05.005

Mackworth, N. H. (1948). The breakdown of vigilance durning prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*, 6–21. doi:10.1080/17470214808416738

Matthews, G., Warm, J. S., Reinerman-Jones, L. E., Langheim, L. K., Washburn, D. A., & Tripp, L. D. (2010). Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. *Journal of Experimental Psychology: Applied*, *16*, 187–203. doi:10.1037/a0019572

Matthews, G., Warm, J. S., Shaw, T. H., & Finomore, V. S. (2010). A multivariate test battery for predicting vigilance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*, 1072–1076. doi:10.1177/154193121005401405

Metzger, M., & Polakow, G. (2011). A survey on applications of agent technology in industrial process control. *IEEE Transactions on Industrial Informatics*, *7*, 570–581.

doi:10.1109/tii.2011.2166781

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *38*, 311–322. doi:10.1177/001872089606380211

Myers, J. L., Well, A., & Lorch, R. F. (2010). *Research design and statistical analysis*. Routledge.

Navon, R. (2005). Automated project performance control of construction projects. *Automation in Construction*, *14*, 467–476. doi:10.1016/j.autcon.2004.09.006

Nuechterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual Sustained Attention : Image Degradation Produces Rapid Sensitivity Decrement over Time. *Science*, *220*, 327–329. doi:10.1126/science.6836276

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*, 924–927. doi:10.1126/science.472714

Parasuraman, R., Warm, J. S., & Dember, W. N. (1987). Vigilance: Taxonomy and utility. In *Ergonomics and Human Factors: Recent Research* (pp. 11–32). doi:10.1007/978-1-4612-4756-2_2

Pigeau, R. A., Angus, R. G., O'Neill, P., & Mack, I. (1995). Vigilance latencies to aircraft detection among NORAD surveillance operators. *Human Factors*, *37*, 622–634. doi:10.1518/001872095779049291

Posner, M. I. (1978). *Chronometric explorations of mind.* Lawrence Erlbaum.

Reinerman-Jones, L. E., Matthews, G., Langheim, L. K., & Warm, J. S. (2011). Selection for vigilance assignments: A review and proposed new direction. *Theoretical Issues in Ergonomics Science*, *12*, 273–296. doi:10.1080/14639221003622620

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology*, *20*, 343–350. doi:10.1037/h0043220

Shaw, T. H., Emfield, A., Garcia, A., de Visser, E., Miller, C., Parasuraman, R., & Fern, L. (2012). Evaluating the Benefits and Potential Costs of Automation Delegation for Supervisory Control of Multiple UAVs. In *Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, pp. 1498–1502). doi:10.1177/154193121005401930

Shaw, T. H., Funke, M. E., Dillard, M., Funke, G. J., Warm, J. S., & Parasuraman, R. (2013). Event-related cerebral hemodynamics reveal target-specific resource allocation for both “go” and “no-go” response-based vigilance tasks. *Brain and Cognition*, *82*, 265–273. doi:10.1016/j.bandc.2013.05.003

Shaw, T. H., Satterfield, K., Ramirez, R., & Finomore, V. (2013). Using cerebral hemovelocity to measure workload during a spatialised auditory vigilance task in novice and experienced observers. *Ergonomics*, *56*, 1251–1263. doi:10.1080/00140139.2013.809154

Sheridan, T. B. (1970). On how often the supervisor should sample. *IEEE Transactions on Systems Science and Cybernetics*, *6*, 140–145. doi:10.1109/TSSC.1970.300289

Small, A. J., Wiggins, M. W., & Loveday, T. (2014). Cue-based processing capacity, cognitive load and the completion of simulated short-duration vigilance tasks in power transmission control. *Applied Cognitive Psychology*, *28*, 481–487. doi:10.1002/acp.3016

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors*, *42*, 183–194. doi:10.1518/001872000779656480

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

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology*, *25*, 229–235. doi:10.1002/acp.1668

Bridging Section 2

The results from Studies 1 and 2 provided evidence indicating that a general capacity for cue utilisation predicts differences in the consumption of cognitive resources during novel sustained attention tasks. However, Studies 1 and 2 relied on cross-task cue utilisation, where cue utilisation in the domain of driving was used to predict performance in the domain of rail control. Therefore, participants in Studies 1 and 2 consisted of novice operators, who had no previous exposure to rail control. Consequently, the observed differences in response latency and resource consumption based on cue utilisation in these studies likely reflected differences in the speed with which participants acquired new cue-based relationships during the rail control task. However, system operators in industrial environments are typically qualified, and therefore have some level of experience with domain-related tasks.

Qualified system operators differ from novice operators in that they have had prior opportunities to acquire cue-based associations through exposure to the operational environment. Therefore, it is unclear whether differences in cognitive load based on cue utilisation, similar to those observed in Studies 1 and 2, would be demonstrated amongst more experienced practitioners in an operational context. Consequently, there is a need to examine sustained attention and cue utilisation amongst qualified operators. However, due to time constraints and costs, the availability of experienced practitioners in an operational context is time restricted, which limits the utility of longer-duration sustained attention tasks for research within this population.

Study 3 was designed to validate a newly adapted short-duration sustained visual search task that could be used with qualified power control room operators. The sustained visual search task required participants to monitor simulated operating power transmission interfaces, and identify system faults. The findings from Study 3 indicated that the sustained visual search task is a valid alternative to a longer-duration process control task for experimental studies.

Using the sustained visual search task, Study 4 was designed to determine whether qualified practitioners' cue utilisation is predictive of their cognitive load and sustained attention in an operational context. To achieve this, system operators from Australian Distribution Network Service Providers (DNSPs) were recruited to participate in two experiments. During each experiment, system operators completed an assessment of cue utilisation within the context of power distribution and the sustained visual search task validated by Study 3. This experimental design enabled context-based assessments of cue utilisation, whereby cue utilisation and sustained attention were both assessed within the domain of power distribution.

STUDY 4

Publication history

The paper arising from Study 4 was entitled, “Cue Utilisation Predicts Control Room Operators’ Performance in a Sustained Visual Search Task”. A version of this paper has been submitted to the *Ergonomics* Journal, and is currently under review. The author of the present dissertation wrote approximately 80% of this paper.

Cue Utilisation Predicts Control Room Operators' Performance in a Sustained Visual Search Task

Daniel Sturman[1], Mark W. Wiggins[1], Jaime C. Auton[2], and William S. Helton[3]

[1] *Macquarie University, NSW, Australia*

[2] *The University of Adelaide, SA, Australia*

[3] *George Mason University, VA, United States*

Acknowledgements

Support for this research was provided by the Australian Research Council [ARC LP160100211].

Abstract

This research was designed to determine whether qualified practitioners' cue utilisation is predictive of their sustained attention in an operational context. Australian Distribution Network Service Provider (DNSP) operators were recruited for two experiments, and were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of power distribution. Operators' sustained attention was assessed using a domain-related sustained visual search task. In both Experiment 1 and Experiment 2, power distribution operators with higher cue utilisation demonstrated shorter mean response latencies during the sustained visual search task, compared to operators with lower cue utilisation. Further, no differences in accuracy based on cue utilisation were observed during the sustained visual search task. The results are consistent with the proposition that power operators with higher cue utilisation retain greater residual cognitive resources during domain-related sustained attention tasks, compared to operators with lower cue utilisation.

Cue Utilisation Predicts Control Room Operators' Performance in a Sustained Visual Search Task

Control rooms in industrial environments, including power control, rail control and aviation, are designed to allow operators to monitor and control networks remotely (Noyes & Bransby, 2001; O'Hara & Hall, 1992). Control room operators are typically required to make scheduled adjustments to a network, while remaining alert to unexpected deviations in the system state (Lee, Kim, & Jang, 2011; McCalley et al., 2004). In high risk industrial environments, operators must sustain attention, and take appropriate action, to ensure optimal efficiency and the safety of the system (O'Hara & Hall, 1992; Stanton, Salmon, Walker, & Jenkins, 2009). The failure to sustain attention can result in negative outcomes ranging from system inefficiencies to catastrophic disasters (Helton & Warm, 2008; Reason, 2000).

Despite the potential for negative outcomes, during sustained attention tasks that require operators to monitor systems for infrequent critical signals, performance typically declines over time (Davies & Parasuraman, 1982; Mackworth, 1948). This decline in performance, referred to as the vigilance decrement, is evident in an increased response latency and/or a decreased accuracy in response to critical signals (Langner & Eickhoff, 2013; Parasuraman, 1979). While sustained attention research has focused primarily on pure monitoring tasks, control room operations typically require a combination of monitoring and process control, whereby operators make routine adjustments in response to regular system deviations (Cuny, 1979; Nachreiner, Nickel, & Meyer, 2006). However, it is only relatively recently that theories of vigilance decrement have been tested in the context of process control sustained attention tasks.

The resource depletion theory posits that sustaining attention requires effort, and results in the depletion of attentional resources over time (Grier et al., 2003; Helton & Russell, 2013; Warm, Parasuraman, & Matthews, 2008). The depletion of attentional resources over an extended vigil eventually results in fewer resources than are necessary to sustain attention

(Grier et al., 2003; Hancock & Warm, 2003). Consequently, performance efficiency begins to decline, as evident in increased response latency or decreased accuracy when responding to critical signals (Parasuraman, 1979; Shaw, Funke, et al., 2013).

Support for the resource depletion account of the vigilance decrement is drawn from research demonstrating that more cognitively demanding tasks result in a more rapid decline in performance (Helton & Russell, 2013; Shaw, Funke, et al., 2013). For instance, increasing the effort required for task completion, by either decreasing the salience of signals or by adding a secondary task, increases response latency and reduces accuracy in response to critical signals (Helton & Russell, 2011; Helton & Warm, 2008; Smit, Eling, & Coenen, 2004). Conversely, reducing the effort required, by increasing signal salience or by inserting warning signals, increases the proportion of signals that are detected correctly (Helton & Warm, 2008; MacLean et al., 2009).

The resource depletion theory is based on the proposition that operators have limited attentional resources, which are depleted throughout effortful tasks (Thomson, Besner, & Smilek, 2015). It follows that operators who consume fewer resources during a task will retain greater residual cognitive resources, thereby enabling the performance of a task for extended periods.

Cue utilisation and cognitive load

The application of cues during the performance of a task is one strategy by which operators can reduce the consumption of cognitive resources (Brouwers, Wiggins, Helton, O'Hare, & Griffin, 2016; Perry, Wiggins, Childs, & Fogarty, 2012; Sturman, Wiggins, Auton, & Loft, 2019; Wiggins, 2011). Cues are recognition-driven associations between situation-specific environmental features and an event or object (Brunswik, 1955; Klein, Calderwood, & Clinton-Cirocco, 1986). Cue utilisation refers to the application of cue-based associations from memory (Lansdale, Underwood, & Davies, 2010; Wiggins, Loveday, & Auton, 2015).

Effective cue utilisation allows operators to attend to features of greater relevance, reducing the overall number of features to which they attend, and thereby reducing the rate at which cognitive resources are consumed (Sturman et al., 2019; Weiss & Shanteau, 2003; Williams, Ward, Knowles, & Smeeton, 2002).

Brouwers et al. (2016) suggest that participants with higher cue utilisation adopt visual search strategies that are associated with the consumption of fewer resources during sustained attention tasks. Evidence to support this assertion can be inferred from experiments demonstrating that participants with higher cue utilisation report relatively lower perceived cognitive load, and record smaller increases in response latency, during sustained attention tasks (Brouwers, Wiggins, Griffin, Helton, & O'Hare, 2017; Brouwers et al., 2016).

Sturman et al. (2019) examined the relationship between cue utilisation and cognitive load using eye tracking data to assess fixation rates, and Functional Near Infrared Spectroscopy (fNIRS) to measure cerebral oxygenation in the prefrontal cortex. Lower fixation rates and smaller increases in cerebral oxygenation in the prefrontal cortex are indicative of more efficient search patterns and a relatively lower consumption of cognitive resources during sustained attention tasks (Goldberg & Kotval, 1999; Helton et al., 2010; Ong, Russell, & Helton, 2013; Poole & Ball, 2006).

In the case of Sturman et al. (2019), participants classified with higher cue utilisation in the domain of driving and wayfinding demonstrated greater decreases in fixation rates, and smaller increases in cerebral oxygenation in the prefrontal cortex compared to participants with lower cue utilisation during simulated sustained attention rail control tasks. Further, participants with higher cue utilisation recorded smaller increases in response latency throughout the rail control tasks. These outcomes suggest that participants with higher cue utilisation maintained sufficient cognitive resources to remain alert and respond quickly, providing further evidence to the effect that individuals with higher cue utilisation consume cognitive resources at a rate that enables them to sustain performance over time (Brouwers et al., 2016; Perry et al., 2012; Wiggins, 2011).

Cue utilisation in an operational context

Investigations examining the relationship between cue utilisation and sustained attention have typically involved naïve practitioners engaging in simplified representations of the domain (Brouwers et al., 2016, 2017; Sturman et al., 2019). Differences in response latency and the consumption of cognitive resources based on cue utilisation generally reflected differences in the rate with which naïve participants acquired new cue-based relationships during novel tasks. Therefore, it remains unclear whether these differences would also be evident amongst more experienced system operators who differ from novice operators in that they have had prior opportunities to acquire cue-based associations through exposure to the operational environment. Consequently, while higher cue utilisation should be associated with a greater frequency of previously acquired cues (McCormack, Wiggins, Loveday, & Festa, 2014; Morrison, Wiggins, Bond, & Tyler, 2013), there is a need to establish whether similar effects are demonstrated amongst more experienced practitioners in an operational context.

Experiment 1

The aim of the present study was to establish whether experienced practitioners' cue utilisation is predictive of their sustained attention in an operational context. To assess cue utilisation and sustained attention, control room operators from 12 Australian Distribution Network Service Providers (DNSPs) were recruited to participate in the study. Cue utilisation, assessed in the context of power control, was used to predict performance on a domain-related sustained visual search task. Based on the proposition that operators with higher cue utilisation would consume fewer cognitive resources during a sustained attention task, it was hypothesised that these operators would demonstrate a greater capacity to sustain performance during a power control sustained visual search task, compared to operators with lower cue utilisation.

Method

Design

Experiment 1 comprised a 2 x 5 mixed factorial design incorporating two levels of cue utilisation (higher and lower) as a between-subjects factor, and five blocks of trial pairs (Block 1, Block 2, Block 3, Block 4, and Block 5) as a within-subjects factor. Participants were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of power distribution. Trial pairs constituted five blocks of two consecutive scenarios within the sustained visual search task.

Participants

Participants comprised 113 distribution network power controllers (112 males), recruited from 12 Australian DNSPs. Network controllers ranged in age from 26 to 63 years (M = 45.1, SD = 9.0), had acquired a mean 10.1 years (SD = 6.9) of experience as network controllers, and had acquired a mean 21.1 years (SD = 10.3) working in power distribution.

EXPERTise 2.0

EXPERT Intensive Skills Evaluation (EXPERTise 2.0) is an online assessment tool designed to assess behaviour consistent with the utilisation of cues within a specific context (Wiggins et al., 2015). For the current study, EXPERTise 2.0 was tailored to the domain of power distribution. Tasks in the EXPERTise battery include a Feature Identification Task (FIT), a Feature Recognition Task (FRT), a Feature Association Task (FAT), a Feature Discrimination Task (FDT), and a Feature Prioritisation Task (FPT).

The FIT requires participants to identify key features, and is based on the observation that, compared to novices, experts are able to identify and utilise visual features in the environment that are more diagnostic of the system state (Müller, Abernethy, & Farrow, 2006; Schriver, Morrow, Wickens, & Talleur, 2008; Wiggins, 2014). Consequently, higher cue utilisation is generally associated with a lower mean response latency (Loveday, Wiggins, & Searle, 2013; Schriver et al., 2008). Participants are presented with six line diagrams using electrical symbology with which the participants are familiar. Each line diagram comprises a scenario that might be experienced as a network controller, including a voltage under or overload, a transformer failure, an erroneous indication, or a normal condition. Participants are asked to review each line diagram, and to identify the area of concern by selecting the specific feature of interest. They are able to select an icon that reflects normal operations if they consider the system to be operating normally. Response latency is measured as the time in milliseconds from the initial presentation of the diagram to the selection of an area of concern.

The FRT consists of 10 trials, during which participants are presented with line diagrams similar to those used in the FIT. Exposure to the line diagrams is limited to a short period, with the period of exposure varying between 20 seconds and 60 seconds depending on the complexity of the scenario. Following exposure, the line diagram is removed, and participants are asked to select from a list the condition represented in the preceding display (e.g., "There has been a complete loss of indications at this substation", "There has been a Tx1 fault resulting in all load being lost", etc.). The FRT is designed to assess the capacity to rapidly extract key information from power distribution scenarios and form an accurate judgement. Higher cue utilisation is associated with a greater number of correct judgements (Brouwers et al., 2017; Wiggins & O'Hare, 2003a).

The FAT consists of 13 trials, during which participants are presented with power distribution-related terms (e.g., 'Feeder Trip, 'Substation'). For each trial, two terms are displayed for a duration of 2000 ms, after which participants rate the extent to which they believe that the two terms are related on a 6-point Likert scale ranging from 1 (*Extremely unrelated*) to 6 (*Extremely related*). As cue utilisation requires the identification of predictive feature-event relationships, higher cue utilisation is associated with greater variance in the perceived relatedness of terms (Morrison et al., 2013; Schvaneveldt, Beringer, & Lamonica, 2001).

In the FDT, participants are presented with two detailed power distribution scenarios, and are asked to make a decision based on their typical response to similar scenarios. After initiating a response for each scenario, participants are then asked to rate a list of features based on their relevance to the decision using a 10-point Likert scale ranging from 1 (*Not important at all*) to 10 (*Extremely important*). Effective cue utilisation requires features to be identified as more or less relevant (Pauley, O'Hare, & Wiggins, 2009; Weiss & Shanteau, 2003), and therefore, higher cue utilisation is associated with a greater variance in ratings (Brouwers et al., 2017).

The FPT consists of two problem scenarios (e.g., advised that there is no supply to a specific region), which are accompanied by drop-down lists of key features relating to the scenario. Each drop-down menu is feature-labelled (e.g., 'Number of Substations', 'Location of Field Staff'), and upon selection, provides participants with information pertaining to the scenario. Only one feature is accessible at any one time, and access to the information is time-limited to 90 seconds. The FPT assesses participants' capacity to prioritise feature cue acquisition (Wiggins & O'Hare, 1995; Wiggins, Stevens, Howard, Henley, & O'Hare, 2002). Lower cue utilisation is more likely to be associated with the selection of information in the sequence in which it is presented (e.g., from top to bottom as they appear on the display screen), while higher cue utilisation is associated with a lower ratio of menus selected in the sequence in which they are presented (Wiggins et al., 2002).

Sustained visual search task

During the sustained visual search task (Small, Wiggins, & Loveday, 2014), participants view a series of 10 short, dynamic scenarios. Each scenario displays a Supervisory

Control and Data Acquisition (SCADA) interface consisting of an operating power transmission grid (see Figure 1) and displays a different transmission grid (Small et al., 2014).

Written instructions explain to participants that a fault (a failed circuit breaker), represented by a flashing red square, will appear during each scenario. Participants are instructed to select the fault, as quickly as possible, using the mouse. A failed circuit breaker always appears on a section of the grid 30 seconds from the onset of each scenario and remains visible until the operator selects the fault. Response latencies are calculated as the time, in milliseconds, from the initial appearance of a fault to the correct selection of the circuit breaker. The sustained visual search task continues for approximately seven minutes, with the total duration dependent upon the response latency of the operator.

Response latency and accuracy were recorded for each individual across each of 10 experimental scenarios that comprised the sustained visual search task. Accuracies were recorded as correct if responses were less than 10 mm from the fault location.

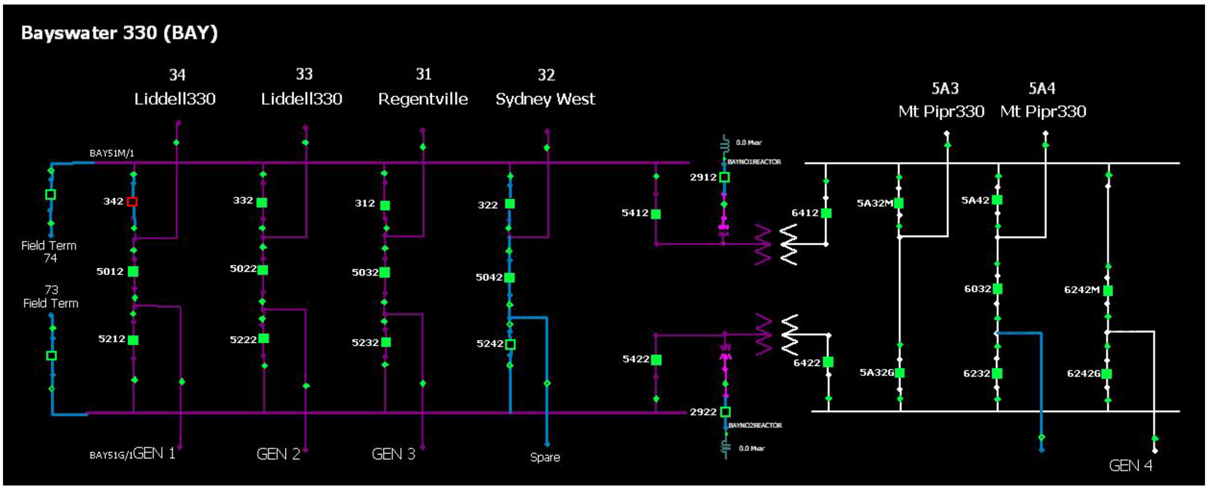

Figure 1. Simulated Supervisory Control and Data Acquisition screen for a power transmission network. A failed circuit appears as a red square on one of the transmission lines located within the network.

Procedure

Following approval from the University Human Research Ethics Committee, power controllers from the 12 DNSPs were provided general information regarding the research, and were given a URL to the EXPERTise 2.0 website. After logging in, participants were asked to read the information sheet and give consent to their participation in the research. Upon giving their consent, participants answered a series of questions that were incorporated to generate a unique participant code.

Network controllers were asked a series of demographic questions, including their age, sex, the number of years they had been employed as a network controller, and the number of years they had worked in power distribution. On completion of the demographic questions, participants completed the EXPERTise tasks and the sustained visual search task. Once participants had completed all the tasks, they were thanked for their participation and offered the opportunity to enter the draw to win an iPad mini.

Results

Response latencies for the sustained visual search task were calculated from the initial appearance of a failed circuit breaker to the correct selection of the circuit breaker. Response latencies for correct responses were divided into five blocks of two consecutive scenarios, and these five variables comprised the dependent variables in subsequent analyses. Accuracy scores were calculated as the proportion of correctly identified targets during each block of the sustained visual search task.

Cue utilisation typologies

Consistent with a standard approach to assessing cue utilisation, EXPERTise data were used to identify cue utilisation typologies that corresponded to higher or lower levels of cue utilisation (Brouwers et al., 2016; Small et al., 2014; Sturman et al., 2019). Consistent with previous research (Loveday, Wiggins, Searle, Festa, & Schell, 2013; Wiggins, Brouwers, Davies, & Loveday, 2014), scores for each task were converted to *z* scores and cluster analyses using the *z* scores were performed to identify two groups. The first cluster, labelled the higher cue utilisation typology, consisted of participants who recorded a shorter response latency on the FIT, greater accuracy on the FRT, a higher mean ratio of variance to reaction time on the FAT, a greater variance in ratings on the FDT, and a lower mean ratio of sequential selections in the FPT.

The second cluster, labelled the lower cue utilisation typology, consisted of participants who had a greater response latency on the FIT, lower accuracy on the FRT, a lower variance in ratings on the FDT, a lower mean ratio of variance to reaction time on the FAT, and a higher mean ratio of sequential selections in the FPT. Independent samples *t* tests demonstrated statistically significant differences in FIT, FRT, FAT and FDT scores between the higher and lower cue utilisation typologies (see Table 1). In the case of the FPT, the difference was non-significant. Nevertheless, the pattern of responses was generally consistent with the pattern which would normally be expected to characterise higher or lower cue utilisation. Thirty-four participants were classified in the higher cue utilisation typology and 79 participants were classified in the lower cue utilisation typology.

Table 1. *Experiment 1 cluster centroids for the EXPERTise task scores.*

	Cluster 1 (n=34)	Cluster 1 (n=79)			
	Higher cue utilisation	Lower cue utilisation	*t*	df	*p*
FIT	-0.35	0.15	2.49*	111	=.014
FRT	0.48	-0.21	3.52**	111	=.001
FAT	1.03	-0.44	9.80**	111	<.001
FDT	0.72	-0.31	5.66**	111	<.001
FPT	0.21	-0.09	1.50	111	=.137

* Significant at the 0.05 level (two-tailed); **Significant at the 0.01 level (two-tailed)

Covariates

Pearson's correlations indicated that the years that participants had been employed as network controllers and the number of years they had worked in power distribution was not associated with response latency or accuracy during the sustained visual search task (ps > .05). Independent samples t tests indicated that cue utilisation was not associated with the number of years participants had been employed as network controllers, nor the number of years they had worked in power distribution (ps > .05). Consequently, it was not necessary to include these variables as potential covariates in the main analyses.

Response latency

A 2 x 5 Analysis of Variance (ANOVA) was conducted with cue utilisation (higher and lower) as a between-subjects factor, and trial pairs (Block 1, Block 2, Block 3, Block 4, and Block 5) as a within-subjects factor. Due to technical issues with the DNSP's firewall security, 18 participants were unable to complete the sustained visual search task. Consequently, response latency and accuracy were not assessed for these 18 participants. The assumption of sphericity was violated ($p < .001$). Therefore, a Greenhouse-Geisser adjustment was used for all within-subjects' effects. A statistically significant main effect was evident for cue utilisation, $F(1,88) = 4.09$, $p = .046$, partial $\eta^2 = .044$, indicating a statistically significant, greater response latency for participants with lower cue utilisation compared to participants with higher cue utilisation (see Figure 2). There was no statistically significant main effect of trial pairs, $F(1.8,157.4) = 2.15$, $p = .126$, partial $\eta^2 = .024$, and no statistically significant interaction between cue utilisation and trial pairs, $F(1.8,157.4) = 0.12$, $p = .864$, partial $\eta^2 = .001$.

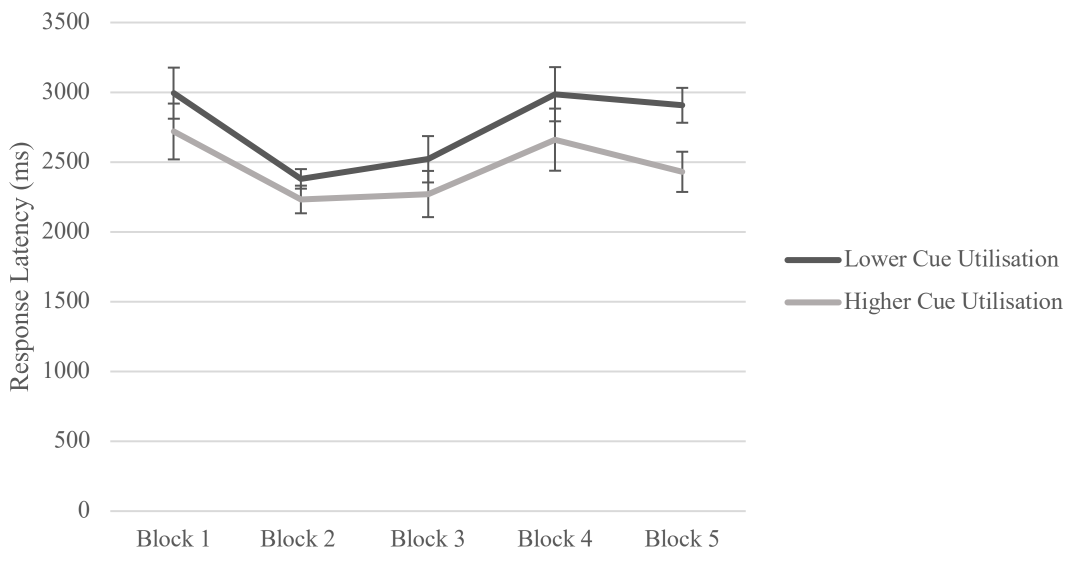

Figure 2. Mean response latencies by cue utilisation typology and trial pairs. Error bars represent ± 1 SE.

Accuracy

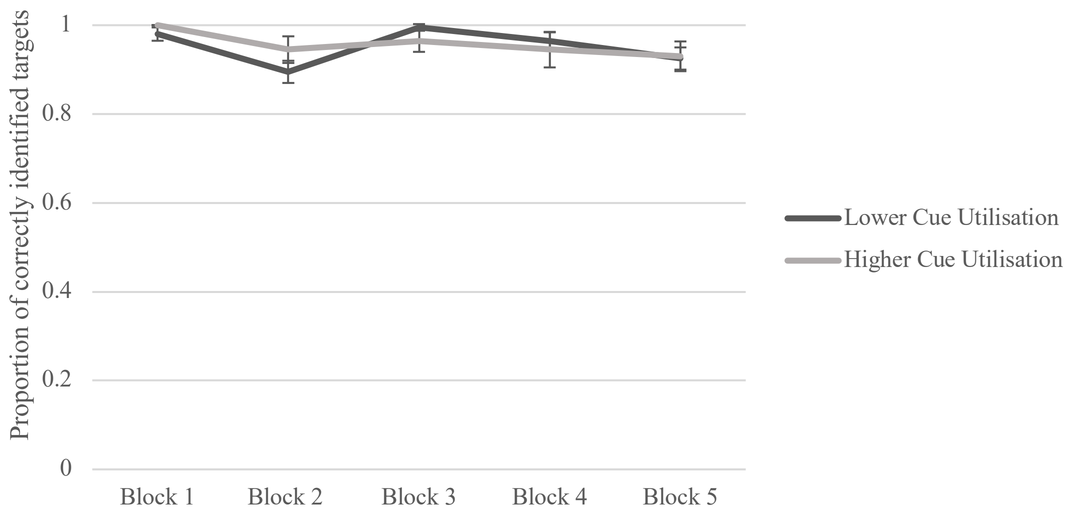

Figure 3. Mean proportion of correctly identified targets by cue utilisation typology and trial pairs. Error bars represent ± 1 SE.

A series of Mann-Whitney U non-parametric tests revealed that, for each block of the sustained visual search task, there was no significant difference in accuracy between operators with higher and lower cue utilisation (ps > .05; see Figure 3). This indicates that differences in mean response latencies based on cue utilisation were not the result of differences in speed-accuracy trade-off strategies.

Discussion

The aim of Experiment 1 was to examine whether experienced practitioners' cue utilisation is predictive of their sustained attention in an operational context. Cue utilisation amongst experienced distribution power controllers was assessed using a power distribution version of EXPERTise 2.0, and operators were classified as having either higher or lower cue utilisation. The results from Experiment 1 indicate that operators with higher cue utilisation demonstrated shorter mean response latencies during a power network sustained visual search task, compared to operators with lower cue utilisation. Further, no differences in accuracy were evident during the visual search task. In combination, these results indicate that power operators with higher cue utilisation are able to maintain greater performance during a domain-specific sustained attention task, compared to participants with lower cue utilisation. This suggests that higher cue utilisation amongst experienced practitioners is associated with the consumption of fewer cognitive resources during domain-related sustained attention tasks (Brouwers et al., 2016; Perry et al., 2012; Sturman et al., 2019).

While the results demonstrated a significant difference in mean response latency based on cue utilisation, the level of significance was marginal (p = .046). Consequently, it would be valuable to replicate the results of Experiment 1 to evaluate the replicability of the findings (Pashler & Wagenmakers, 2012; Schooler, 2014). Further, while the sustained visual search task has previously demonstrated concurrent validity (Small et al., 2014; Sturman et al., 2019), there is a need to establish the reliability of the task by demonstrating that operators' responses remain relatively consistent over time.

Experiment 2

Experiment 2 was designed to replicate the results of Experiment 1, to provide stronger evidence that higher cue utilisation amongst experienced practitioners is associated with the capacity to sustain performance for extended periods. Further, Experiment 2 was intended to establish the reliability of the sustained visual search task as a measure of sustained attention. To demonstrate the replicability of results using a broader sample, additional power distribution operators, who did not participate in Experiment 1, were recruited to participate in Experiment 2. However, to demonstrate test-retest reliability, operators who participated in Experiment 1 were also invited to participate in Experiment 2. Based on the proposition that participants with higher cue utilisation consume fewer resources during the performance of a task and thereby retain greater residual cognitive resources, it was hypothesised that these operators would demonstrate a greater capacity to sustain performance, compared to operators with lower cue utilisation. Further, it was hypothesised that, amongst those participants who engaged in Experiments 1 and 2, performance on the sustained visual search task during Experiment 1 would be positively correlated with performance on the task during Experiment 2.

Method

Design

Experiment 2 comprised the same 2 x 5 mixed factorial design as Experiment 1, incorporating two levels of cue utilisation (higher and lower) as a between-subjects factor, and five levels of trial pairs (Block 1, Block 2, Block 3, Block 4, and Block 5) as a within-subjects factor.

Participants

Participants in Experiment 2 comprised 86 distribution power controllers, of whom 46 had completed Experiment 1. Participants ranged in age from 26 to 62 years (*M* = 45.3, *SD* = 8.9), and had acquired a mean 9.9 years (*SD* = 6.7) of experience as network controllers, and had acquired a mean 21.8 years (*SD* = 11.0) working in power distribution.

EXPERTise

The same five tasks from the EXPERTise 2.0 power distribution battery utilised in Experiment 1 were used to assess participants' level of cue utilisation. To reduce the likelihood of ceiling effects in the FPT, the time for each scenario was reduced from 90 seconds to 60 seconds. The FIT, FRT, FAT and FDT remained unchanged from Experiment 1.

Sustained visual search task

The sustained visual search task contained the same 10 short, dynamic scenarios as those used in Experiment 1, presented in a new randomised order.

Procedure

The procedure for Experiment 2 was identical to Experiment 1.

Results

Cue utilisation typologies

Consistent with Experiment 1, a cluster analysis was undertaken using *z* scores for all five EXPERTise tasks to identify the cue utilisation profiles that corresponded with relatively higher and lower cue utilisation. As with Experiment 1, the first cluster, labelled the higher cue utilisation typology, consisted of participants who had a shorter response latency on the FIT, greater accuracy on the FRT, a higher mean ratio of variance to reaction time on the FAT, a greater variance in ratings on the FDT, and a lower mean ratio of sequential selections in the FPT. The second cluster, labelled the lower cue utilisation typology, consisted of participants who had a greater response latency on the FIT, lower accuracy on the FRT, a lower variance in ratings on the FDT, a lower mean ratio of variance to reaction time on the FAT, and a higher mean ratio of sequential selections in the FPT. Independent samples *t* tests demonstrated significant differences in FIT, FRT, and FDT scores between the higher and lower cue utilisation typologies (see Table 2). In the case of the FDT and FPT, the differences were non-significant. Nevertheless, the pattern of responses was generally consistent with the pattern which would normally be expected to characterise higher or lower cue utilisation. Fifty-four participants were classified in the higher cue utilisation typology and 34 participants were classified in the lower cue utilisation typology.

Table 2. *Experiment 2 cluster centroids for the EXPERTise task scores.*

	Cluster 1 (n=54)	Cluster 1 (n=34)			
	Higher cue utilisation	Lower cue utilisation	*t*	df	*p*
FIT	-0.45	0.71	6.38**	86	<.001
FRT	0.49	-0.76	7.18**	86	<.001
FAT	0.25	-0.37	2.92**	86	=.004
FDT	0.14	-0.24	1.75	86	=.083
FPT	0.14	-0.21	1.82	86	=.073

* Significant at the 0.05 level (two-tailed); **Significant at the 0.01 level (two-tailed)

Covariates

Pearson's product-moment correlations indicated that years of employment as network controllers and the number of years working within power distribution was not associated with response latency nor accuracy during the sustained visual search task (*p*s > .05). Independent samples *t* tests indicated that, for both experiments, cue utilisation was not associated with the number of years participants had been employed as network controllers, nor the number of years they had worked in power distribution (ps > .05). Consequently, it was not necessary to include covariates as potential explanatory variables in the main analyses.

Response latency

Consistent with Experiment 1, a 2 x 5 Analysis of Variance (ANOVA) was conducted with cue utilisation as a between-subjects factor, and trial pairs as a within-subjects factor. Due to a software error, response latency during the sustained visual search task was not recorded for eight of the participants. Nevertheless, a statistically significant main effect was evident for cue utilisation, $F(1,78) = 4.76$, $p = .032$, partial $\eta^2 = .057$, indicating a significantly greater mean response latency for participants with lower cue utilisation compared to participants with higher cue utilisation (see Figure 4). There was no statistically significant main effect for trial pairs, $F(4,312) = 0.28$, $p = .893$, partial $\eta^2 = .004$, and no statistically significant interaction between cue utilisation and trial pairs, $F(4,312) = 0.30$, $p = .880$, partial $\eta^2 = .004$.

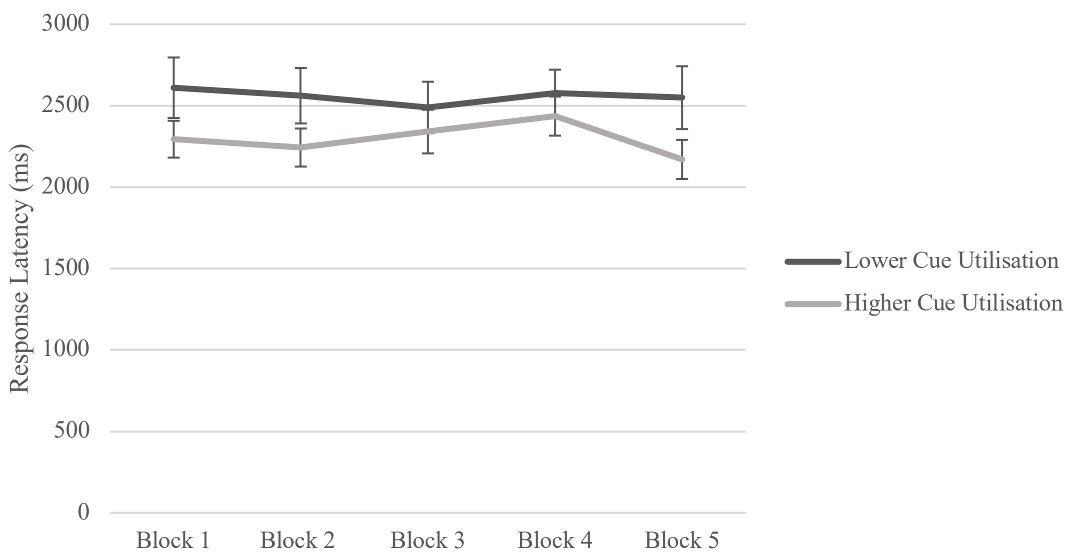

Figure 4. Mean response latencies by cue utilisation typology and trial pairs. Error bars represent ± 1 SE.

Accuracy

A series of Mann-Whitney U non-parametric tests revealed that, for each block of the sustained visual search task, there was no significant difference in accuracy between operators with higher and lower cue utilisation (*p*s > .05; see Figure 5). This indicates that differences in mean response latencies based on cue utilisation were not the result of differences in speed-accuracy trade-off strategies.

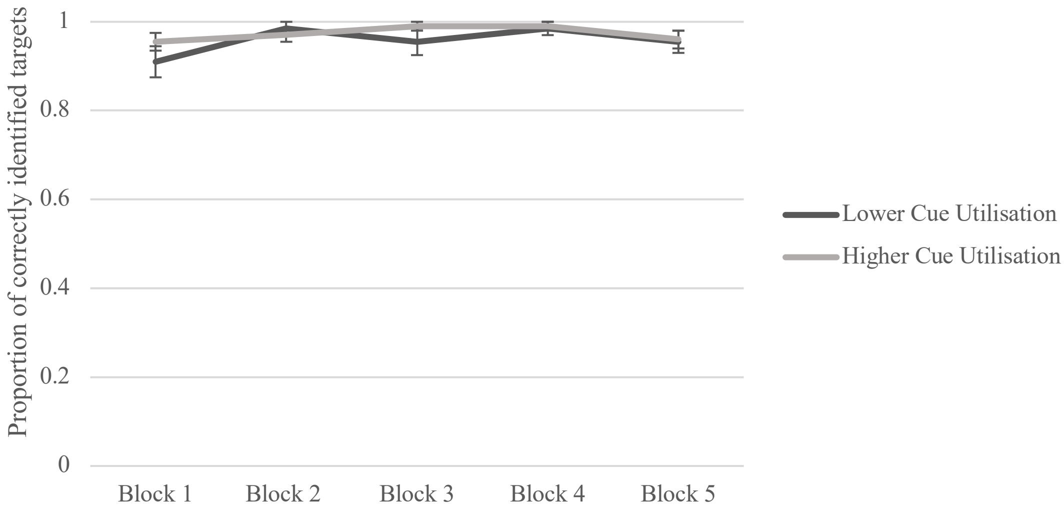

Figure 5. Mean proportion of correctly identified targets by cue utilisation typology and trial pairs. Error bars represent ± 1 SE.

Test-retest reliability

Correlational analyses were conducted to examine the relationship between performance on the sustained visual search task in Experiment 1 and Experiment 2 for the 46 operators who participated in both experiments. A Pearson's correlation indicated a strong positive relationship for mean response latency in Experiment 1 and Experiment 2, $r = .690$, $p < .001$. Further, a Spearman's correlation indicated a moderate positive correlation between mean accuracy scores in Experiment 1 and Experiment 2, $r = .369$, $p = .019$. In combination, these results provide evidence supporting the reliability of the sustained visual search task as a measure of sustained attention.

Discussion

The primary aim of Experiment 2 was to replicate and extend the outcomes of Experiment 1, and test whether qualified practitioners' cue utilisation is predictive of their sustained attention. Experiment 2 was also intended to establish the reliability of the sustained visual search task as a measure of sustained attention in the context of a process control task.

Consistent with the results from Experiment 1, operators with higher cue utilisation demonstrated shorter mean response latencies during the sustained visual search task, and no difference in mean accuracy, compared to operators with lower cue utilisation. These results provide further support for the proposition that power operators with higher cue utilisation are able to maintain greater performance during a domain-specific sustained attention task, compared to participants with lower cue utilisation. This outcome is consistent with the notion that higher cue utilisation amongst experienced practitioners is associated with lower cognitive load during domain-related sustained attention tasks (Brouwers et al., 2016; Perry et al., 2012; Sturman et al., 2019). Further, mean response latency and accuracy during Experiment 1 were positively associated with mean response latency and accuracy during Experiment 2, providing evidence to support the test-retest reliability of the sustained visual search task.

General Discussion

The empirical findings from Experiment 1 and Experiment 2 indicate that control room operators with higher cue utilisation are able to maintain a greater level of performance during sustained visual search tasks, compared to operators with lower cue utilisation. The outcomes replicate previous research involving inexperienced operators, which has demonstrated that cue utilisation is associated with differences in response latency during novel sustained attention tasks (Brouwers et al., 2016, 2017; Sturman et al., 2019). Further, the findings extend previous research (e.g., Small et al., 2014), demonstrating that differences in response latency based on cue utilisation are consistently demonstrated by experienced practitioners when completing a validated task which is relevant to their domain of expertise.

Attentional resource theory posits that cognitive resources are drawn from a limited supply, and that the allocation of cognitive resources to a task depletes the residual attentional resources available (Kahneman, 1973; Warm et al., 2008; Wickens, 1980). A reduction in residual cognitive resources can reduce accuracy and increase operational errors (Reason, 1990; Wickens, Hollands, Banbury, & Parasuraman, 2015). Consequently, participants who consume fewer cognitive resources during the sustained visual search task should demonstrate shorter mean response latencies while maintaining a high level of accuracy. The shorter mean response latencies recorded by participants with higher cue utilisation during the sustained visual search task is consistent with the proposition that effective cue utilisation allows operators to consume fewer cognitive resources during sustained attention tasks. These results also provide support for the assertion that a greater capacity for cue utilisation amongst experienced practitioners reduces cognitive load by creating efficiencies in information processing without impacting negatively on performance (Brouwers et al., 2016; Lansdale et al., 2010; Sturman et al., 2019).

The results of the present study have applied implications for the selection of qualified operators in high-risk industrial environments. Operators who are able to utilise cues in their operational environment should be able to respond quickly and adaptively to meet the needs of critical situations (Klein, 2008). Further, by consuming fewer cognitive resources during the completion of their primary tasks, operators with higher cue utilisation should retain greater cognitive resources, allowing them to better manage the demands of secondary tasks (Brouwers et al., 2017; Wickens, 2002). Consequently, the capacity to identify cue utilisation performance may aid in the selection of job applicants who are better able to sustain attention and maintain performance during demanding situations. Further, by reducing cognitive load, higher cue utilisation may enable the acquisition of new features and the opportunity to refine existing cues. Therefore, measures of cue utilisation may provide the basis to select job applicants who are more likely to acquire new skills in the absence of formal training.

Consistent with previous research, cue utilisation was not a function of the number of years that operators had been employed as network controllers, nor the number of years they had worked in power distribution (Crane et al., 2018; Loveday, Wiggins, Festa, Schell, & Twigg, 2013). This suggests experience alone does not predict cue utilisation (Crane et al., 2018). Consequently, measures of cue utilisation might also be applied to identify employees who are most in need of cue-based training interventions. For example, cue utilisation could be increased through cue discovery, which utilises simulated operating systems to enable the ad-hoc acquisition of cues through trial and error (Klayman, 1988; Morrison & Morrison, 2015; Wiggins, 2015). Evidence supporting this approach was reported Visser, Tichon, and Diver (2012), who demonstrated that enabling participants to discover cues through trial and error using a heavy equipment simulator, improved civil construction operators truck-loading performance following training.

Cue-based training could also be used to change the way that information is presented to operators at particular stages of training to assist with the acquisition, interpretation, and integration of cues in their operational environment (Scherer et al., 2008; Wiggins & O'Hare, 2003b). For example, by presenting qualified pilots with weather-related features and the associated consequences, and allowing the pilots to practice using these feature-event relationships in a safe environment, Wiggins and O'Hare (2003) demonstrated improved timeliness of weather-related decision making.

Limitations and future direction

While it is hypothesised that differences in sustained attention based on cue utilisation reflect differences in cognitive resource consumption, these results need to be interpreted with caution due to the potential confounding effect of common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Cue utilisation and sustained attention were both assessed using similar methodologies. For instance, both assessments relied on measures of response latency and accuracy. Further, all of the assessments were run using EXPERTise 2.0 software, and data for each operator were typically collected in a single session. Consequently, the relationship between cue utilisation and sustained attention may be partially attributable to the measurement method, rather than differences in cognitive resource consumption (Podsakoff et al., 2003). To control for common method bias, there is a need to use alternative measures of performance and cognitive load amongst qualified operators. For instance, physiological measures, including measures of cerebral blood flow and eye behaviour metrics, could provide additional insight regarding the consumption of cognitive resources during operational tasks (Evans & Fendley, 2017; Langner & Eickhoff, 2013; Shaw, Satterfield, Ramirez, & Finomore, 2013), while also reducing potential common method biases.

Conclusion

The current study was designed to determine whether qualified practitioners' cue utilisation is predictive of their sustained attention in an operational context. In both Experiment 1 and Experiment 2, power distribution operators with higher cue utilisation demonstrated shorter mean response latencies during a power network sustained visual search task, compared to operators with lower cue utilisation. Further, no differences in accuracy, based on cue utilisation, were observed during the visual search task. The results of the study are consistent with the proposition that power operators with higher cue utilisation consume cognitive resources at a lower rate during domain-related sustained attention tasks, compared to operators with lower cue utilisation, thereby enabling sustained performance. From an applied perspective, these results suggest that measures of cue utilisation may assist in the selection, training and management of qualified operators in high risk industrial environments.

References (Study 4)

Brouwers, S., Wiggins, M. W., Griffin, B., Helton, W. S., & O'Hare, D. (2017). The role of cue utilisation in reducing the workload in a train control task. *Ergonomics*, *60*, 1–16. doi:10.1080/00140139.2017.1330494

Brouwers, S., Wiggins, M. W., Helton, W., O'Hare, D., & Griffin, B. (2016). Cue utilization and cognitive load in novel task performance. *Frontiers in Psychology*, *7*, 1–12. doi:10.3389/fpsyg.2016.00435

Brunswik, E. (1955). Representative Design and Probabilistic Theory in a Functional Psychology. *Psychological Review*, *62*, 193–217. doi:10.1037/h0047470

Crane, M. F., Brouwers, S., Wiggins, M. W., Loveday, T., Forrest, K., Giselle, S., … Cyna, A. M. (2018). " Experience Isn ' t Everything ": How Emotion Affects the Relationship Between Experience and Cue Utilization. *Human Factors*, *60*, 685–698. doi:10.1177/0018720818765800

Cuny, X. (1979). Different levels of analysing process control tasks. *Ergonomics*, *22*, 415–425. doi:10.1080/00140137908924626

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

Evans, D. C., & Fendley, M. (2017). A multi-measure approach for connecting cognitive workload and automation. *Journal of Human Computer Studies*, *97*, 182–189. doi:10.1016/j.ijhcs.2016.05.008

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, *24*, 631–645. doi:10.1016/S0169-8141(98)00068-7

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

Hancock, P. A., & Warm, J. S. (2003). A Dynamic Model of Stress and Sustained Attention. *Journal of Human Performance in Extreme Environments*, *7*, 519–537. doi:10.7771/2327-2937.1024

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*, 429–437. doi:10.1007/s00221-011-2749-1

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*, 429–436. doi:10.1007/s00221-012-3322-2

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*, 18–25. doi:10.1016/j.actpsy.2008.04.002

Helton, W. S., Warm, J. S., Tripp, L. D., Matthews, G., Parasuraman, R., & Hancock, P. A. (2010). Cerebral lateralization of vigilance: A function of task difficulty. *Neuropsychologia*, *48*, 1683–1688. doi:10.1016/j.neuropsychologia.2010.02.014

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ.

Klayman, J. (1988). Cue discovery in probabilistic environments: Uncertainty and experimentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 317. doi:10.1037//0278-7393.14.2.317

Klein, G. A. (2008). Naturalistic Decision Making. *Human Factors*, *50*, 456–460. doi:10.1518/001872008X288385.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 30, pp. 576–580). doi:10.1177/154193128603000616

Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin*, *139*, 870–900. doi:10.1037/a0030694

Lansdale, M., Underwood, G., & Davies, C. (2010). Something Overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, *24*, 213–225. doi:10.1002/acp.1552

Lee, S. J., Kim, J., & Jang, S. C. (2011). Human error mode identification for NPP main control room operations using soft controls. *Journal of Nuclear Science and Technology*, *48*, 902–910. doi:10.1080/18811248.2011.9711776

Loveday, T., Wiggins, M. W., Festa, M., Schell, D., & Twigg, D. (2013). Pattern recognition as an indicator of diagnostic expertise. In P. L. Carmona, J. S. Sanchez, & A. L. N. Fred (Eds.), *Pattern Recognition - Application and Methods* (Vol. 204, pp. 1–11). Berlin: Springer. doi:10.1007/978-3-642-36530-0

Loveday, T., Wiggins, M. W., & Searle, B. J. (2013). Cue utilization and broad indicators of workplace expertise. *Journal of Cognitive Engineering and Decision Making*, *8*, 98–113. doi:10.1177/1555343413497019

Loveday, T., Wiggins, M. W., Searle, B. J., Festa, M., & Schell, D. (2013). The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*, 125–137. doi:10.1177/0018720812448475

Mackworth, N. H. (1948). The breakdown of vigilance durning prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*, 6–21.

doi:10.1080/17470214808416738

MacLean, K. A., Aichele, S. R., Bridwell, D. A., Mangun, G. R., Wojciulik, E., & Saron, C. D. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception, & Psychophysics*, *71*, 1042–1058. doi:10.3758/app.71.5.1042

McCalley, J., Asgarpoor, S., Bertling, L., Billinion, R., Chao, H., Chen, J., … Grigg, C. (2004). Probabilistic security assessment for power system operations. In *Power Engineering Society General Meeting, 2004. IEEE* (pp. 212–220). IEEE.

McCormack, C., Wiggins, M. W., Loveday, T., & Festa, M. (2014). Expert and competent non-expert visual cues during simulated diagnosis in intensive care. *Frontiers in Psychology*, *5*, 1–7. doi:10.3389/fpsyg.2014.00949

Morrison, B. W., & Morrison, N. (2015). Diagnostic Cues in Major Crime Investigation. In M. W. Wiggins & T. Loveday (Eds.), *Diagnostic Expertise in Organizational Environments* (pp. 91–98). Surrey, England: Ashgate Publishing Limited, Surrey, England.

Morrison, B. W., Wiggins, M. W., Bond, N. W., & Tyler, M. D. (2013). Measuring relative cue strength as a means of validating an inventory of expert offender profiling cues. *Journal of Cognitive Engineering and Decision Making*, *7*, 211–226. doi:10.1177/1555343412459192

Müller, S., Abernethy, B., & Farrow, D. (2006). How do world-class cricket batsmen anticipate a bowler’s intention? *The Quarterly Journal of Experimental Psychology*, *59*, 2162–2186. doi:10.1080/02643290600576595

Nachreiner, F., Nickel, P., & Meyer, I. (2006). Human factors in process control systems: The design of human-machine interfaces. *Safety Science*, *44*, 5–26. doi:10.1016/j.ssci.2005.09.003

Noyes, J., & Bransby, M. (2001). *People in control: human factors in control room design*. London, UK: IET.

O'Hara, J. M., & Hall, R. E. (1992). Advanced control rooms and crew performance issues: implications for human reliability. *IEEE Transactions on Nuclear Science*, *39*, 919–923. doi:10.1109/23.159733

Ong, M., Russell, P. N., & Helton, W. S. (2013). Frontal cerebral oxygen response as an indicator of initial attention effort during perceptual learning. *Experimental Brain Research*, *229*, 571–578. doi:10.1007/s00221-013-3619-9

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*, 924–927. doi:10.1126/science.472714

Pashler, H., & Wagenmakers, E. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, *7*, 528–530. doi:10.1177/1745691612465253

Pauley, K., O'Hare, D., & Wiggins, M. W. (2009). Measuring expertise in weather-related aeronautical risk perception: The validity of the Cochran–Weiss–Shanteau (CWS) Index. *The International Journal of Aviation Psychology*, *19*, 201–216. doi:10.1080/10508410902979993

Perry, N. C., Wiggins, M. W., Childs, M., & Fogarty, G. (2012). Can reduced processing decision support interfaces improve the decision-making of less-experienced incident commanders? *Decision Support Systems*, *52*, 497–504. doi:10.1016/j.dss.2011.10.010

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903. doi:10.1037/0021-9010.88.5.879

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211–219). Hershey, PA: IGI Publishing. doi:10.4018/978-1-59140-562-7.ch034

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge university press.

Reason, J. (2000). Human error: Models and management. *BMJ*, *320*, 768–770. doi:10.1136/bmj.320.7237.768

Scherer, R., Lee, F., Schlogl, A., Leeb, R., Bischof, H., & Pfurtscheller, G. (2008). Toward self-paced brain–computer communication: navigation through virtual worlds. *IEEE Transactions on Biomedical Engineering*, *55*, 675–682. doi:10.1109/tbme.2007.903709

Schooler, J. W. (2014). Metascience could rescue the 'replication crisis.' *Nature News*, *515*, 9–9. doi:10.1038/515009a

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50*, 864–878. doi:10.1037/e578082012-006

Schvaneveldt, R. W., Beringer, D. B., & Lamonica, J. A. (2001). Priority and organization of information accessed by pilots in various phases of flight. *The International Journal of Aviation Psychology*, *11*, 253–280. doi:10.1037/e406222004-001

Shaw, T. H., Funke, M. E., Dillard, M., Funke, G. J., Warm, J. S., & Parasuraman, R. (2013). Event-related cerebral hemodynamics reveal target-specific resource allocation for both "go" and "no-go" response-based vigilance tasks. *Brain and Cognition*, *82*, 265–273. doi:10.1016/j.bandc.2013.05.003

Shaw, T. H., Satterfield, K., Ramirez, R., & Finomore, V. (2013). Using cerebral hemovelocity to measure workload during a spatialised auditory vigilance task in novice and experienced observers. *Ergonomics*, *56*, 1251–1263. doi:10.1080/00140139.2013.809154

Small, A. J., Wiggins, M. W., & Loveday, T. (2014). Cue-based processing capacity, cognitive load and the completion of simulated short-duration vigilance tasks in power transmission control. *Applied Cognitive Psychology*, *28*, 481–487. doi:10.1002/acp.3016

Smit, A. S., Eling, P. A. T. ., & Coenen, A. M. . (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica*, *115*, 35–42. doi:10.1016/j.actpsy.2003.11.001

Stanton, N. A., Salmon, P., Walker, G. H., & Jenkins, D. (2009). *Human factors in the design and evaluation of central control room operations*. CRC Press. doi:10.1201/9781439809921

Sturman, D., Wiggins, M. W., Auton, J. C., & Loft, S. (2019). Cue utilization differentiates resource allocation during sustained attention simulated rail control tasks. *Journal of Experimental Psychology: Applied*, Advance online publication. doi:10.1037/xap0000204

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*, 82–96. doi:10.1177/1745691614556681

Visser, T., Tichon, J. G., & Diver, P. (2012). Reducing the dangers of operator distraction through simulation training. *Asia-Pacific Simulation and Training Conference and Exhibition*, 18–21.

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

Weiss, D. J., & Shanteau, J. (2003). Empirical assessment of expertise. *Human Factors*, *45*, 104–116. doi:10.1518/hfes.45.1.104.27233

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickers (Ed.), *Attention and Performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159–177. doi:10.1080/14639220210123806

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. Psychology Press.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology*, *25*, 229–235. doi:10.1002/acp.1668

Wiggins, M. W. (2014). Differences in situation assessments and prospective diagnoses of simulated weather radar returns amongst experienced pilots. *International Journal of Industrial Ergonomics*, *44*, 18–23. doi:10.1016/j.ergon.2013.08.006

Wiggins, M. W. (2015). Diagnosis and Instructional Systems Design. In M. W. Wiggins & T. Loveday (Eds.), *Diagnostic Expertise in Organizational Environments* (pp. 69–80). Surrey, England: Ashgate Publishing Limited, Surrey, England.

Wiggins, M. W., Brouwers, S., Davies, J., & Loveday, T. (2014). Trait-based cue utilization and initial skill acquisition: Implications for models of the progression to expertise. *Frontiers in Psychology*, *5*, 1–8. doi:10.3389/fpsyg.2014.00541

Wiggins, M. W., Loveday, T., & Auton, J. C. (2015). *EXPERT Intensive Skills Evaluation (EXPERTise) Test*. Sydney: Macquarie University.

Wiggins, M. W., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental*

Psychology: Applied, *1*, 305–320. doi:10.1037/1076-898X.1.4.305

Wiggins, M. W., & O'Hare, D. (2003a). Expert and novice pilot perceptions of static in-flight images of weather. *International Journal of Aviation Psychology*, *13*, 173–187. doi:10.1207/S15327108IJAP1302_05

Wiggins, M. W., & O'Hare, D. (2003b). Weatherwise: Evaluation of a Cue-Based Training Approach for the Recognition of Deteriorating Weather Conditions during Flight. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*, 337–345. doi:10.1518/hfes.45.2.337.27246

Wiggins, M. W., Stevens, C., Howard, A., Henley, I., & O'Hare, D. (2002). Expert, intermediate and novice performance during simulated pre-flight decision-making. *Australian Journal of Psychology*, *54*, 162–167. doi:10.1080/00049530412331312744

Williams, A. M., Ward, P., Knowles, J. M., & Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied*, *8*, 259–270. doi:10.1037/1076-898X.8.4.259

Bridging Section 3

Study 4 was designed to examine whether qualified practitioners' cue utilisation is predictive of their sustained attention performance in an operational context. The primary differences from Studies 1 and 2 were the recruitment of qualified DNSP system operators, rather than novice operators, and the use of context-based assessments of cue utilisation, whereby cue utilisation and task performance are both assessed within the same domain. This allowed the relationship between cue utilisation and sustained attention to be evaluated amongst operators who had had prior opportunities to acquire cue-based associations through exposure to the operational environment. It was anticipated that operators with higher cue utilisation would demonstrate greater sustained attention during the visual search task, compared to operators with lower cue utilisation.

The results from Study 4 indicated that system operators classified with higher cue utilisation, recorded shorter mean response latencies during the sustained visual search task, compared to those operators who were classified with lower cue utilisation. This outcome supports the proposition that qualified operators' cue utilisation is associated with lower cognitive load during sustained attention tasks. However, like the rail control tasks used in Studies 1 and 2, the sustained visual search task relied on a simplified representation of the operating system, containing relatively static and predictable task features. Consequently, the extent to which the findings from Study 4 can be generalised to complex and dynamic operating systems is limited.

In more dynamic operational tasks, such as driving, emerging features are likely to be less predictable compared to the repetitious control tasks used in Studies 1 to 4. Therefore, Study 5 was designed to examine whether cue utilisation differentiates qualified drivers' cognitive resource consumption during a simulated driving task. As the simulated driving task contains a greater number of less predictable, emerging features compared to the sustained attention tasks used in Studies 1 to 4, reduced fixations in the absence of any other strategies may be an inefficient search strategy, as this would reduce the number of emerging features to which an operator could attend. Consequently, it was reasoned that purposive sampling, whereby operators visually fixate on key features within close proximity (Henderson, 2003; Underwood et al., 2003), would be a more efficient strategy during the dynamic scenario, as it would reduce the time and effort spent scanning for emerging features. It was predicted that, during the simulated driving task, participants with higher cue utilisation would record lower visual fixation rates, smaller mean saccade amplitudes, smaller mean fixation dispersions, and lower increases in cerebral oxygenation from baseline, compared to participants with lower cue utilisation.

STUDY 5

Publication history

The paper arising from Study 5 was entitled, "Drivers' Cue Utilisation Predicts Cognitive Resource Consumption during a Simulated Driving Scenario". A version of this paper has been submitted to the *Journal of Applied Research in Memory and Cognition*, and is currently under review. The author of the present dissertation wrote approximately 80% of this paper.

Drivers' Cue Utilisation Predicts Cognitive Resource Consumption during a Simulated Driving Scenario

Daniel Sturman[1] and Mark W. Wiggins[1]

[1] *Macquarie University, NSW, Australia*

Acknowledgements

Support for this research was provided by the Australian Research Council [ARC LP160100211].

Abstract

This study was designed to examine whether cue utilisation differentiates drivers' consumption of cognitive resources during a simulated driving task. Participants were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of driving. During the simulated driving task, participants with higher cue utilisation demonstrated smaller mean visual saccade amplitudes, smaller mean fixation dispersions, smaller increases in cerebral oxygenation and recorded fewer missed traffic signals, compared to participants with lower cue utilisation. These results suggest that drivers with higher cue utilisation experienced lower cognitive load during the simulated driving task, while maintaining a higher level of performance, compared to drivers with lower cue utilisation. The results provide support for the assertion that, amongst qualified operators, a relatively greater capacity for cue utilisation is associated with the consumption of fewer cognitive resources during operational tasks.

Drivers' Cue Utilisation Predicts Cognitive Resource Consumption during a Simulated Driving Scenario

Despite increasing levels of automation, human operators remain necessary for many high risk operating systems, including motor vehicles (Norman, 2015; Wickens, Hollands, Banbury, & Parasuraman, 2015). To maintain a sufficiently high level of performance, these operators are required to allocate cognitive resources to monitoring and controlling the operating systems (Luck, Hillyard, Mouloua, & Hawkins, 1996; Waller, Gupta, & Giambatista, 2004; Wickens, 2002). However, attentional resource theories posit that cognitive resources are drawn from a limited source, and that the allocation of cognitive resources to a task can result in fewer residual resources than are necessary to remain alert and respond quickly and accurately to deviations in the system state (Kahneman, 1973; Parasuraman, 1979; Wickens, 2002).

Expert performance is associated with the capacity to sustain rapid and accurate performance for extended periods (Farrington-Darby & Wilson, 2006; Sherbino et al., 2012). This level of expertise derives from the gradual development of specialised associations and routines, which are retained in memory, and which can be activated rapidly in the absence of conscious processing (Ericsson & Lehmann, 1996; Ericsson & Towne, 2010; Salthouse, 1991). The activation of these specialised associations reduces the demands on working memory resources (Chung & Byrne, 2008).

Associations between situation-specific environmental features and task events are referred to as cues (Brunswik, 1955; Klein, Calderwood, & Clinton-Cirocco, 2010). Cue utilisation is the application of cue-based processing (Lansdale, Underwood, & Davies, 2010). The capacity for cue utilisation relies upon the identification of predictive features in the environment, the association between these features and events in memory, the retention of these cue-based associations, and the appropriate application of cues in response to environmental features (Wiggins, 2012, 2015).

The effective utilisation of cue-based associations leads to efficiencies in information processing (Brouwers, Wiggins, Griffin, Helton, & O'Hare, 2017; Brouwers, Wiggins, Helton, O'Hare, & Griffin, 2016; Lansdale et al., 2010). The activation and retrieval of cues from long-term memory has the advantage of imposing relatively fewer demands on working memory resources (Chung & Byrne, 2008; Evans & Fendley, 2017). Consequently, higher cue utilisation is associated with lower perceived cognitive demands (Brouwers et al., 2016). Further, the rapid identification of cues is associated with a greater rate of skill acquisition and improvements in performance (Brouwers et al., 2017; Perry, Wiggins, Childs, & Fogarty, 2013).

Individual differences in cue utilisation are based on the capacity of individuals to recognise features with greater or lesser predictive validity, allowing them to isolate predictive features and disregard features with little predictive validity (Wiggins, 2015). However, assessments of cue utilisation based on the identification of specific features are difficult, as the predictive validity of features tends to be idiosyncratic (Wiggins, 2012; Wigton, 1996; Yuki, Maddux, & Masuda, 2007). For example, different patterns of cues may be used by two different experts to form the same correct diagnosis (Wigton, 1996). Consequently, due to the subjective nature of classifying features as having greater or lesser predictive validity, an alternative approach to measuring cue utilisation involves assessing behaviour that is indicative of the utilisation of cues.

Domain-related cue utilisation has been assessed using the EXPERT Intensive Skills Evaluation (EXPERTise; Wiggins et al., 2015) situation judgment test (Brouwers et al., 2016, 2017; Small, Wiggins, & Loveday, 2014; Sturman et al., 2019). EXPERTise is an online, composite assessment tool, incorporating experimental tasks which assess distinct but complementary aspects of behaviour that reflect the utilisation of cues (Loveday, Wiggins, Harris, O'Hare, & Smith, 2013; Wiggins, Azar, Hawken, Loveday, & Newman, 2014). Tasks are selected to represent features to which operators must attend in the operational context, comprising feature identification, feature recognition, feature association, feature discrimination, and feature prioritisation (Loveday, Wiggins, Harris, et al., 2013; Wiggins, Loveday, & Auton, 2015). Using standardised task scores, cluster analyses can be employed to classify participants by cue utilisation typologies that correspond to higher or lower levels of cue utilisation (Loveday, Wiggins, Harris, et al., 2013; Wiggins, Whincup, & Auton, 2018). Typically, there is a dissociation between years of experience in a domain and cue utilisation typology (Loveday, Wiggins, Searle, Festa, & Schell, 2013).

Context-based assessments of cue utilisation, whereby cue utilisation and task performance are assessed in the same domain, differentiate performance and decision making amongst experienced personnel (Loveday, Wiggins, & Searle, 2013; Small et al., 2014; Wiggins, Azar, et al., 2014). For instance, Loveday et al. (2013) observed that higher cue utilisation amongst software engineers was associated with superior error management when developing solutions to software engineering problems. Small et al. (2014) demonstrated that, amongst qualified power control room operators, those participants with higher cue utilisation became familiar with a novel, domain-related short vigilance task more rapidly than those participants with lower cue utilisation. Further, Wiggins et al. (2014) noted that levels of cue utilisation differentiated pilots' decisions to divert or continue a flight as planned during deteriorating weather conditions.

Cross-task cue utilisation, whereby cue utilisation evaluated in one context is used to predict performance in another context, has also differentiated accuracy, response latency, the capacity to sustain attention, and skill acquisition during novel tasks (Brouwers et al., 2016, 2017; Wiggins, Brouwers, Davies, & Loveday, 2014). For instance, Wiggins et al. (2014) demonstrated that a general capacity for cue utilisation was associated with greater improved accuracy in landing a simulated aircraft following a limited number of trials. Similarly, higher cue utilisation assessed in the driving domain, is associated with lower perceptions of cognitive load, and a greater capacity to maintain rapid and accurate responses to critical targets during sustained attention rail control tasks (Brouwers et al., 2016, 2017; Sturman, Wiggins, Auton, & Loft, 2019).

Differences in sustained attention-related performance based on cue utilisation are posited to derive from the capacity of individuals with higher cue utilisation to recognise and utilise predictive features, allowing them to adopt strategies that reduce the consumption of cognitive resources during a task (Loveday, Wiggins, Festa, Schell, & Twigg, 2013; Schriver, Morrow, Wickens, & Talleur, 2008). Consequently, these individuals should have greater residual cognitive resources, allowing them to sustain rapid and accurate responses for longer periods of time, compared to individuals who consume greater cognitive resources during the task.

Brouwers et al. (2017) provide evidence to support differences in the recognition and utilisation of predictive features based on cue utilisation, demonstrating that participants with higher cue utilisation were significantly more likely to recognise an implicit pattern in a sustained attention rail control task, which in turn was associated with significantly lower response latencies to critical signals. Presumably, by recognising the implicit pattern, individuals with higher cue utilisation were able to anticipate events by attending to the most predictive features, thereby reducing the cognitive resources required to maintain performance (Brouwers et al., 2017). However, as this evidence relies primarily on self-reports and inferences derived from response latencies during sustained attention tasks, it could potentially be explained by alternative factors, such as participants' level of motivation or engagement.

To establish whether participants with higher cue utilisation consume fewer attentional resources during sustained attention tasks, Sturman et al. (2019) examined the relationship between cue utilisation and cognitive load using eye tracking data to assess fixation rates, and

Functional Near Infrared Spectroscopy (fNIRS) to measure cerebral oxygenation in the prefrontal cortex. During repetitious, and therefore predictable tasks, the identification of key features was predicted to lead to a reduction in the frequency of fixations which are presumed indicative of more efficient search patterns and the consumption of fewer cognitive resources during the task (Ikehara & Crosby, 2005; Poole & Ball, 2006; Salvucci & Goldberg, 2000). More efficient search patterns should reduce the cognitive resources required for a task, and therefore be associated with smaller increases in cerebral oxygenation in the prefrontal cortex (Helton et al., 2010; Shaw et al., 2009).

Sturman et al. (2019) observed that, compared to participants with lower cue utilisation, participants who demonstrated higher cue utilisation in the context of automobile driving, recorded smaller increases in cerebral oxygenation in the prefrontal cortex and greater decreases in fixation rates during a sustained attention rail control task containing predictable features. Furthermore, participants with higher cue utilisation demonstrated smaller increases in response latency over time, indicating that these individuals had sufficient residual cognitive resources to remain alert and respond quickly. These findings provide further support for the proposition that higher cue utilisation is associated with lower cognitive load during tasks containing predictable features (Brouwers et al., 2016; Perry, Wiggins, Childs, & Fogarty, 2012; Wiggins, 2011).

Evidence to support differences in the rate at which cognitive resources are consumed has typically relied on cross-task measures cue-utilisation, whereby cue utilisation evaluated in one context is used to predict performance in another context (e.g., Brouwers et al., 2017; Sturman et al., 2019). For instance, Sturman et al. (2019) relied on general cue utilisation assessed in the driving domain to predict performance during a novel rail control task. Consequently, differences in response latency and the consumption of cognitive resources based on cue utilisation likely reflect differences in the speed with which participants acquire new cue-based relationships during novel tasks. However, it is unclear whether these differences would also be observed with qualified system operators who differ from novice operators in that they have had prior opportunities to acquire relevant cues. While higher cue utilisation should be associated with a greater frequency of previously acquired cues controlling for years of experience (McCormack, Wiggins, Loveday, & Festa, 2014; Morrison, Wiggins, Bond, & Tyler, 2013), it remains unclear whether differences in the rate at which cognitive resources are consumed based on cue utilisation are also evident amongst qualified operators.

Dynamic operational environments

In dynamic tasks, such as driving, emerging features are likely to be less predictable, compared to repetitious control tasks. During tasks containing constantly emerging features, reduced fixations in the absence of any other strategies may be inefficient, as this would reduce the number of emerging features to which an operator could attend. To reduce cognitive load in less predictable dynamic environments, purposive sampling may be undertaken, whereby operators visually fixate on key features within close proximity (Henderson, 2003; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003). This could increase efficiency by reducing the time and effort spent scanning for emerging features (Watson, Brennan, Kingstone, & Enns, 2010).

While fixation rates provide an indication of cognitive load and search pattern efficiencies in predictable environments, additional eye behaviour metrics can provide a more detailed measure of search patterns, particularly in dynamic environments such as driving. For instance, when assessing hazard detection during driving tasks, researchers often examine a range of eye behaviour metrics, including saccade amplitude and fixation dispersion (Chapman & Underwood, 1998; Grüner & Ansorge, 2017; Underwood, Crundall, & Chapman, 2011). Saccade amplitude and fixation dispersion in particular can also be used to assess workload, or the consumption of cognitive resources during tasks (Camilli, Terenzi, & Di Nocera, 2007; Di Nocera, Camilli, & Terenzi, 2007; Fu et al., 2011).

Saccade amplitude refers to the change in the degrees of visual angle from the pre-saccade fixation to the post-saccade fixation. Greater performance in the detection of critical targets has been associated with search strategies involving saccades of smaller amplitude (Bertram, Helle, Kaakinen, & Svedström, 2013). In a driving simulation study, Underwood et al. (2011) observed that saccadic amplitude reduced when approaching hazards, presumably as drivers are more likely to visually track, and concentrate multiple fixations around the potential hazard. Consequently, individuals who identify key areas of concern more rapidly, should demonstrate saccades of smaller amplitude.

Fixation dispersion is the extent to which fixations are spread out versus focused while completing a task. Greater fixation dispersion has been associated with greater workload during simulated flight (Di Nocera et al., 2007) and during video game tasks (Camilli et al., 2007). For instance, Di Nocera et al. (2007) observed that fixation dispersion was greater during more demanding sections of a flight. Camilli et al. (2007) further demonstrated concurrent validity between fixation dispersion and a subjective measure of workload. Therefore, participants who perceive fewer cognitive demands during a task should also demonstrate smaller mean fixation dispersions.

The aim of the present study was to establish whether qualified operators' cue utilisation is predictive of their cognitive resource consumption during a dynamic sustained attention task. To recruit a sufficient number of qualified operators, cue utilisation and cognitive resource consumption were both assessed within the context of driving, as a significant proportion of the adult population possess driving experience. Based on the proposition that cue utilisation is associated with the capacity to draw from patterns in memory to respond to potential sources of threat in the environment, drivers with relatively higher cue utilisation were expected to spend relatively more time attending to specific areas of the environment and visually tracking features of greater concern. Further, as the activation and retrieval of cues from long-term memory has the advantage of imposing relatively fewer demands on working memory resources, participants with higher cue utilisation consume fewer cognitive resources during the driving task. Consequently, these participants should have greater residual cognitive resources, allowing them to maintain a high level of driving performance. Specifically, it was hypothesised that, during a simulated driving task, participants with higher cue utilisation would record lower visual fixation rates, smaller mean saccade amplitudes, smaller mean fixation dispersions, and lower increases in cerebral oxygenation from baseline, compared to participants with lower cue utilisation. Further, it was hypothesised that participants with higher cue utilisation would record fewer missed traffic signals and fewer speed exceedances during the simulated driving task, compared to participants with lower cue utilisation.

Material and Methods

Design

The experiment comprised a 2 x 4 mixed factorial design incorporating two levels of cue utilisation (higher and lower) as a between-subjects factor and four levels of distance travelled (Block 1, Block 2, Block 3, and Block 4) as a within-subjects factor. Participants were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of driving. The blocks comprised four, three-kilometre quartiles of the simulated driving task, which enabled sustained attention to be assessed through changes in the dependent variables as a function of distance driven.

Participants

Sixty-two undergraduate university students (42 females and 20 males) were recruited and received course credit in return for participation. The participants ranged in age from 18 to 22 years (M = 19.16, SD = 1.20), and ranged in driving experience from 24 to 60 months (M =

37.7, *SD* = 10.8). Participants consisted of motor vehicle drivers who were qualified to drive a vehicle without supervision by a licensed driver. The inclusion criteria comprised motor vehicle drivers with a minimum of two years experience driving a class-C motor vehicle and a maximum age of 22 years. These selection criteria were imposed to control for exposure to driving, which enabled comparative assessments of cue utilisation.

Instruments

The participants were asked to indicate their age, sex, years of driving experience and hours of daily driving. Cue utilisation was assessed using the Expert Skills Evaluation (EXPERTise 2.0) software platform (Wiggins, Loveday, & Auton, 2015).

EXPERTise

EXPERTise 2.0 software assesses aspects of cue utilisation within a specific domain or context. The driving edition of EXPERTise employed by Brouwers et al. (2016) was used in the present study as it assesses the acquisition of cues in the context of driving. EXPERTise tasks include a Feature Identification Task (FIT), a Feature Recognition Task (FRT), a Feature Association Task (FAT), a Feature Discrimination Task (FDT), and a Feature Prioritisation Task (FPT).

The FIT consists of 21 trials, during which participants are required to identify key features from photographs of a road as viewed from the driver's seat of a car. Participants are presented with one image at a time and asked to use a computer mouse to identify, as quickly as possible, the area of greatest concern in the scene (e.g. pedestrian or ball on the road). Response latency is recorded as the time in milliseconds from the presentation of the stimulus to selection of an area of concern. The FIT assesses participants' capacity to identify and utilise more diagnostic visual features in the environment, with a lower mean response latency indicative of higher cue utilisation (Brouwers et al., 2016; Schriver et al., 2008).

During the FRT participants are presented with images of a road for 500 ms, after which the image is removed and participants are asked to estimate the speed limit of the road from one of four options (50–60, 70–80, 90–100 or 110+ km/h). One image is presented at a time, with a total of 18 trials. The FRT assesses participants' capacity to rapidly extract information from a scene, and their ability to utilise this information to form an accurate judgement. A greater number of correct judgements is indicative of higher cue utilisation (Brouwers et al., 2017; Wiggins & O'Hare, 2003).

The FAT is designed to assess participants' capacity to rapidly assess the relatedness of features in their environment. Over 30 trials, participants are presented with driving-related terms (e.g., 'pedestrian crossing', 'breaking'). Each trial consists of two terms which are displayed for 1500 milliseconds, after which the terms are removed. Participants are then asked to rate the extent to which the terms are related using a six-point scale, from 1 (*Extremely unrelated'*) to 6 (*Extremely related*). FAT scores are calculated by dividing the variance in scores by the mean reaction time. Higher cue utilisation is associated with a greater ratio of variance to reaction time (Morrison et al., 2013; Schvaneveldt, Beringer, & Lamonica, 2001).

During the FDT, participants are required to read a short driving scenario, and to make a decision based on their typical response to similar scenarios. Following their decision, participants are presented with a list of 14 features and asked to rate the relative importance of each feature to their decision. Each feature is rated on a 10-point scale ranging from 1 (*Not important at all*) to 10 (*Extremely important*). As cue utilisation requires individuals to identify features as more or less relevant, a greater variance of ratings is indicative of higher cue utilisation (Brouwers et al., 2017; Pauley, O'Hare, & Wiggins, 2009).

The FPT is designed to assess participants' capacity to prioritise the acquisition of information from the environment (Wiggins & O'Hare, 1995; Wiggins, Stevens, Howard, Henley, & O'Hare, 2002). During the FPT, participants are presented with a way-finding scenario, and asked to select one of five possible modes of transport. Accompanying the scenario is a list of 17 drop-down menus, which are feature-labelled (e.g., 'Closest Bus Stop, 'Current Traffic Conditions). Upon selection, each menu provides information pertaining to the scenario. Only one drop-down menu can be selected at a time, and participants are given two minutes to access the menus prior to making their decision. Lower cue utilisation is typically associated with the selection of information in the sequential order in which it is presented (e.g., from top to bottom as presented on the computer screen), while higher cue utilisation is associated with a lower ratio of sequentially selected items (Wiggins et al., 2002).

Simulated driving task

The simulated driving task was conducted using the 'STISIM' (version 8, model 100) driving simulator, which includes a steering wheel, a brake and an accelerator set in front of three 22-inch computer screens that mimic the view through a windscreen, allowing a 135-degree field of view. Displays on both side screens simulated side-view mirrors while a rear-view mirror was present at the top of the centre screen. The experimental trial consisted of a 12-kilometre drive designed to simulate typical urban driving conditions, including regular traffic lights, stop signs and pedestrian crossings. For the purpose of data analysis, the drive consisted of four, three-kilometre blocks, which were each matched for traffic conditions and the number of vehicles and pedestrians. Each block contained three sets of traffic lights, two pedestrian crossings and two stop signs. The experimental trial was designed to simulate a moderate workload task. Adhering to the road rules, the drive took approximately 20 minutes for participants to complete.

Performance was assessed by the number of missed traffic signals and the speed exceedances recorded during each block of the simulated driving task. A missed traffic signal included any failure to stop before a red light, stop sign, or pedestrian crossing when required. Speed exceedances were recorded each time that participants exceeded the speed limit by more than one km/h.

Eye tracking

Prior to the simulated driving task, participants were fitted with SMI Eye Tracking Glasses (version 2) using the system's standard operating procedures, including a three-point calibration. The system has a $60° \times 46°$ recording visual angle, a resolution of 1280 x 960 pixels, and a sampling frequency of 60 Hz. Eye tracking data were recorded for the duration of the simulated driving task, and later analysed using BeGaze software. Eye behaviour metrics consisted of fixation rates (mean number of fixations made per minute), saccade amplitude (mean change in degrees of the visual angle per saccade), and fixation dispersion (mean dispersion of fixations in degrees). The velocity threshold for a saccade was set at 40 deg/sec, and therefore smooth pursuit of objects in the dynamic environment, which typically have a velocity less than 30 deg/sec (Berg, Boehnke, Marino, Munoz, & Itti, 2009; Robinson, 1965), did not contribute to measures of saccade amplitude. Similarly, BeGaze algorithms did not record multiple fixations during smooth pursuit (SensoMotoric Instruments, Teltow, Germany).

Near Infrared Spectroscopy (NIRS)

Prior to completing the simulated driving task, participants were fitted with a Portalite Near Infrared Spectroscope (NIRS) sensor approximately one centimetre above their right eyebrow. The right frontal lobe is typically associated with increased activity during sustained attention tasks (Ong, Russell, & Helton, 2013), and was therefore selected as the brain region of interest for the current study.

Using light in the near-infrared spectrum, the NIRS system estimates relative levels of oxyhaemoglobin and deoxyhaemoglobin in the cerebral tissue. Cerebral oxygenation is assessed using regional oxygen saturation (rSO2), which is calculated as the ratio of oxyhaemoglobin total haemoglobin (Ekkekakis, 2009; Gratton & Fabiani, 2006). During a five-minute baseline period, participants were asked to sit quietly and minimise body movements. rSO2 during the final minute of the baseline period was used as a baseline index.

Changes in rSO2 relative to baseline were used to assess cerebral activation during the simulated driving task.

Procedure

Following approval from the University Human Research Ethics Committee, participants were tested individually in 90-minute sessions. After reading and signing an information and consent form, participants subsequently completed an online demographic questionnaire, the five EXPERTise 2.0 tasks, and the simulated driving task. Computer prompts directed participants through the demographic questionnaire and the EXPERTise tasks. Standardised instructions were provided verbally for the simulated driving task, with participants instructed to follow the main road, drive as they normally would, and adhere to the speed limit. Following the verbal instructions, participants were fitted with the fNIRS sensor. After the five-minute baseline period, participants completed a 1.3-kilometre practice drive before completing the experimental trial.

Results

Cue utilisation typologies

Expertise data were used to identify cue utilisation typologies that corresponded to higher or lower levels of cue utilisation. Consistent with standardised practice in the calculation of cue utilisation typologies (Brouwers et al., 2017; Sturman et al., 2019), scores for each task were converted to *z* scores. A cluster analysis using the *z* scores was performed to identify two groups. Consistent with previous research (Brouwers et al., 2016; Small et al., 2014), the first cluster, labelled the higher cue utilisation group, consisted of participants who recorded a shorter response latency on the FIT, greater accuracy on the FRT, a higher mean ratio of variance to reaction time on the FAT, a greater variance in ratings on the FDT, and a lower ratio of sequential selections in the FPT. The second cluster, labelled the lower cue utilisation group, consisted of participants who recorded a greater response latency on the FIT, lower accuracy on the FRT, a lower mean ratio of variance to reaction time on the FAT, a lower variance in ratings on the FDT, and a higher ratio of sequential selections in the FPT. Twenty-one participants were classified in the higher cue utilisation group and 41 participants were classified in the lower cue utilisation group. The ratio of participants in the higher and lower cue utilisation groups is consistent with previous research (Brouwers et al., 2016). Independent samples *t* tests demonstrated significant differences in FRT, FAT, FDT and FPT scores between the higher and lower cue utilisation groups (see Table 1).

Table 1. *Cluster centroids for the EXPERTise task scores.*

	Cluster 1 (n=19)	Cluster 2 (n=42)			
	Higher cue utilisation	Lower cue utilisation	*t*	df	*p*
FIT	-0.13	0.07	-0.72	60	.475
FRT	0.75	-0.39	5.01**	60	<.001
FAT	0.34	-0.17	2.21*	60	.031
FDT	0.86	-0.45	6.43**	60	<.001
FPT	-0.69	0.35	-4.44**	60	<.001

* Significant at the 0.05 level (two-tailed); **Significant at the 0.01 level (two-tailed)

Covariates

Independent samples *t* tests indicated that cue utilisation was not related to participants' age, $t(60) = -1.21$, $p = .233$, $d = .34$, years of driving experience, $t(60) = -1.77$, $p = .082$, $d = .49$, or hours of daily driving, $t(60) = 0.48$, $p = .630$, $d = .12$. A chi-square test indicated that cue utilisation was not related to participants' sex, $\chi^2(1) = 0.20$, $p = .657$. A series of ANCOVAs indicated that participants' age, sex, and hours of daily driving were not associated with missed traffic signals, speed exceedances, or rSO2 levels during the driving simulation (all *p*s > .05). The ANCOVA results also indicated that years of driving experience was not associated with fixation rates ($p = .704$), rSO2 levels ($p = .937$), saccade amplitude ($p = .496$), or fixation dispersion ($p = .996$). However, the ANCOVA results revealed positive associations between years of driving experience and missed traffic signals (p = .021) and speed exceedances (p = .049). Consequently, years of driving experience was included as a covariate for the main analyses.

Physiological measures

For each block, fixation rates were calculated as the mean number of eye fixations recorded per minute, saccade amplitude was calculated as the mean change in degrees of the visual angle per saccade, and fixation dispersion was calculated as the mean dispersion of fixations in degrees. Due to technical problems, movement artefacts and/or calibration difficulties, eye behaviour metrics were unable to be calculated accurately for six participants, four of whom were in the lower cue utilisation group. Consequently, data for these six participants were excluded from analyses involving fixation rates, saccade amplitude, and fixation dispersion. Fixation rates, saccade amplitudes, and fixation dispersions were approximately normally distributed for each block and cue utilisation group.

Relative measures of Regional Oxygen Saturation (rSO2) compared to baseline were used to calculate rSO2 scores for each block of the driving simulation. The mean rSO2 during each block was compared to the baseline measure taken prior to the simulation, and scores represent the percentage change in rSO2 from baseline. Therefore, a positive score represented an increase in rSO2 compared to baseline, while a negative score represented a decrease in rSO2 compared to baseline. rSO2 scores were approximately normally distributed for each block and cue utilisation group.

Cue utilisation and eye behaviour metrics

To investigate whether participants' information acquisition differed based on their cue utilisation, three 2 x 4 ANCOVAs were conducted with cue utilisation group as a between-groups variable (Higher, Lower), distance travelled as a within-groups variable (Blocks 1-4), and years of driving experience as a covariate. The dependent variables for the three ANCOVAs comprised fixation rates, saccade amplitude, and fixation dispersion respectively.

With fixation rate as the dependent variable, the assumption of sphericity was violated (p = .001). Therefore, a Greenhouse-Geisser adjustment was used for all within-subjects' effects. There was no statistically significant main effect for cue utilisation, $F(1,53) = 0.44$, $p = .510$, $\eta^2 = .01$, or distance travelled, $F(2.0,107.9) = 0.42$, $p = .700$, $\eta^2 = .01$. No statistically significant interaction was evident between distance travelled and cue utilisation, $F(2.0,107.9) = 0.80$, $p = .454$, $\eta^2 = .02$ (see Figure 1). This indicates that there were no differences in the rate of fixation based on either cue utilisation or the distance travelled.

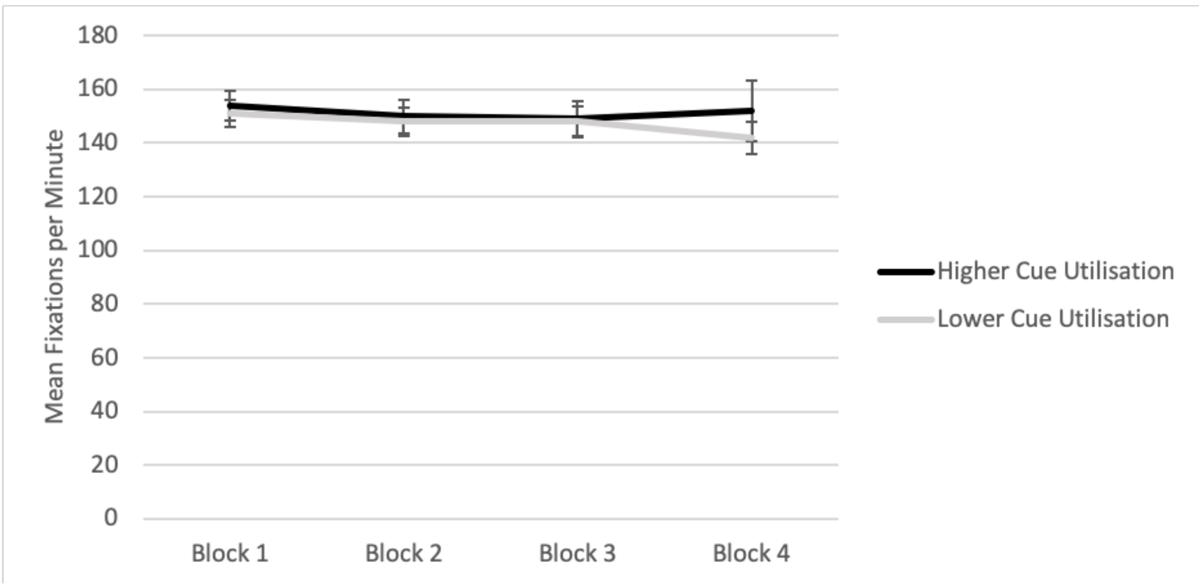

Figure 1. Mean number of eye fixations per minute by cue utilisation typology and distance travelled (represented by blocks). Error bars represent ± 1 SE.

With saccade amplitude as the dependent variable, a statistically significant main effect was evident for cue utilisation, $F(1,53) = 6.86$, $p = .011$, $\eta^2 = .115$, with the lower cue utilisation group recording significantly greater mean saccade amplitudes ($M = 7.76$, $SD = 2.91$) compared to the higher cue utilisation group (*M* = 6.02, *SD* = 1.93). There was no statistically significant main effect for distance travelled, $F(3,159) = 0.50$, $p = .685$, $\eta^2 = .01$, and no statistically significant interaction was evident between distance travelled and cue utilisation, $F(3,159) = 0.15$, $p = .932$, $\eta^2 = .003$ (see Figure 2).

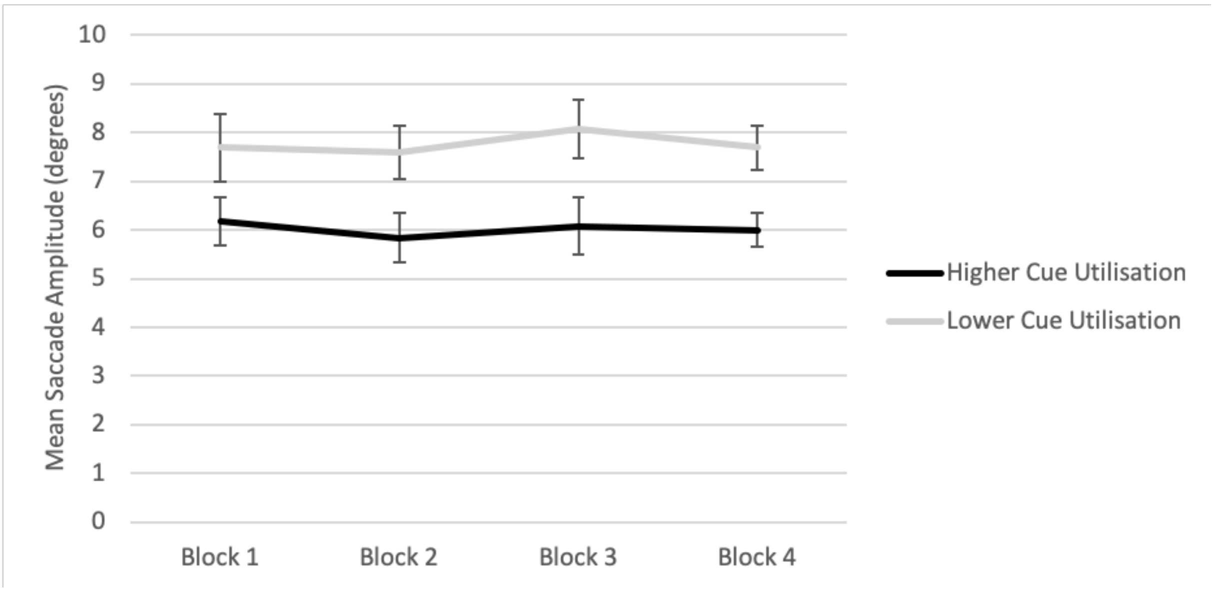

Figure 2. Mean saccade amplitude by cue utilisation typology and distance travelled (represented by blocks). Error bars represent ± 1 SE.

With fixation dispersion as the dependent variable, a statistically significant main effect was evident for cue utilisation, $F(1,53) = 5.02$, $p = .029$, $\eta^2 = .09$, with the lower cue utilisation group recording significantly greater mean fixation dispersions (*M* = 41.9, *SD* = 15.2) compared to the higher cue utilisation group (*M* = 34.3, *SD* = 7.5). There was no statistically significant main effect for distance travelled, $F(3,159) = 0.34$, $p = .769$, $\eta^2 = .01$, and no statistically significant interaction was evident between distance travelled and cue utilisation, $F(3,159) = 1.62$, $p = .187$, $\eta^2 = .03$ (see Figure 3).

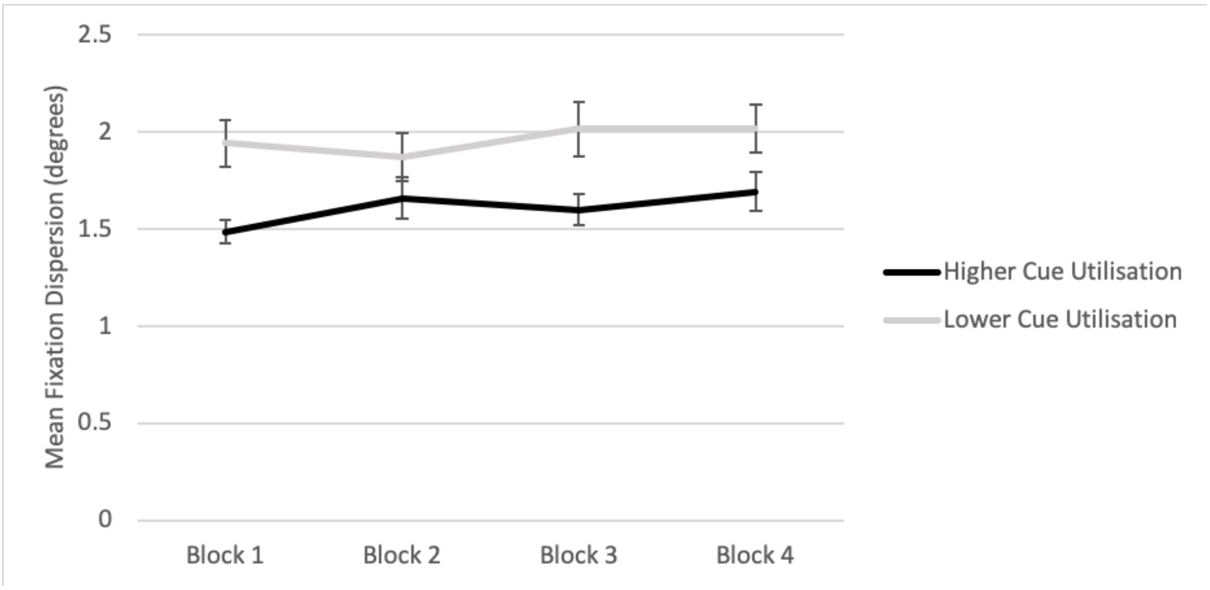

Figure 3. Mean fixation dispersion by cue utilisation typology and distance travelled (represented by blocks). Error bars represent ± 1 SE.

Cue utilisation and rSO2

To investigate whether participants with higher cue utilisation consumed fewer cognitive resources during the simulated driving task, a 2 x 4 ANCOVA was conducted with cue utilisation group as a between-groups variable (Higher, Lower), distance travelled as a within-groups variable (Blocks 1-4), years of driving experience as a covariate, and rSO2 scores as the dependent variable. The assumption of sphericity was violated (p = .002). Therefore, a Greenhouse-Geisser adjustment was used for all within-subjects' effects. A statistically significant main effect was evident for cue utilisation, $F(1,58) = 4.66$, $p = .035$, $\eta^2 = .074$, with the lower cue utilisation group recording significantly greater increases in rSO2 from baseline ($M = 1.32$, $SD = 1.05$) compared to the higher cue utilisation group ($M = 0.68$, $SD = 1.18$). There was no statistically significant main effect for distance travelled, $F(2.4,141.7) = 0.39$, $p = .720$, $\eta^2 = .01$, and no statistically significant interaction between cue utilisation and distance travelled, $F(2.4,141.7) = 0.47$, $p = .662$, $\eta^2 = .01$ (see Figure 4).

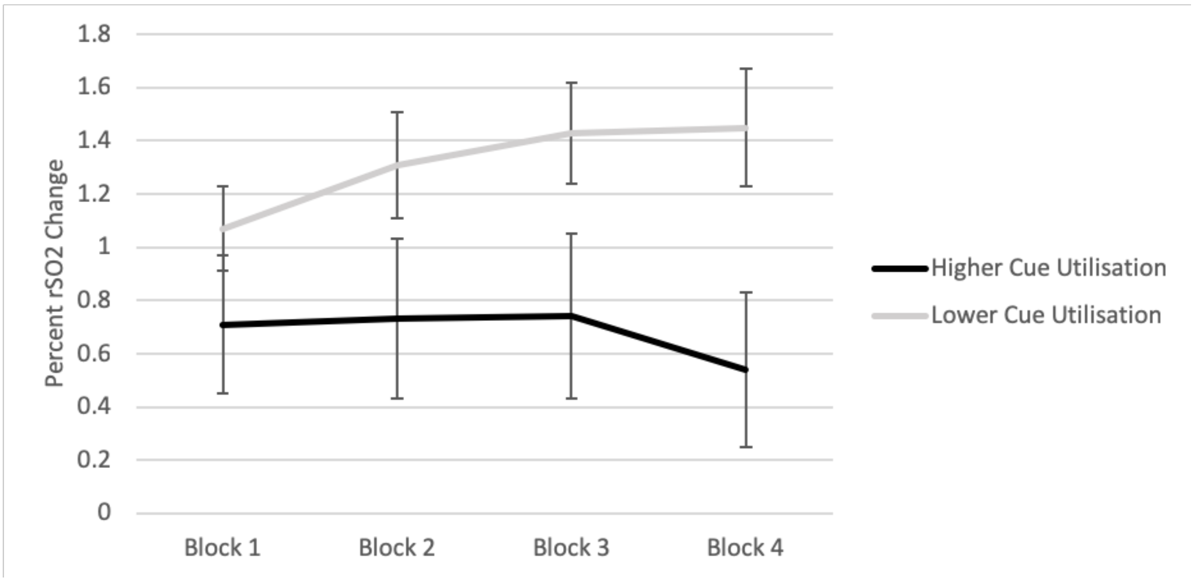

Figure 4. Mean oxygenation scores in the right hemisphere by cue utilisation typology and distance travelled (represented by blocks). Oxygenation scores are based on percent change relative to baseline. Error bars represent ± 1 SE.

Cue utilisation and driving performance

To test whether cue utilisation was associated with performance during the simulated driving task, two 2 x 4 ANCOVAs were conducted with cue utilisation group as a between-groups variable (Higher, Lower), distance travelled as a within-groups variable (Blocks 1-4), and years of driving experience as a covariate. The dependent variables for the two ANCOVAs comprised missed traffic signals and speed exceedances respectively. Missed traffic signals were calculated as the frequency of failures to stop at red lights, stop signs and pedestrian crossings. Speed exceedances were calculated as the frequency with which participants exceeded the speed limit.

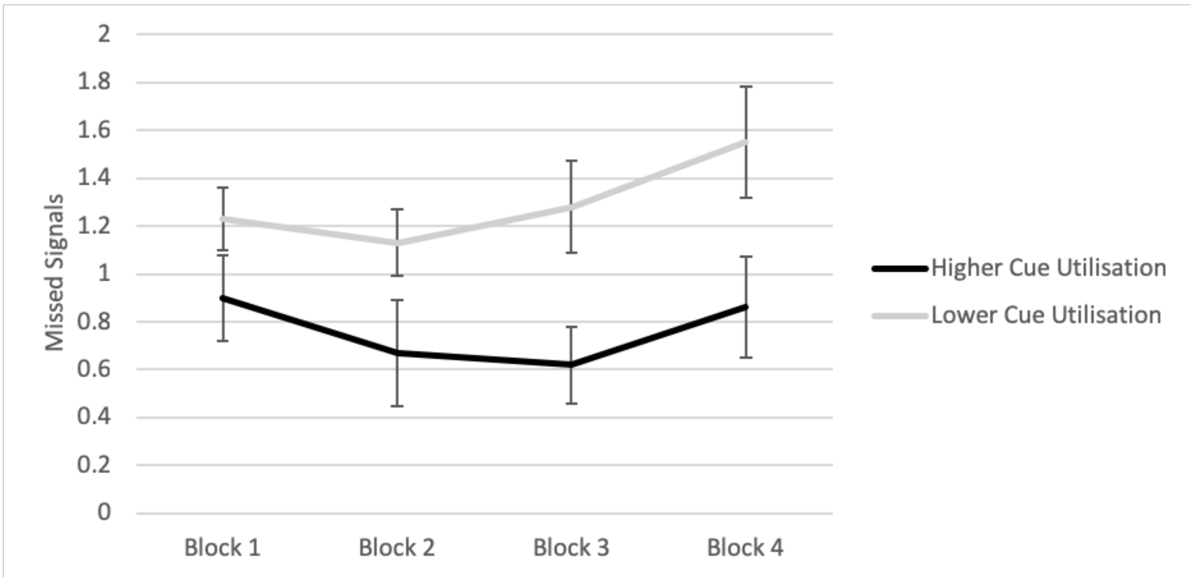

Figure 5. Mean number of missed traffic signals by cue utilisation typology and distance travelled (represented by blocks). Error bars represent ± 1 SE.

With missed traffic signals as the dependent variable, a statistically significant main effect was evident for cue utilisation, $F(1,58) = 4.33$, $p = .042$, $\eta^2 = .07$, with a significantly greater frequency of traffic signals missed in the lower cue utilisation group ($M = 1.29$, $SD = 0.88$) compared to the higher cue utilisation group ($M = 0.76$, $SD = 0.60$). This indicates that participants with higher cue utilisation made fewer errors, on average, compared to participants with lower cue utilisation (see Figure 5). There was no statistically significant main effect for distance travelled, $F(3,174) = 1.12$, $p = .341$, $\eta^2 = .02$, and no statistically significant interaction between cue utilisation and distance travelled, $F(3,174) = 1.14$, $p = .334$, $\eta^2 = .02$.

With speed exceedances as the dependent variable, there was no statistically significant main effect for cue utilisation, $F(1,58) = 0.82$, $p = .369$, $\eta^2 = .01$, nor distance travelled, $F(3,174) = 2.28$, $p = .081$, $\eta^2 = .04$. No statistically significant interaction was evident between distance travelled and cue utilisation, $F(3,174) = 0.17$, $p = .917$, $\eta^2 = .003$ (see Figure 6).

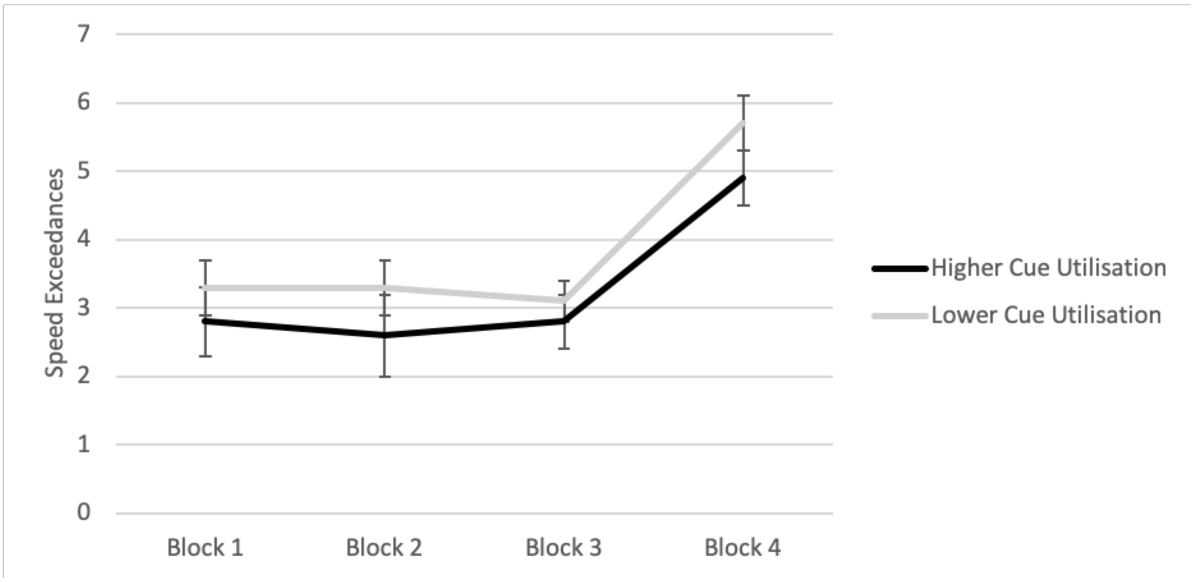

Figure 6. Mean number of speed exceedances by cue utilisation typology and distance. Error bars represent ± 1 SE.

Discussion

The primary aims of this study were to determine whether qualified drivers differ in their physiological responses and driving performance during a simulated driving task, based on their level of driving cue utilisation. Participants were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of driving. Using a 12-kilometre, simulated driving task, physiological responses were assessed through eye behaviour metrics (fixation rates, saccade amplitude, and fixation dispersion) and changes in cerebral oxygenation compared to baseline, while performance was assessed through the frequency of missed traffic signals and speed exceedances.

Higher cue utilisation is associated with the identification of more predictive features and greater efficiencies in information processing (Lansdale et al., 2010; Wiggins, 2015). Consequently, participants with higher cue utilisation should demonstrate more efficient search patterns and lower cognitive load during driving tasks, while maintaining a relatively high level of driving performance, compared to participants with lower cue utilisation. Therefore, it was hypothesised that, during a simulated driving task, participants with higher cue utilisation would record lower visual fixation rates, smaller mean saccade amplitudes, smaller mean fixation dispersions, smaller increases in cerebral oxygenation, fewer missed traffic signals, and fewer speed exceedances, compared to participants with lower cue utilisation.

As hypothesised, participants with higher cue utilisation demonstrated smaller mean saccade amplitudes, smaller mean fixation dispersions, and smaller changes in cerebral oxygenation in the prefrontal cortex, on average, compared to participants with lower cue utilisation. Participants with higher cue utilisation also recorded fewer missed traffic signals during the driving simulation task compared to participants with lower cue utilisation. There was no statistically significant difference in fixation rates nor speed exceedances based on levels of cue utilisation. In combination, these results indicate that higher cue utilisation amongst qualified operators is associated with the consumption of fewer resources, and the capacity to maintain relatively higher levels of performance during familiar operational tasks.

Theoretical and practical implications

Context-based assessments of cue utilisation, whereby cue utilisation and task performance are assessed in the same domain, and cross-task cue utilisation, whereby cue utilisation evaluated in one context is used to predict performance in another context, appear to differentiate operational performance (Brouwers et al., 2016, 2017; Loveday, Wiggins, & Searle, 2013; Small et al., 2014). Further, cue utilisation assessed in one domain appears to be associated with the consumption of cognitive resources during a novel task in a different domain (Brouwers et al., 2016; Sturman et al.., 2019). However, as the tasks in these experiments were novel, differences in resource consumption were likely due to differences in the rate of cue acquisition, rather than differences in the application of previously acquired cue-based associations. Consequently, it was unclear whether context-based cue utilisation would predict the consumption of cognitive resources amongst qualified operators.

The present study extends previous research, providing support for the assertion that, amongst qualified operators, a relatively greater capacity for cue utilisation is associated with the consumption of fewer cognitive resources necessary to maintain system operations. Further, differences in cerebral oxygenation were evident, controlling for participants' years of driving experience. This suggests that despite relatively equivalent opportunities for the acquisition of cue-based associations in the context of driving, higher cue utilisation was associated with a lower cognitive load during the simulated driving task.

The finding that higher cue utilisation is associated with smaller saccade amplitudes and smaller fixation dispersions extends previous research, providing evidence to suggest that there are differences in the patterns by which information is acquired. These results are consistent with the notion that individuals with higher cue utilisation target potential sources of threat, and consequently spend relatively more time than participants with lower cue utilisation attending to more specific areas of the environment and visually tracking features, rather than broadly scanning the visual scene. Indeed, participants with higher cue utilisation appear to adopt a strategy that allows them to examine co-located features within the context of the driving scenario, thereby reducing the frequency of large saccades between disparate features embodied within the scenario. This in turn suggests that higher cue utilisation may be associated with a more flexible representation of the environment, that can potentially be adapted as the scenario develops.

The present study also revealed that participants with higher cue utilisation recorded fewer missed traffic signals, as measured by failures to stop at red lights, stop signs and pedestrian crossings, compared to participants with lower cue utilisation. Further, there were no differences in speed exceedances based on the level of cue utilisation. These outcomes indicate that, despite consuming relatively fewer cognitive resources during a task, participants with higher cue utilisation demonstrate equivalent or greater performance compared to those participants with relatively lower cue utilisation. This is consistent with the notion that cue utilisation reduces cognitive load by creating efficiencies in information processing without negatively impacting performance (Brouwers et al., 2016; Lansdale et al., 2010; Sturman et al., 2019).

Attentional resource theory is based on the proposition that cognitive resources are drawn from a limited supply, and that the allocation of cognitive resources to a task depletes the residual attentional resources available (Kahneman, 1973; Warm, Parasuraman, & Matthews, 2008; Wickens, 1980). The outcomes of the present study indicate that drivers with lower cue utilisation are likely to consume cognitive resources at a greater rate during a task, and consequently should have fewer residual attentional resources when exposed to a task of similar duration. A reduction in the availability of attentional resources can reduce accuracy and increase operational errors (Reason, 1990; Wickens et al., 2015), and may account for the greater frequency of missed traffic signals recorded by participants with lower cue utilisation.

At an applied level, assessments of cue utilisation have implications for the management of performance amongst qualified drivers. Assessments of driving are typically conducted during the period when individuals are learning to drive. However, the outcomes of the present study demonstrate that drivers, with similar years of exposure, differ in their driving performance, based on their level of cue utilisation. Consequently, cue-based assessments of driving may be beneficial for predicting performance and assisting with the training of more experienced or older drivers. Alternatively, cue-based assessments of driving may assist the selection, training and supervision of professional drivers, such as taxi, bus or truck drivers.

Since cue utilisation was not a function of years of experience or hours of daily driving, it might be concluded that exposure alone does not predict cue utilisation. Consequently, regardless of operators' years of experience, proactive approaches, such as cue-based training, may be beneficial for increasing cue utilisation. For instance, cue discovery, whereby the ad-hoc acquisition of cues is enabled through trial and error with simulated operating systems

(Klayman, 1988), could be used to increase cue utilisation. Evidence supporting this approach was reported by Ivancic and Hesketh (2000), who demonstrated that allowing drivers to learn cue-based associations through trial-and-error during a simulated driving task resulted in significantly improved transfer of driving skills, compared to errorless learning. Similarly, information could be presented to drivers at various stages of their training to aid in the acquisition, interpretation, and integration of cues that are associated with their operational environment.

Limitations and future direction

Due to the cross-sectional design of the present study, the interpretation of the outcomes is limited by the extent to which causal relationships can be established. For instance, it is unclear whether a third variable, not controlled for in the present study, better explains the relationship between cue utilisation, resource consumption, and driving performance. Longitudinal studies, tracking cue utilisation and cognitive load with increased operator experience, may be beneficial in establishing the causal relationship between cue utilisation and cognitive resource consumption.

The generalisability of results may also be limited by the driving simulation, which was designed as a 20-minute, moderate demand task, to reflect the demands typically experienced by drivers in urban environments. However, the cognitive demands experienced by drivers may vary considerably with changes in factors such as traffic conditions and the length of drive. Therefore, the simulated drive may not accurately reflect the demands experienced by professional drivers who are more likely to drive for longer periods with varied traffic conditions. Further, cue utilisation amongst professional drivers may be influenced by factors including driver selection, training and performance management. Consequently, it remains unclear whether similar differences in resource consumption and performance based on cue utilisation would be observed amongst skilled professional drivers. Establishing whether cue utilisation predicts the rate at which cognitive resources are consumed amongst experienced professionals, is necessary before assessments of cue utilisation can be applied to the selection, training and management of professional operators.

Future research is also required to examine the relationship between cue utilisation and patterns of visual search amongst more experienced drivers. Despite the association between smaller fixation dispersions and lower cognitive load, there is also evidence to suggest that more experienced drivers typically demonstrate greater fixation dispersions in comparison to novice drivers (Crundall, Chapman, Phelps, & Underwood, 2003; Mourant & Rockwell, 1972; Underwood, Chapman, Bowden, & Crundall, 2002). The outcomes of the present study indicate that, amongst less experienced drivers of broadly equivalent exposure, higher cue utilisation is associated with smaller fixation dispersions. This suggests that there are individual differences in cue utilisation and visual search patterns and that greater driving experience may not necessarily be associated with higher cue utilisation. In fact, it may be the case that previous research has confounded driving experience with cue utilisation and there is now a need to examine differences in eye behaviour metrics based on driver experience, controlling for cue utilisation.

Conclusion

The current study was designed to determine whether qualified drivers differ in rate at which cognitive resources are consumed and driving performance during a simulated driving task, based on their level of cue utilisation. The results indicated that participants with higher cue utilisation demonstrated smaller mean visual saccade amplitudes, smaller mean fixation dispersions, smaller increases in cerebral oxygenation, and recorded fewer missed traffic signals, compared to participants with lower cue utilisation. These results suggest that drivers with higher cue utilisation experienced lower levels of cognitive load during the simulated driving task, while maintaining a higher level of performance, compared to drivers with lower cue utilisation. From an applied perspective, this suggests that domain-specific assessments of cue utilisation may predict the consumption of cognitive resources and performance amongst qualified drivers.

References (Study 5)

Berg, D. J., Boehnke, S. E., Marino, R. A., Munoz, D. P., & Itti, L. (2009). Free viewing of dynamic stimuli by humans and monkeys. *Journal of Vision*, *9*, 19–19. doi:10.1167/9.5.19

Bertram, R., Helle, L., Kaakinen, J. K., & Svedström, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PloS One*, *8*, e66169. doi:10.1371/journal.pone.0066169

Brouwers, S., Wiggins, M. W., Griffin, B., Helton, W. S., & O'Hare, D. (2017). The role of cue utilisation in reducing the workload in a train control task. *Ergonomics*, *60*, 1–16. doi:10.1080/00140139.2017.1330494

Brouwers, S., Wiggins, M. W., Helton, W., O'Hare, D., & Griffin, B. (2016). Cue utilization and cognitive load in novel task performance. *Frontiers in Psychology*, *7*, 1–12. doi:10.3389/fpsyg.2016.00435

Brunswik, E. (1955). Representative Design and Probabilistic Theory in a Functional Psychology. *Psychological Review*, *62*, 193–217. doi:10.1037/h0047470

Camilli, M., Terenzi, M., & Di Nocera, F. (2007). Concurrent validity of an ocular measure of mental workload. In *Human factors issues in complex system performance. In D. de Waard, G.R.J. Hockey, P. Nickel, and K.A. Brookhuis (Eds.)* (pp. 117–129). Shaker Publishing Maastricht, the Netherlands.

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, *27*, 951–964. doi:10.1068/p270951

Chung, P. H., & Byrne, M. D. (2008). Cue effectiveness in mitigating postcompletion errors in a routine procedural task. *International Journal of Human-Computer Studies*, *66*, 217–232. doi:10.1016/j.ijhcs.2007.09.001

Crundall, D., Chapman, P., Phelps, N., & Underwood, G. (2003). Eye Movements and Hazard Perception in Police Pursuit and Emergency Eye Movements and Hazard Perception in Police Pursuit and Emergency Response Driving. doi:10.1037/1076-898X.9.3.163

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A Random Glance at the Flight Deck: Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload. *Journal of Cognitive Engineering and Decision Making*, *1*, 271–285. doi:10.1518/155534307X255627.

Ekkekakis, P. (2009). Illuminating the Black Box: Investigating Prefrontal Cortical Hemodynamics during Exercise with Near-Infrared Spectroscopy. *Journal of Sport and Exercise Psychology*, *31*, 505–553. doi:10.1123/jsep.31.4.505

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305. doi:10.1146/annurev.psych.47.1.273

Ericsson, K. A., & Towne, T. J. (2010). Expertise. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 404–416. doi:10.1002/wcs.47

Evans, D. C., & Fendley, M. (2017). A multi-measure approach for connecting cognitive workload and automation. *Journal of Human Computer Studies*, *97*, 182–189. doi:10.1016/j.ijhcs.2016.05.008

Farrington-Darby, T., & Wilson, J. R. (2006). The nature of expertise: A review. *Applied Ergonomics*, *37*, 17–32. doi:10.1016/j.apergo.2005.09.001

Fu, R., Guo, Y., Chen, Y., Yuan, W., Ma, Y., Peng, J., … Wang, C. (2011). Research on heart rate and eye movement as indicators of drivers' mental workload. In *Road Safety and Simulation* (pp. 1–20).

Gratton, G., & Fabiani, M. (2006). Optical imaging of brain function in R. Parasuraman, M. Rizzo (Eds.), Neuroergonomics: The Brain At Work, Oxford University. In *Oxford Series in Human-Technology Interactions* (pp. 65–81). Cambridge, MA.

Grüner, M., & Ansorge, U. (2017). Mobile Eye Tracking During Real-World Night Driving: A Selective Review of Findings and Recommendations for Future Research. *Journal of Eye Movement Research*, *1011*, 1–18. doi:10.16910/jemr.10.2.1

Helton, W. S., Warm, J. S., Tripp, L. D., Matthews, G., Parasuraman, R., & Hancock, P. A. (2010). Cerebral lateralization of vigilance: A function of task difficulty. *Neuropsychologia*, *48*, 1683–1688. doi:10.1016/j.neuropsychologia.2010.02.014

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504. doi:10.1016/j.tics.2003.09.006

Ikehara, C. S., & Crosby, M. E. (2005). Assessing Cognitive Load with Physiological Sensors. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, *00*, 295a–295a. doi:10.1109/HICSS.2005.103

Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence. *Ergonomics*, *43*, 1966–1984. doi:10.1080/00140130050201427

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ.

Klayman, J. (1988). Cue discovery in probabilistic environments: Uncertainty and experimentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 317. doi:10.1037//0278-7393.14.2.317

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. In *Journal of Cognitive Engineering and Decision Making* (Vol. 4, pp. 186–209). SAGE Publications Sage CA: Los Angeles, CA. doi:10.1518/155534310x12844000801203

Lansdale, M., Underwood, G., & Davies, C. (2010). Something Overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, *24*, 213–225. doi:10.1002/acp.1552

Loveday, T., Wiggins, M. W., Festa, M., Schell, D., & Twigg, D. (2013). Pattern recognition as an indicator of diagnostic expertise. In P. L. Carmona, J. S. Sanchez, & A. L. N. Fred (Eds.), *Pattern Recognition - Application and Methods* (Vol. 204, pp. 1–11). Berlin: Springer. doi:10.1007/978-3-642-36530-0

Loveday, T., Wiggins, M. W., Harris, J. M., O'Hare, D., & Smith, N. (2013). An objective approach to identifying diagnostic expertise among power system controllers. *Human Factors*, *55*, 90–107. doi:10.1177/0018720812450911

Loveday, T., Wiggins, M. W., & Searle, B. J. (2013). Cue utilization and broad indicators of workplace expertise. *Journal of Cognitive Engineering and Decision Making*, *8*, 98–113. doi:10.1177/1555343413497019

Loveday, T., Wiggins, M. W., Searle, B. J., Festa, M., & Schell, D. (2013). The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*, 125–137. doi:10.1177/0018720812448475

Luck, S. J., Hillyard, S. a, Mouloua, M., & Hawkins, H. L. (1996). Mechanisms of visual-spatial attention: resource allocation or uncertainty reduction? *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 725–737. doi:10.1037/0096-

1523.22.3.725

McCormack, C., Wiggins, M. W., Loveday, T., & Festa, M. (2014). Expert and competent non-expert visual cues during simulated diagnosis in intensive care. *Frontiers in Psychology*, *5*, 1–7. doi:10.3389/fpsyg.2014.00949

Morrison, B. W., Wiggins, M. W., Bond, N. W., & Tyler, M. D. (2013). Measuring relative cue strength as a means of validating an inventory of expert offender profiling cues. *Journal of Cognitive Engineering and Decision Making*, *7*, 211–226. doi:10.1177/1555343412459192

Mourant, R. R., & Rockwell, T. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors*, *14*, 325–335. doi:10.1177/001872087201400405

Norman, D. A. (2015). The human side of automation. In *Road Vehicle Automation 2* (pp. 73–79). Springer.

Ong, M., Russell, P. N., & Helton, W. S. (2013). Frontal cerebral oxygen response as an indicator of initial attention effort during perceptual learning. *Experimental Brain Research*, *229*, 571–578. doi:10.1007/s00221-013-3619-9

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*, 924–927. doi:10.1126/science.472714

Pauley, K., O'Hare, D., & Wiggins, M. W. (2009). Measuring expertise in weather-related aeronautical risk perception: The validity of the Cochran–Weiss–Shanteau (CWS) Index. *The International Journal of Aviation Psychology*, *19*, 201–216. doi:10.1080/10508410902979993

Perry, N. C., Wiggins, M. W., Childs, M., & Fogarty, G. (2012). Can reduced processing decision support interfaces improve the decision-making of less-experienced incident commanders? *Decision Support Systems*, *52*, 497–504. doi:10.1016/j.dss.2011.10.010

Perry, N. C., Wiggins, M. W., Childs, M., & Fogarty, G. (2013). The Application of Reduced-Processing Decision Support Systems to Facilitate the Acquisition of Decision-Making Skills. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*, 535–544. doi:10.1177/0018720812467367

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211–219). Hershey, PA: IGI Publishing. doi:10.4018/978-1-59140-562-7.ch034

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge university press.

Robinson, D. A. (1965). The mechanics of human smooth pursuit eye movement. *The Journal of Physiology*, *180*, 569–591. doi:10.1113/jphysiol.1965.sp007718

Salthouse, T. A. (1991). Expertise as the circumvention of human processing limitations. In *Toward a general theory of expertise: Prospects and limits* (pp. 286–300). Cambridge, NY: Cambridge University Press.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Eye Tracking Research & Applications (ETRA) Symposium* (pp. 71–78). Palm Beach Gardens, FL: ACM.

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50*, 864–878. doi:10.1037/e578082012-006

Schvaneveldt, R. W., Beringer, D. B., & Lamonica, J. A. (2001). Priority and organization of information accessed by pilots in various phases of flight. *The International Journal of Aviation Psychology*, *11*, 253–280. doi:10.1037/e406222004-001

Shaw, T. H., Warm, J. S., Finomore, V., Tripp, L., Matthews, G., Weiler, E., & Parasuraman, R. (2009). Effects of sensory modality on cerebral blood flow velocity during vigilance. *Neuroscience Letters*, *461*, 207–211. doi:10.1016/j.neulet.2009.06.008

Sherbino, J., Dore, K. L., Wood, T. J., Young, M. E., Gaissmaier, W., Kreuger, S., & Norman, G. R. (2012). The relationship between response time and diagnostic accuracy. *Academic Medicine*, *87*, 785–791. doi:10.1097/ACM.0b013e318253acbd

Small, A. J., Wiggins, M. W., & Loveday, T. (2014). Cue-based processing capacity, cognitive load and the completion of simulated short-duration vigilance tasks in power transmission control. *Applied Cognitive Psychology*, *28*, 481–487. doi:10.1002/acp.3016

Sturman, D., Wiggins, M. W., Auton, J. C., & Loft, S. (2019). Cue utilization differentiates resource allocation during sustained attention simulated rail control tasks. *Journal of Experimental Psychology: Applied*, Advance online publication. doi:10.1037/xap0000204

Underwood, G., Chapman, P., Bowden, K., & Crundall, D. (2002). Visual search while driving : skill and awareness during inspection of the scene, *5*, 87–97.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, *46*, 629–646. doi:10.1080/0014013031000090116

Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*, 435–446. doi:10.1016/j.trf.2011.04.008

Waller, M. J., Gupta, N., & Giambatista, R. C. (2004). Effects of Adaptive Behaviors and Shared Mental Models on Control Crew Performance. *Management Science*, *50*, 1534–1544. doi:10.1287/mnsc.1040.0210

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

Watson, M. R., Brennan, A. A., Kingstone, A., & Enns, J. T. (2010). Looking versus seeing: Strategies alter eye movements during visual search. *Psychonomic Bulletin and Review*, *17*, 543–549. doi:10.3758/PBR.17.4.543

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickers (Ed.), *Attention and Performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159–177. doi:10.1080/14639220210123806

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. Psychology Press.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology*, *25*, 229–235. doi:10.1002/acp.1668

Wiggins, M. W. (2012). The role of cue utilisation and adaptive interface design in the management of skilled performance in operations control. *Theoretical Issues in Ergonomics Science*, *15*, 283–292. doi:10.1080/1463922X.2012.724725

Wiggins, M. W. (2015). Diagnosis and Instructional Systems Design. In M. W. Wiggins & T. Loveday (Eds.), *Diagnostic Expertise in Organizational Environments* (pp. 69–80). Surrey, England: Ashgate Publishing Limited, Surrey, England.

Wiggins, M. W., Azar, D., Hawken, J., Loveday, T., & Newman, D. (2014). Cue-utilisation typologies and pilots' pre-flight and in-flight weather decision-making. *Safety Science*, *65*, 118–124. doi:10.1016/j.ssci.2014.01.006

Wiggins, M. W., Brouwers, S., Davies, J., & Loveday, T. (2014). Trait-based cue utilization and initial skill acquisition: Implications for models of the progression to expertise. *Frontiers in Psychology*, *5*, 1–8. doi:10.3389/fpsyg.2014.00541

Wiggins, M. W., Loveday, T., & Auton, J. C. (2015). *EXPERT Intensive Skills Evaluation (EXPERTise) Test*. Sydney: Macquarie University.

Wiggins, M. W., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, *1*, 305–320. doi:10.1037/1076-898X.1.4.305

Wiggins, M. W., & O'Hare, D. (2003). Expert and novice pilot perceptions of static in-flight images of weather. *International Journal of Aviation Psychology*, *13*, 173–187. doi:10.1207/S15327108IJAP1302_05

Wiggins, M. W., Stevens, C., Howard, A., Henley, I., & O'Hare, D. (2002). Expert, intermediate and novice performance during simulated pre-flight decision-making. *Australian Journal of Psychology*, *54*, 162–167. doi:10.1080/00049530412331312744

Wiggins, M. W., Whincup, E., & Auton, J. C. (2018). Cue utilisation reduces effort but increases arousal during a process control task. *Applied Ergonomics*, *69*, 120–127. doi:10.1016/j.apergo.2018.01.012

Wigton, R. S. (1996). Social Judgement Theory and Medical Judgement. *Thinking & Reasoning*, *2*, 175–190. doi:10.1080/135467896394492

Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, *43*, 303–311. doi:10.1016/j.jesp.2006.02.004

Bridging Section 4

Study 5 was designed to examine whether an assessment of cue utilisation within the domain of driving predicts differences in performance and cognitive resource consumption during a simulated driving task. A primary difference from Studies 1 to 4 was the use of a dynamic, sustained attention task, which contained features that were less predictable to those used in earlier studies. It was anticipated that participants with higher cue utilisation would demonstrate more efficient visual search patterns by fixating on key features within close proximity, thereby reflecting behaviour that reduced cognitive load.

The results of Study 5 indicated that participants with higher cue utilisation recorded smaller mean visual saccade amplitudes, smaller mean fixation dispersions, and smaller increases in cerebral oxygenation, while recording fewer missed traffic signals, compared to participants with lower cue utilisation. Overall, these outcomes suggest that cue utilisation is associated with more efficient search patterns and lower levels of cognitive load during dynamic, sustained attention tasks. Further, the outcomes indicate that, despite consuming fewer cognitive resources, participants with higher cue utilisation were able to maintain a greater level of performance, compared to those participants with lower cue utilisation. As in Study 2, this outcome supports the proposition that differences in cognitive resource consumption based on cue utilisation are not due to reductions in effort at the cost of performance. Rather, the results are consistent with models of cue utilisation, which posit that the activation of patterns of cues is a non-conscious process that enables situations to be recognised as familiar, thereby facilitating accurate and rapid responses (Klein, 1993; Wiggins, 2015a).

The experimental paradigms used in Studies 1 to 5 have provided a degree of experimental control, enabling differences in the consumption of cognitive resources to be identified. However, controlled experimental studies necessitate the use of simplified simulated operational environments, which differ from typical operational environments in terms of the complexity of feature-event relationships. Therefore, Study 6 was designed to evaluate the relationship between cue utilisation and cognitive load during qualified operators' regular operational tasks. During periods of DNSP control room operators' regular workdays, physiological measures of workload were assessed through changes in cerebral oxygenation in the prefrontal cortex compared to baseline, and through eye behaviour metrics (fixation rates, saccade amplitude, and fixation dispersion). Based on the proposition that the activation and retrieval of cues from long-term memory has the advantage of imposing relatively fewer demands on working memory resources, it was anticipated that higher cue utilisation would be associated with smaller increases in cerebral oxygenation during regular operational tasks, reflecting lower levels of cognitive load. Further, based on the assumption that cue utilisation enables operators to draw on patterns in memory to anticipate events, it was anticipated that operators with higher cue utilisation would demonstrate more efficient search patterns, compared to those operators with lower cue utilisation.

STUDY 6

Publication History

The paper arising from Study 6 was entitled, "Control Room Operators' Cue Utilisation Predicts Cognitive Resource Consumption during Regular Operational Tasks". A version of this paper has been submitted to the *Human Factors* Journal, and is currently under review. The author of the present dissertation wrote approximately 80% of this paper.

Control Room Operators' Cue Utilisation Predicts Cognitive Resource Consumption during Regular Operational Tasks

Daniel Sturman[1], Mark W. Wiggins[1], Jaime C. Auton[2], Shayne Loft[3], William S. Helton[3], Johanna I. Westbrook[1], and Jeffery Braithwaite[1]

[1] *Macquarie University, NSW, Australia*

[2] *The University of Adelaide, SA, Australia*

[3] *The University of Western Australia, WA, Australia*

[4] *George Mason University, VA, United States*

Acknowledgements

Support for this research was provided by the Australian Research Council [ARC LP160100211].

Abstract

Simulated laboratory studies have demonstrated that cue utilisation differentiates cognitive load during process control tasks. This study was designed to examine whether qualified practitioners' cue utilisation is predictive of their sustained attention during regular operational tasks. Australian Distribution Network Service Provider (DNSP) operators were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of power distribution. During two, 20-minute periods of operators' regular workdays, physiological measures of workload were assessed through changes in cerebral oxygenation in the prefrontal cortex compared to baseline, and through eye behaviour metrics (fixation rates, saccade amplitude, and fixation dispersion). The results indicated that there were no statistically significant differences in eye behaviour metrics, based on levels of cue utilisation. However, as hypothesised, during both sessions, operators with higher cue utilisation demonstrated smaller increases from baseline in cerebral oxygenation in the prefrontal cortex, compared to operators with lower cue utilisation. This outcome is consistent with the proposition that operators with higher cue utilisation experience lower cognitive load during periods of their regular workday, compared to operators with lower cue utilisation.

Control Room Operators' Cue Utilisation Predicts Cognitive Resource Consumption during Regular Operational Tasks

Control room operators in high risk industrial environments, including power control and rail control, must respond rapidly and accurately to deviations in the system state (O'Hara & Hall, 1992; Stanton, Salmon, Walker, & Jenkins, 2009). This requires that operators sustain a visual search to monitor and control various operating systems, often for extended periods (Mumaw, Roth, Vicente, & Burns, 2000; Vicente, Roth, & Mumaw, 2001). The failure to sustain attention can result in inaccurate or delayed responses, increasing the likelihood of potentially catastrophic errors (Helton & Warm, 2008; Reason, 2000; Warm, 1984).

Attentional Resource Theory posits that sustaining attention over extended periods results in the consumption of attentional resources (Grier et al., 2003; Helton & Russell, 2012; Kahneman, 1973). As attentional resources are drawn from a limited pool of cognitive resources, this can eventually result in fewer residual resources than are necessary to manage changes in the system state (Kahneman, 1973; Parasuraman, 1979; Wickens, 2002). Consequently, performance efficiency declines throughout sustained attention tasks, as evident in an increased response latency or decreased accuracy when responding to critical signals (Parasuraman, 1979; Shaw et al., 2009). This decline in performance is referred to as the vigilance decrement (Parasuraman, 1979).

A resource depletion account of the vigilance decrement is supported by evidence to the effect that more cognitively demanding tasks result in a more rapid decline in performance (Helton & Russell, 2013; Shaw et al., 2013). For example, increasing the demands of a sustained attention task, by decreasing signal salience or through the addition of a secondary task, results in a more rapid increase in response latency and the frequency of errors when responding to critical signals (Helton & Russell, 2011; Helton & Warm, 2008; Smit, Eling, & Coenen, 2004). Conversely, reducing task demands, by increasing signal salience or by inserting warning signals, results in a greater number of correctly detected critical signals (Helton & Warm, 2008; MacLean et al., 2009).

Based on the proposition that more demanding tasks result in a more rapid consumption of cognitive resources, it follows that operators who experience lesser cognitive load during operational tasks will consume fewer resources over a given period. Consequently, they should retain greater residual cognitive resources, thereby enabling them to sustain their performance for longer periods (Matthews, Warm, Shaw, & Finomore, 2010). Cue utilisation is one strategy that operators appear to engage to reduce the rate at which cognitive resources are consumed (Brouwers, Wiggins, Helton, O'Hare, & Griffin, 2016; Perry, Wiggins, Childs, & Fogarty, 2012; Wiggins, 2011).

Cue utilisation

Cues are associations between situation-specific environmental features and task-related objects or events (Brunswik, 1955; Klein, Calderwood, & Clinton-Cirocco, 2010). Cue utilisation is the application of cue-based processing, which is dependent upon individuals' capacity to develop and recognise cues (Lansdale, Underwood, & Davies, 2010; Wiggins, 2012). It requires the identification of predictive features in the operational environment, the association between these features and events in memory, the retention of these cue-based associations, and the application of cues in response to environmental features (Wiggins, 2012, 2015)

The activation and retrieval of cues from long-term memory has the advantage of enabling performance while imposing relatively fewer demands on working memory resources (Chung & Byrne, 2008; Evans & Fendley, 2017; Lansdale et al., 2010). Further, effective cue utilisation should enable operators to attend to features of greater relevance, reducing the overall number of features to which they attend, and thereby reducing the rate at which cognitive resources are consumed (Sturman et al., 2019; Weiss & Shanteau, 2003; Williams, Ward, Knowles, & Smeeton, 2002). Evidence to support the assertion that higher cue utilisation is associated with the consumption of fewer cognitive resources can be drawn from research demonstrating that, during sustained attention tasks, participants with higher cue utilisation report relatively lower perceived cognitive load, and record relatively smaller increases in response latency, compared to participants with lower cue utilisation (Brouwers, Wiggins, Griffin, Helton, & O'Hare, 2017; Brouwers et al., 2016).

Using a simulated rail control task that incorporated an implicit pattern of train movements, Brouwers et al. (2016) noted that the addition of a concurrent, secondary task failed to impact the performance of participants with higher cue utilisation. However, it resulted in a significant reduction in performance for participants with lower cue utilisation. It was posited that participants with higher cue utilisation had recognised the implicit pattern, which enabled the adoption of a cue-based strategy that reduced the rate at which cognitive resources were consumed, and thereby provided additional residual resources that minimised the impact of the secondary task (Brouwers et al., 2016). Sturman et al. (in submission) observed similar results, demonstrating that relatively higher cue utilisation amongst qualified power distribution control room operators was associated with higher performance during a simulated power control sustained visual search task.

Although the effects appear relatively consistent, evidence to support the association between cue utilisation and the rate at which cognitive resources are consumed relies primarily on inferences derived from mean response latencies (Brouwers et al., 2016, 2017; Small, Wiggins, & Loveday, 2014; Sturman et al., in submission). However, this association could potentially be explained by alternative factors, such as participants' level of motivation or engagement. Further, assessments of cue utilisation and sustained attention both typically rely on measures of response latency and accuracy, and data for each participant is typically collected in a single session (Brouwers et al., 2016, 2017; Small et al., 2014). Consequently, the relationship between cue utilisation and cognitive resource consumption may be partially attributable to common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). To overcome these potential methodological issues, complementary evidence using alternative measures of cognitive load is required.

Sturman et al. (2019) examined the relationship between cue utilisation and the consumption of cognitive resources amongst novice operators using eye tracking data to assess fixation rates (number of fixations per minute), and Functional Near Infrared Spectroscopy (fNIRS) to measure cerebral oxygenation in the prefrontal cortex. Lower fixation rates and smaller increases in cerebral oxygenation in the prefrontal cortex are indicative of more efficient search patterns and lower consumption of cognitive resources during sustained attention tasks (Goldberg & Kotval, 1999; Helton et al., 2010; Ong, Russell, & Helton, 2013; Poole & Ball, 2006).

During novel rail control simulations containing repetitious patterns of train movement, participants with higher cue utilisation assessed in the domain of driving demonstrated greater decreases in fixation rates and smaller increases in cerebral oxygenation in the prefrontal cortex, while maintaining a higher level of performance, compared to participants with lower cue utilisation (Sturman et al., 2019). This evidence provides additional support for the proposition that higher cue utilisation is associated with the consumption of fewer cognitive resources during novel tasks containing predictable features (Brouwers et al., 2016; Perry et al., 2012; Wiggins, 2011). However, Sturman et al. (2019) relied on cross-task cue utilisation, whereby cue utilisation evaluated in one context (driving) was used to predict cognitive load in another novel context (rail control). As these novice operators had no prior opportunities to acquire relevant cues, differences in cognitive load based on cue utilisation likely reflect differences in the rate of cue acquisition during the novel tasks. Consequently, it remains unclear whether differences in cognitive load based on cue utilisation are also evident amongst qualified personnel in familiar operating environments.

In complex operating environments, the emergence of critical features is likely to be less predictable compared to experimental tasks containing repetitious patterns. While fixation rates provide an indication of search pattern efficiencies in predictable environments, additional eye behaviour metrics can provide a more detailed measure of search patterns in less predictable environments. For example, researchers often analyse a range of eye behaviour metrics, including saccade amplitude and fixation dispersion, to examine hazard detection during dynamic driving tasks (Chapman & Underwood, 1998; Grüner & Ansorge, 2017; Underwood, Crundall, & Chapman, 2011).

Saccade amplitude and fixation dispersion can be used to assess workload and the allocation of cognitive resources during sustained attention tasks (Camilli, Terenzi, & Di Nocera, 2007; Di Nocera, Camilli, & Terenzi, 2007; Fu et al., 2011). Saccade amplitude refers to the change in the degrees of visual angle from the pre-saccade fixation to the post-saccade fixation. Greater performance in the detection of critical targets has been associated with search strategies involving saccades of smaller amplitude (Bertram, Helle, Kaakinen, & Svedström, 2013). Fixation dispersion is the extent to which fixations are distributed while completing a task. Smaller fixation dispersions have been associated with lower subjective ratings of workload during simulated flight (Di Nocera et al., 2007) and during visuomotor tasks (Camilli et al., 2007).

To examine whether physiological measures of cognitive resource consumption differ based on cue utilisation with qualified operators in dynamic environments, Sturman and Wiggins (in submission) examined indicators of qualified drivers' consumption of cognitive resources during a simulated driving task. Qualified drivers with higher cue utilisation demonstrated smaller mean visual saccade amplitudes, smaller mean fixation dispersions, smaller increases in cerebral oxygenation, and recorded fewer missed traffic signals during the simulated driving task, compared to participants with lower cue utilisation. These results provide support for the assertion that, amongst qualified operators, a relatively greater capacity for cue utilisation is associated with the consumption of fewer cognitive resources during simulated, sustained attention tasks.

The aim of the present study was to examine whether qualified operators' cue utilisation is associated with the allocation of cognitive resources to regular operational tasks. To assess cue utilisation and cognitive resource consumption in an operational context, operators from four Australian Distribution Network Service Provider (DNSP) control rooms were recruited to participate in the study. Cue utilisation, assessed in the context of power control, was used to predict physiological measures of cognitive load during periods of operators' regular workdays. As the activation and retrieval of cues from long-term memory has the advantage of imposing relatively fewer demands on working memory resources, operators with higher cue utilisation should allocate fewer cognitive resources to regular operational tasks. Further, based on the proposition that operators with higher cue utilisation will draw on patterns in memory to anticipate events and enable more efficient search patterns, these operators should spend more time attending to specific areas associated with their operational tasks, rather than broadly scanning their environment. Specifically, it was hypothesised that, during periods of their regular workday, control room operators with higher cue utilisation would record lower increases in cerebral oxygenation from baseline, lower visual fixation rates, smaller mean fixation dispersions, and smaller mean saccade amplitudes, compared to operators with lower cue utilisation.

Method

Design

Testing sessions were conducted during two, 20-minute periods of each participant's regular workday. The two sessions were not necessarily comparable, as the duration between the first and second sessions varied. Consequently, the two sessions were analysed separately. Further, considering the two sessions separately also allowed for a more robust evaluation of any effects that were evident in the first session. Each session comprised a 2 x 4 mixed factorial design incorporating two levels of cue utilisation (higher and lower) as a between-subjects factor, and four, 5-minute time periods (Period 1, Period 2, Period 3, and Period 4) as a within-subjects factor. Participants were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of power distribution. Time constituted the four quartiles of the 20-minute testing sessions.

Participants

Participants comprised 38 male power distribution network controllers, recruited from 4 Australian DNSP control rooms. Participants ranged in age from 27 to 60 years (M = 42.2, SD = 7.6), had acquired a mean 8.8 years (SD = 4.8) of experience as network controllers, and had acquired a mean 19.9 years (SD = 9.7) working in power distribution.

EXPERTise 2.0

EXPERT Intensive Skills Evaluation (EXPERTise 2.0) is an on-line assessment tool designed to assess behaviour consistent with the utilisation of cues within a specific context (Wiggins, Loveday, & Auton, 2015). For the current study, EXPERTise 2.0 was tailored to the domain of power distribution. Tasks in the EXPERTise battery include a Feature Identification Task (FIT), a Feature Recognition Task (FRT), a Feature Association Task (FAT), a Feature Discrimination Task (FDT), and a Feature Prioritisation Task (FPT).

The FIT consists of six scenarios that network controllers might typically experience, including a transformer failure, a voltage under or overload, an erroneous indication, or a normal condition. For each scenario, participants are presented with a line diagram consisting of electrical symbology with which they are familiar. Participants are instructed to review each diagram, and to identify the area of concern as quickly as possible by selecting the specific feature of interest, or by selecting an icon that indicates the system to be operating normally. Response latency is measured as the time in milliseconds from the initial presentation of the diagram to the selection of an area of concern. Higher cue utilisation is associated with shorter mean response latencies (Loveday, Wiggins, & Searle, 2013; Schriver, Morrow, Wickens, & Talleur, 2008; Wiggins, 2014).

During the FRT, participants are presented with line diagrams representing electricity distribution systems. The FRT consists of 10 trials, during which a line diagram is presented for a period of exposure varying between 20 seconds and 60 seconds. After each line diagram has been displayed, the diagram is removed and participants are asked to select, from one of five options, the condition represented in the preceding display (e.g., "The substation has suffered a loss of all indications", "The probability is that the flow on Tx2 is correct; the 11kV West Bus is still in service and no load has been lost", etc.). The FRT assesses operators' ability to rapidly extract information from the scenarios, and use this information to form accurate judgements. Consequently, a greater number of correct judgements is indicative of higher cue utilisation (Wiggins & O'Hare, 2003a).

During the FAT, participants are presented with two phrases related to power distribution (e.g., 'Overhead Lines', 'Low Voltage'). Participants are presented with a total of 13 trials. Each pair of phrases is presented for a period of two seconds, after which time the phrases are removed and participants are asked to rate the extent to which they believe that the two phrases are related on a six-point Likert scale ranging from 1 (*Extremely unrelated*) to 6 (*Extremely related*). As cue utilisation requires operators to rapidly differentiate predictive and non-predictive feature-event relationships, higher cue utilisation is associated with a greater ratio of variance in ratings to mean response latency (Schvaneveldt, Beringer, & Lamonica, 2001).

The FDT consists of two detailed power distribution scenarios. For each scenario, participants are asked to formulate a decision. Following each decision, participants rate a list of features based on their perceived relevance to the decision using a 10-point Likert scale ranging from 1 (*Not important at all*) to 10 (*Extremely important*). As effective cue utilisation requires features to be identified as more or less relevant, higher cue utilisation is associated with a greater variance in ratings of perceived relevance (Brouwers et al., 2017; Pauley, O'Hare, & Wiggins, 2009; Weiss & Shanteau, 2003).

In the FPT, participants are presented with two problem scenarios (e.g., received notification from the call centre that a member of the public has lodged an emergency call), which are accompanied by drop-down lists of key features relating to the scenario. The drop-down menus are feature-labelled (e.g., 'Location of the Report', 'Content of the Customer Report'), and upon selection, provide information which is relevant to that scenario. Access to the information is limited to a 90-second period. The FPT is designed to assess an individual's ability to prioritise feature cue acquisition (Wiggins & O'Hare, 1995; Wiggins, Stevens, Howard, Henley, & O'Hare, 2002). Lower cue utilisation is associated with the sequential selection of drop-down lists (e.g., from top to bottom of the screen, in the order in which they are presented), while higher cue utilisation is associated with a less sequential selection of information (Wiggins et al., 2002).

Eye tracking

Prior to each testing session, participants were fitted with SMI Eye Tracking Glasses (version 2) using the system's standard operating procedures. The SMI glasses have a $60° \times 46°$ recording visual angle, and a sampling frequency of 60 Hz. Eye tracking data were recorded for the duration of each 20-minute testing session, and later analysed using BeGaze software. Fixation rates (mean number of fixations made per minute), fixation dispersion (mean dispersion of fixations in pixels), and saccade amplitude (mean change in degrees of the visual angle per saccade) were calculated for each of the four time periods throughout each testing session.

Near Infrared Spectroscopy (NIRS)

Prior to each testing session, participants were fitted with a Portalite Near Infrared Spectroscopy (NIRS) sensor. NIRS utilises light in the near-infrared spectrum to assess haemoglobin levels in targeted brain regions. Cerebral oxygenation (rSO2) is calculated as the ratio of oxyhaemoglobin to total haemoglobin, and represents a measure of cerebral activation (Ekkekakis, 2009; Gratton & Fabiani, 2006).

The NIRS sensor was positioned approximately one centimetre above the participants' right eyebrow. Due to time constraints associated with testing during periods of participants' regular workdays, the baseline period was restricted to two minutes, during which time participants were asked to sit quietly, minimise body movements, and to remain as relaxed as possible. rSO2 during the second minute of the baseline period was used as a baseline index.

Subjective workload

Subjective workload was measured with the NASA Task Load Index (NASA-TLX: Hart & Staveland, 1988). The NASA-TLX is a widely used and validated rating procedure, which involves rating workload along six dimensions including physical demands, mental demands, temporal demands, effort, frustration, and performance (Hart & Staveland, 1988). Every 10 minutes throughout each testing session, participants were asked to rate their perception of workload during the preceding 10 minutes, using a seven-point Likert scale for each dimension of workload.

Procedure

Testing sessions were conducted following approval from the University Human Research Ethics Committee. Testing consisted of an online component and an in situ component. For the online component, power controllers from the 4 DNSP control rooms were provided general information regarding the research, and were given the URL to the EXPERTise 2.0 website. After logging in, participants were asked to read an information sheet and give consent to their participation in the online component of the research. Having given their consent, participants answered a series of questions that were incorporated to generate a unique participant code. Participants were then asked a series of demographic questions, including their age, sex, the number of years that they had been employed as a network controller, and the number of years they had worked in power distribution. On completion of the demographic questions, the participants completed the EXPERTise tasks. The identity of participants who elected to participate in the online component of the study remained anonymous.

During the in situ testing sessions, participants from the four DNSP control rooms were invited to wear eye-tracking glasses and a near-infrared spectroscope during two, 20-minute periods of their regular workday, once near the beginning of their shift, and once towards the end of their shift. Participants were asked to read an information sheet and give consent to their participation in the in situ component of the research. Having given their consent, participants answered the same series of questions that were incorporated to generate their unique participant code in the online component, which later allowed their data from the two components of the study to be matched anonymously. Participants were then fitted with the eye-tracking glasses and near-infrared spectroscope, and subsequently instructed to continue with their current work tasks as they would during a typical workday. Participants were asked, if time permitted, to complete the NASA-TLX every 10 minutes during each testing session.

Results

Physiological measures

NIRS data were used to calculate relative measures of Regional Oxygen Saturation (rSO2) for each period of each testing session. rSO2 scores were calculated by comparing mean rSO2 during each period to the baseline measure taken prior to each testing session. Scores represent the percentage change in rSO2 from baseline, with positive scores representing an increase in rSO2 compared to baseline, and negative scores represented a decrease in rSO2 compared to baseline. rSO2 scores were approximately normally distributed for each period and cue utilisation typology.

Fixation rates were calculated as the mean number of eye fixations recorded per minute. Saccade amplitude was calculated as the mean change in degrees of the visual angle per saccade, while fixation dispersion was calculated as the mean dispersion of fixations in pixels. Due to calibration difficulties, eye behaviour metrics were unable to be accurately calculated for two participants, one in the high cue utilisation typology and one in the low cue utilisation typology. Consequently, data for these two participants were excluded from analyses involving fixation rates, saccade amplitude and fixation dispersion. Eye behaviour metrics were approximately normally distributed for each period and cue utilisation typology.

Cue utilisation typologies

EXPERTise data were used to identify cue utilisation typologies that corresponded to higher or lower levels of cue utilisation (Brouwers et al., 2016; Small et al., 2014; Sturman et al., 2019). Consistent with a standard approach for classifying participants into cue utilisation typologies (Loveday, Wiggins, Searle, Festa, & Schell, 2013; Wiggins, Brouwers, Davies, & Loveday, 2014), scores for each task were converted to *z* scores, and a cluster analysis was used to identify two typologies. The first cluster, labelled the higher cue utilisation typology, consisted of participants, the centroids for whom reflected a shorter response latency on the

FIT, greater accuracy on the FRT, a higher mean ratio of variance to reaction time on the FAT, a greater variance in ratings on the FDT, and a higher mean ratio of sequential selections in the FPT. The second cluster, labelled the lower cue utilisation typology, consisted of participants, the centroids for whom reflected a greater response latency on the FIT, lower accuracy on the FRT, a lower variance in ratings on the FDT, a lower mean ratio of variance to reaction time on the FAT, and a lower mean ratio of sequential selections in the FPT.

Independent samples *t* tests demonstrated significant differences in FIT, FRT, and FDT mean scores between the higher and lower cue utilisation typologies (see Table 1). In the case of the FDT and FPT, the differences were non-significant. Nevertheless, the pattern of responses was generally consistent with the pattern which would normally be expected to characterise higher or lower cue utilisation. Fourteen participants were classified in the higher cue utilisation typology and 24 participants were classified in the lower cue utilisation typology.

Table 1. *Cluster centroids for the EXPERTise task scores.*

EXPERTise	Cluster 1 (n=14)	Cluster 1 (n=24)			
tasks	Higher cue utilisation	Lower cue utilisation	*t*	df	*p*
FIT	-0.52	0.31	2.66*	36	.012
FRT	0.58	-0.34	3.00**	36	.005
FAT	0.64	-0.37	3.44**	36	.001
FDT	0.20	-0.12	0.94	36	.338
FPT	0.38	-0.23	1.87	36	.117

* Significant at the 0.05 level (two-tailed); **Significant at the 0.01 level (two-tailed)

Covariates

Power distribution experience. Independent samples *t* tests indicated that cue utilisation was not associated with the number of years in which participants had been employed as network controllers, nor the number of years they had worked in power distribution (*p*s > .05). Pearson's correlations indicated that the years that participants had been employed as network controllers and the number of years they had worked in power distribution were not associated with rSO2 levels or saccade amplitude (*p*s > .05). The number of years that participants reported working in power distribution was not associated with fixation rates nor fixation dispersion (*p*s > .05). However, the number of years as a network controller was positively correlated with mean fixation rate during Session 1 (p = .014) and Session 2 (p = .028), and negatively correlated with fixation dispersion during Session 1 (p = .043) and Session 2 (p = .032). Consequently, years of experience as a network controller was included as a covariate for the main analyses involving eye behaviour metrics.

Subjective workload. Mean ratings of subjective workload for each dimension of the NASA-TLX were calculated for each 20-minute testing session. Independent samples *t* tests indicated that cue utilisation was not associated with ratings of physical demands, mental demands, temporal demands, effort, frustration, or performance (*p*s > .05). Pearson's correlations for Session 1 revealed a statistically significant positive association between rSO2 levels and temporal demands, r = .320, p = .023, and a statistically significant positive association between rSO2 levels and effort, r = .323, p = .022. Other correlations between the dimensions of subjective workload and the outcome variables during Session 1 were not statistically significant (*p*s > .05). Pearson's correlations for Session 2 revealed a statistically significant, positive association between rSO2 levels and mental demands, r = .350, p = .014, and a statistically significant positive association between rSO2 levels and effort, r = .375, p = .008. Other correlations between the dimensions of subjective workload and the outcome variables during Session 2 were not statistically significant (*p*s > .05). Consequently, mental demands, temporal demands, and effort were included as covariates for the main analyses involving rSO2.

Cue utilisation and rSO2

To investigate whether operators with higher cue utilisation consume fewer cognitive resources during their everyday work tasks, a 2 x 4 ANCOVA was conducted for each session, with cue utilisation as a between-groups variable (Higher, Lower), time as a within-groups variable (Periods 1-4), mental demands, temporal demands, and effort as covariates, and rSO2 scores as the dependent variable. For Session 1, a statistically significant main effect was evident for cue utilisation, $F(1,33) = 5.21$, $p = .029$, $\eta^2 = .136$, with the lower cue utilisation typology recording significantly greater increases in rSO2 from baseline ($M = 2.21$, $SD = 1.28$) compared to the higher cue utilisation typology ($M = 1.10$, $SD = 1.69$). For Session 1, there was no statistically significant main effect for time, $F(3,99) = 1.21$, $p = .309$, $\eta^2 = .035$, and no statistically significant interaction between cue utilisation and time, $F(3,99) = 1.07$, $p = .366$, $\eta^2 = .031$ (see Figure 1).

For Session 2, a statistically significant main effect was evident for cue utilisation, $F(1,33) = 4.65$, $p = .038$, $\eta^2 = .123$, with the lower cue utilisation typology associated with significantly greater increases in rSO2 from baseline ($M = 2.23$, $SD = 2.05$) compared to the higher cue utilisation typology ($M = 0.95$, $SD = 1.43$). For Session 2, the assumption of sphericity was violated ($p < .001$). Therefore, a Greenhouse-Geisser adjustment was used for all within-subjects' effects. There was no statistically significant main effect for time, $F(1.8,60.1) = 1.96$, $p = .145$, $\eta^2 = .056$, and no statistically significant interaction between cue utilisation and time, $F(1.8,60.1) = 1.24$, $p = .298$, $\eta^2 = .036$ (see Figure 1).

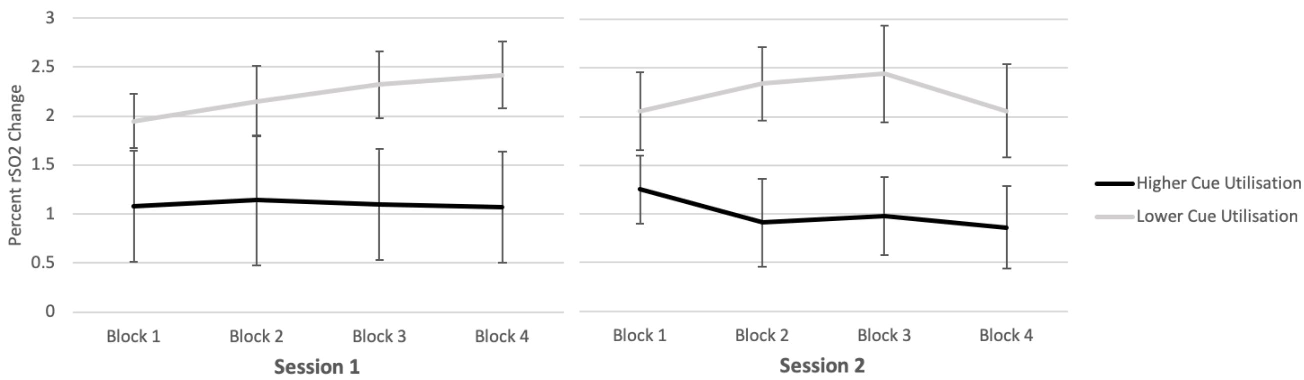

Figure 1. Mean oxygenation scores in the right hemisphere by cue utilisation typology and time during Session 1 (left) and Session 2 (right). Oxygenation scores are based on percent change relative to baseline. Error bars represent ± 1 SE.

Cue utilisation and eye behaviour metrics

To investigate whether participants' information acquisition differed based on their cue utilisation, two 2 x 4 MANCOVAs were conducted with cue utilisation typology as a between-groups variable (Higher, Lower), time as a within-groups variable (Periods 1-4), and years of experience as a network controller as a covariate. The dependent variables for the MANCOVAs comprised fixation rates, saccade amplitude, and fixation dispersion, for Session 1 and Session 2 respectively.

The multivariate effect was not statistically significant by cue utilisation during Session 1, $F(1,33) = 1.44$, $p = .201$, $\eta^2 = .042$, nor during Session 2, $F(1,33) = 0.01$, $p = .985$, $\eta^2 = .001$. There was no statistically significant multivariate effect of time during Session 1, *Wilk's Lambda* = 0.92, $F(3,31) = 0.93$, $p = .440$, $\eta^2 = .082$, nor during Session 2, *Wilk's Lambda* = 0.96, $F(3,31) = 0.80$, $p = .749$, $\eta^2 = .039$. No statistically significant multivariate interaction was evident between time and cue utilisation during Session 1, *Wilk's Lambda* = 0.98, $F(3,31) = 0.24$, $p = .870$, $\eta^2 = .022$, nor during Session 2, *Wilk's Lambda* = 0.92, $F(3,31) = 1.73$, $p =$

.183, $\eta^2 = .147$ (see Figure 2). This indicates that, during both Session 1 and Session 2, there were no differences in eye behaviour metrics based on either cue utilisation or time.

Discussion

The primary aim of this study was to examine whether qualified operators' cue utilisation is associated with the consumption of cognitive resources during regular operational tasks. DNSP control room operators were classified with either higher or lower cue utilisation based on an assessment of cue utilisation within the context of power distribution. During two, 20-minute periods of operators' regular workdays, physiological measures of workload were assessed through changes in cerebral oxygenation in the prefrontal cortex compared to baseline, and through eye behaviour metrics (fixation rates, saccade amplitude, and fixation dispersion).

As higher cue utilisation is associated with the identification of more predictive features and greater efficiencies in information processing (Lansdale et al., 2010; Wiggins, 2015), operators with higher cue utilisation should demonstrate more efficient search patterns, and consume fewer cognitive resources, compared to operators with lower cue utilisation. Consequently, it was hypothesised that, during periods of their regular workday, control room operators with higher cue utilisation would record lower increases in cerebral oxygenation from baseline, lower visual fixation rates, smaller mean fixation dispersions, and smaller mean saccade amplitudes, compared to operators with lower cue utilisation.

The results indicated that there were no statistically significant differences in fixation rates, fixation dispersions, or saccade amplitudes, based on levels of cue utilisation. However, as hypothesised, during both Session 1 and Session 2, operators with higher cue utilisation demonstrated smaller increases in cerebral oxygenation in the prefrontal cortex from baseline, compared to operators with lower cue utilisation. This suggests that operators with higher cue utilisation experience lower levels of cognitive load during periods of their regular workday, compared to operators with lower cue utilisation.

Theoretical and practical implications

The results of the present study provide support for the assertion that a greater capacity for cue utilisation is associated with the consumption of fewer cognitive resources during operational tasks. The outcomes replicate previous research, which has inferred that higher cue utilisation is associated with lower cognitive load during process control tasks, by demonstrating that higher cue utilisation is associated with smaller increases in cerebral oxygenation in the prefrontal cortex (Sturman et al., 2019; Sturman & Wigggins, in submission). However, operational tasks in these previous experiments consisted of laboratory simulations, which likely contained subtle differences in feature-event relationships, compared to operational environments to which participants were typically exposed. Consequently, it was unclear whether context-based cue utilisation would predict cognitive load amongst qualified operators during regular operational tasks.

The present study extends previous research, providing support for the proposition that qualified operators' cue utilisation predicts cognitive load during everyday operational tasks. Further, differences in cerebral oxygenation were evident, controlling for participants' subjective ratings of mental demands, temporal demands, and effort. This indicates that, when completing tasks rated as being similarly demanding, operators with higher cue utilisation consume fewer cognitive resources when completing those tasks.

Attentional Resource Theory is based on the proposition that cognitive resources are drawn from a limited supply, and that the consumption of cognitive resources during operational tasks reduces the availability of residual cognitive resources (Kahneman, 1973; Warm, Parasuraman, & Matthews, 2008; Wickens, 1980). The outcomes of the present study indicate that operators with lower cue utilisation are likely to experience lower cognitive load during routine operational tasks, and consume fewer cognitive resources over a specified time period, compared to operators with higher cue utilisation. A reduction in the availability of cognitive resources is associated with a reduction in accuracy and an increase in operational errors when responding to critical signals (Reason, 1990; Wickens, Hollands, Banbury, & Parasuraman, 2015). Consequently, operators with lower cue utilisation are potentially more likely to demonstrate a decline in performance during operational tasks, compared to those operators with higher cue utilisation.

Greater residual cognitive resources are posited to allow operators with higher cue utilisation to better manage the demands of secondary tasks (Brouwers et al., 2017; Wickens, 2002). Therefore, operators with higher cue utilisation should be able to maintain a higher level of performance during more demanding periods of work, compared to operators with lower cue utilisation. Consequently, assessments of cue utilisation may aid in the selection of job applicants who are better able to sustain attention and maintain performance during demanding situations.

In addition to aiding the selection of operators, assessments of cue utilisation could be used to support the training and professional development of operators. For instance, the ability to predict operators' cognitive load could potentially be used to improve job performance by optimising the length of time between breaks for individual operators. Alternatively, assessments of cue utilisation could help identify operators who would benefit from cue-based training interventions, whereby operators are given the opportunity to acquire cues that can be generalised to the broader operational environment (Klayman, 1988; Scherer et al., 2008; Wiggins, 2015; Wiggins & O'Hare, 2003b). Finally, selecting operators using assessments of cue utilisation that demonstrate predictive validity for technical training may reduce training failure.

Limitations and future direction

The present study is limited by the lack of experimental control which can occur during field testing. For instance, due to time constraints and operators' work schedules, testing sessions for different operators were conducted at different times of the day and during periods of relatively higher or lower work demands. To help control for variances in work demands, two sessions were conducted for each operator during different periods of their shift. Further, subjective self-reports of workload were collected to statistically control for differences in work demands. Although a statistically significant difference in cerebral oxygenation was evident, taking into account mental workload, this was not the case for eye behaviour metrics.

In addition to differences in work demands, operational tasks varied between testing sessions and operators. For example, different control room operations require differing degrees of interaction with Supervisory Control and Data Acquisition (SCADA) displays, face-to-face communication, and radio communication. These different operational tasks are likely to require different visual search patterns. Consequently, the lack of statistically significant differences in eye behaviour metrics based on cue utilisation may be due to the variability in the operational tasks completed by each control room operator.

The cross-sectional design of the present study also limits the extent to which causal relationships can be established. For instance, operators with higher cue utilisation may be more likely to be assigned to tasks with greater or lesser demands, or may opt to take breaks more or less frequently, compared to operators with lower cue utilisation. As these factors are likely to influence cognitive load, differences in cerebral oxygenation based on cue utilisation could potentially be explained by environmental features.

To control for the variability in operational tasks and work demands, future research could utilise high fidelity power control simulations. The use of simulations would allow eye behaviour metrics and cerebral oxygenation to be assessed in a controlled environment. Further, simulations would allow the consumption of cognitive resources and performance to be assessed at successive time points and under similar conditions, which may be beneficial in establishing a causal relationship between cue utilisation and the consumption of cognitive resources.

Conclusion

The current study was designed to determine whether qualified operators' cue utilisation is associated with cognitive load during regular operational tasks. Physiological measures of cognitive resource consumption were assessed during two, 20-minute periods of power distribution network controllers' regular work days. During both sessions of testing, operators with higher cue utilisation demonstrated smaller increases in cerebral oxygenation in the prefrontal cortex from baseline, compared to operators with lower cue utilisation. However, cue utilisation was not associated with differences in eye behaviour. The results of the study are consistent with the proposition that power operators with higher cue utilisation experience differences in cognitive load during regular operational tasks, compared to operators with lower cue utilisation.

References (Study 6)

Bertram, R., Helle, L., Kaakinen, J. K., & Svedström, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PloS One*, *8*, e66169. doi:10.1371/journal.pone.0066169

Brouwers, S., Wiggins, M. W., Griffin, B., Helton, W. S., & O'Hare, D. (2017). The role of cue utilisation in reducing the workload in a train control task. *Ergonomics*, *60*, 1–16. doi:10.1080/00140139.2017.1330494

Brouwers, S., Wiggins, M. W., Helton, W., O'Hare, D., & Griffin, B. (2016). Cue utilization and cognitive load in novel task performance. *Frontiers in Psychology*, *7*, 1–12. doi:10.3389/fpsyg.2016.00435

Brunswik, E. (1955). Representative Design and Probabilistic Theory in a Functional Psychology. *Psychological Review*, *62*, 193–217. doi:10.1037/h0047470

Camilli, M., Terenzi, M., & Di Nocera, F. (2007). Concurrent validity of an ocular measure of mental workload. In *Human factors issues in complex system performance. In D. de Waard, G.R.J. Hockey, P. Nickel, and K.A. Brookhuis (Eds.)* (pp. 117–129). Shaker Publishing Maastricht, the Netherlands.

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, *27*, 951–964. doi:10.1068/p270951

Chung, P. H., & Byrne, M. D. (2008). Cue effectiveness in mitigating postcompletion errors in a routine procedural task. *International Journal of Human-Computer Studies*, *66*, 217–232. doi:10.1016/j.ijhcs.2007.09.001

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A Random Glance at the Flight Deck: Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload. *Journal of Cognitive Engineering and Decision Making*, *1*, 271–285.

doi:10.1518/155534307X255627.

Ekkekakis, P. (2009). Illuminating the Black Box: Investigating Prefrontal Cortical Hemodynamics during Exercise with Near-Infrared Spectroscopy. *Journal of Sport and Exercise Psychology*, *31*, 505–553. doi:10.1123/jsep.31.4.505

Evans, D. C., & Fendley, M. (2017). A multi-measure approach for connecting cognitive workload and automation. *Journal of Human Computer Studies*, *97*, 182–189. doi:10.1016/j.ijhcs.2016.05.008

Fu, R., Guo, Y., Chen, Y., Yuan, W., Ma, Y., Peng, J., … Wang, C. (2011). Research on heart rate and eye movement as indicators of drivers' mental workload. In *Road Safety and Simulation* (pp. 1–20).

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, *24*, 631–645. doi:10.1016/S0169-8141(98)00068-7

Gratton, G., & Fabiani, M. (2006). Optical imaging of brain function in R. Parasuraman, M. Rizzo (Eds.), Neuroergonomics: The Brain At Work, Oxford University. In *Oxford Series in Human-Technology Interactions* (pp. 65–81). Cambridge, MA.

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

Grüner, M., & Ansorge, U. (2017). Mobile Eye Tracking During Real-World Night Driving: A Selective Review of Findings and Recommendations for Future Research. *Journal of Eye Movement Research*, *1011*, 1–18. doi:10.16910/jemr.10.2.1

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (Vol. 52, pp. 139–183). Amsterdam: North-Holland: Elsevier.

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*, 429–437. doi:10.1007/s00221-011-2749-1

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

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*, 429–436. doi:10.1007/s00221-012-3322-2

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*, 18–25. doi:10.1016/j.actpsy.2008.04.002

Helton, W. S., Warm, J. S., Tripp, L. D., Matthews, G., Parasuraman, R., & Hancock, P. A. (2010). Cerebral lateralization of vigilance: A function of task difficulty. *Neuropsychologia*, *48*, 1683–1688. doi:10.1016/j.neuropsychologia.2010.02.014

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ.

Klayman, J. (1988). Cue discovery in probabilistic environments: Uncertainty and experimentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 317. doi:10.1037//0278-7393.14.2.317

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. In *Journal of Cognitive Engineering and Decision Making* (Vol. 4, pp. 186–209). SAGE Publications Sage CA: Los Angeles, CA. doi:10.1518/155534310x12844000801203

Lansdale, M., Underwood, G., & Davies, C. (2010). Something Overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, *24*, 213–225.

doi:10.1002/acp.1552

Loveday, T., Wiggins, M. W., & Searle, B. J. (2013). Cue utilization and broad indicators of workplace expertise. *Journal of Cognitive Engineering and Decision Making*, *8*, 98–113. doi:10.1177/1555343413497019

Loveday, T., Wiggins, M. W., Searle, B. J., Festa, M., & Schell, D. (2013). The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*, 125–137. doi:10.1177/0018720812448475

MacLean, K. A., Aichele, S. R., Bridwell, D. A., Mangun, G. R., Wojciulik, E., & Saron, C. D. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception, & Psychophysics*, *71*, 1042–1058. doi:10.3758/app.71.5.1042

Matthews, G., Warm, J. S., Shaw, T. H., & Finomore, V. S. (2010). A multivariate test battery for predicting vigilance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*, 1072–1076. doi:10.1177/154193121005401405

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors*, *42*, 36–55.

O'Hara, J. M., & Hall, R. E. (1992). Advanced control rooms and crew performance issues: implications for human reliability. *IEEE Transactions on Nuclear Science*, *39*, 919–923. doi:10.1109/23.159733

Ong, M., Russell, P. N., & Helton, W. S. (2013). Frontal cerebral oxygen response as an indicator of initial attention effort during perceptual learning. *Experimental Brain Research*, *229*, 571–578. doi:10.1007/s00221-013-3619-9

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*, 924–927. doi:10.1126/science.472714

Pauley, K., O'Hare, D., & Wiggins, M. W. (2009). Measuring expertise in weather-related aeronautical risk perception: The validity of the Cochran–Weiss–Shanteau (CWS) Index. *The International Journal of Aviation Psychology*, *19*, 201–216. doi:10.1080/10508410902979993

Perry, N. C., Wiggins, M. W., Childs, M., & Fogarty, G. (2012). Can reduced processing decision support interfaces improve the decision-making of less-experienced incident commanders? *Decision Support Systems*, *52*, 497–504. doi:10.1016/j.dss.2011.10.010

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903. doi:10.1037/0021-9010.88.5.879

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211–219). Hershey, PA: IGI Publishing. doi:10.4018/978-1-59140-562-7.ch034

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge university press.

Reason, J. (2000). Human error: Models and management. *BMJ*, *320*, 768–770. doi:10.1136/bmj.320.7237.768

Scherer, R., Lee, F., Schlogl, A., Leeb, R., Bischof, H., & Pfurtscheller, G. (2008). Toward self-paced brain–computer communication: navigation through virtual worlds. *IEEE Transactions on Biomedical Engineering*, *55*, 675–682. doi:10.1109/tbme.2007.903709

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50*, 864–878. doi:10.1037/e578082012-006

Schvaneveldt, R. W., Beringer, D. B., & Lamonica, J. A. (2001). Priority and organization of information accessed by pilots in various phases of flight. *The International Journal of Aviation Psychology*, *11*, 253–280. doi:10.1037/e406222004-001

Shaw, T. H., Funke, M. E., Dillard, M., Funke, G. J., Warm, J. S., & Parasuraman, R. (2013). Event-related cerebral hemodynamics reveal target-specific resource allocation for both "go" and "no-go" response-based vigilance tasks. *Brain and Cognition*, *82*, 265–273. doi:10.1016/j.bandc.2013.05.003

Shaw, T. H., Warm, J. S., Finomore, V., Tripp, L., Matthews, G., Weiler, E., & Parasuraman, R. (2009). Effects of sensory modality on cerebral blood flow velocity during vigilance. *Neuroscience Letters*, *461*, 207–211. doi:10.1016/j.neulet.2009.06.008

Small, A. J., Wiggins, M. W., & Loveday, T. (2014). Cue-based processing capacity, cognitive load and the completion of simulated short-duration vigilance tasks in power transmission control. *Applied Cognitive Psychology*, *28*, 481–487. doi:10.1002/acp.3016

Smit, A. S., Eling, P. A. T. ., & Coenen, A. M. . (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica*, *115*, 35–42. doi:10.1016/j.actpsy.2003.11.001

Stanton, N. A., Salmon, P., Walker, G. H., & Jenkins, D. (2009). *Human factors in the design and evaluation of central control room operations*. CRC Press. doi:10.1201/9781439809921

Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*, 435–446. doi:10.1016/j.trf.2011.04.008

Vicente, K. J., Roth, E. M., & Mumaw, R. J. (2001). How do operators monitor a complex, dynamic work domain? The impact of control room technology. *International Journal of Human-Computer Studies*, *54*, 831–856. doi:10.1006/ijhc.2001.0463

Warm, J. S. (1984). *Sustained attention in human performance*. New York: Wiley.

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

Weiss, D. J., & Shanteau, J. (2003). Empirical assessment of expertise. *Human Factors*, *45*, 104–116. doi:10.1518/hfes.45.1.104.27233

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickers (Ed.), *Attention and Performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159–177. doi:10.1080/14639220210123806

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. Psychology Press.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology*, *25*, 229–235. doi:10.1002/acp.1668

Wiggins, M. W. (2012). The role of cue utilisation and adaptive interface design in the management of skilled performance in operations control. *Theoretical Issues in Ergonomics Science*, *15*, 283–292. doi:10.1080/1463922X.2012.724725

Wiggins, M. W. (2014). Differences in situation assessments and prospective diagnoses of simulated weather radar returns amongst experienced pilots. *International Journal of Industrial Ergonomics*, *44*, 18–23. doi:10.1016/j.ergon.2013.08.006

Wiggins, M. W. (2015). Diagnosis and Instructional Systems Design. In M. W. Wiggins & T. Loveday (Eds.), *Diagnostic Expertise in Organizational Environments* (pp. 69–80). Surrey, England: Ashgate Publishing Limited, Surrey, England.

Wiggins, M. W., Brouwers, S., Davies, J., & Loveday, T. (2014). Trait-based cue utilization and initial skill acquisition: Implications for models of the progression to expertise.

Frontiers in Psychology, *5*, 1–8. doi:10.3389/fpsyg.2014.00541

Wiggins, M. W., Loveday, T., & Auton, J. C. (2015). *EXPERT Intensive Skills Evaluation (EXPERTise) Test*. Sydney: Macquarie University.

Wiggins, M. W., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, *1*, 305–320. doi:10.1037/1076-898X.1.4.305

Wiggins, M. W., & O'Hare, D. (2003a). Expert and novice pilot perceptions of static in-flight images of weather. *International Journal of Aviation Psychology*, *13*, 173–187. doi:10.1207/S15327108IJAP1302_05

Wiggins, M. W., & O'Hare, D. (2003b). Weatherwise: Evaluation of a Cue-Based Training Approach for the Recognition of Deteriorating Weather Conditions during Flight. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*, 337–345. doi:10.1518/hfes.45.2.337.27246

Wiggins, M. W., Stevens, C., Howard, A., Henley, I., & O'Hare, D. (2002). Expert, intermediate and novice performance during simulated pre-flight decision-making. *Australian Journal of Psychology*, *54*, 162–167. doi:10.1080/00049530412331312744

Williams, A. M., Ward, P., Knowles, J. M., & Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied*, *8*, 259–270. doi:10.1037/1076-898X.8.4.259

General Discussion

Cues refer to recognition-driven associations between situation-specific environmental features and an event or object (Brunswik, 1955; Klein, Calderwood, & Clinton-Cirocco, 1986). Through experience, operators develop a repertoire of patterns of cues that describe the causal factors in a situation (Klein, 1993, 2003). These patterns are stored in long-term memory, and provide information regarding the type of situation, including plausible goals, cues to monitor, expectancies about the situation, and typical reactions (Coderre, Mandin, Harasym, & Fick, 2003; Croskerry, 2009; Ericsson & Kintsch, 1995; Klein, 1993, 2003). In addition to enabling rapid and appropriate responses, often in complex and dynamic situations (Beilock, Bertenthal, McCoy, & Carr, 2004; Ericsson & Lehmann, 1996; Salthouse, 1991; Klein, Calder-wood, & Clinton-Cirocco, 1986; 2010; Klein, 1998), the retrieval and activation of cues from long-term memory has the advantage of imposing relatively fewer demands on working memory resources (Chung & Byrne, 2008; Evans, 2008; Norman & Shallice, 1986).

Cue utilisation refers to the application of cue-based processing (Lansdale, Underwood, & Davies, 2010). Individual differences in cue utilisation derive from differences in the capacity of individuals to identify predictive features in the environment, associate these features and events in memory, retain these cue-based associations, and appropriately apply cues in response to environmental features (Wiggins, 2012, 2015). Given that the activation of cues appears to reduce cognitive load (Chung & Byrne, 2008; Evans & Fendley, 2017), it has been argued that operators with higher cue utilisation should consume fewer cognitive resources per unit time during sustained attention tasks, compared to those operators with lower cue utilisation (Brouwers, Wiggins, Helton, O'Hare, & Griffin, 2016; Small, Wiggins, & Loveday, 2014). Consequently, these operators should retain greater residual cognitive resources, thereby enabling them to sustain attention for longer periods (Brouwers, Wiggins, Griffin, Helton, & O'Hare, 2017; Matthews et al., 2010).

Evidence supporting the proposition that higher cue utilisation is associated with a slower consumption of cognitive resources relies primarily on inferences derived from changes in mean response latency during sustained attention tasks (Brouwers et al., 2016, 2017; Small, Wiggins, & Loveday, 2014). For instance, Brouwers et al. (2016) observed that, during a semi-automated process control task, the addition of a concurrent, secondary cognitive task was associated with increases in response latency over time for participants with lower cue utilisation. However, the imposition of a secondary task had no impact on the response latency of participants with higher cue utilisation. Brouwers et al. (2016) suggest that participants with higher cue utilisation were able to adopt a strategy during the sustained attention task that reduced the rate at which their cognitive resources were consumed, thereby providing greater residual cognitive resources that subsequently minimised the impact of the secondary task.

Despite differences in response latencies indicating differences in cognitive load, associations between cue utilisation and response latency could potentially be explained by alternative factors, including participants' level of motivation or engagement. Further, assessments of cue utilisation and sustained attention both typically rely on measures of response latency and accuracy, and data for each participant is typically collected in a single session (Brouwers et al., 2016, 2017; Small et al., 2014). Consequently, the relationship between cue utilisation and the demand for cognitive resources may be partially attributable to common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Finally, previous studies have relied primarily on cross task cue utilisation, whereby cue utilisation evaluated in one context (e.g., driving) is used to predict cognitive load in another novel context (e.g., rail control). As novice operators in these studies had no prior opportunities to acquire relevant cues, differences in cognitive load based on cue utilisation likely reflect differences in the rate of cue acquisition during the novel tasks. Consequently, it was previously unclear whether differences in cognitive load based on cue utilisation would also be evident amongst qualified personnel in familiar operating environments.

The overall aim of the present programme of research was to examine whether differences in cue utilisation are associated with differences in the consumption of cognitive resources during sustained attention tasks. Research questions pertaining to the overall aim included whether:

(a) differences in sustained attention based on cue utilisation differ in process control environments compared to monitoring environments? (Studies 1 and 2);

(b) cue utilisation also differentiates the performance of qualified operators during domain relevant sustained attention tasks? (Studies 3, 4, and 5);

(c) differences in cue utilisation are associated with differences in physiological measures of cognitive load during novel tasks? (Studies 2, 5, and 6);

(d) differences in cognitive load are also observed in more dynamic operational environments? (Study 5); and whether

(e) differences in physiological measures of cognitive load based on cue utilisation are also evident amongst qualified operators during familiar operational tasks? (Study 6)

The six studies presented in this dissertation were designed to investigate these specific research questions.

Study 1 was designed to examine whether differences in cue utilisation are associated with differences in performance decrements during novel monitoring and process control sustained attention tasks. The study was conducted with motor vehicle drivers, who undertook an assessment of cue utilisation using the driving version of the Expert Skills Evaluation (EXPERTise 2.0; Wiggins, Brouwers, Davies, & Loveday, 2014). Participants also completed a 30-minute rail control simulation containing implicit patterns of train movement, that required trains to be rerouted either infrequently (monitoring task) or periodically (process control task). As higher cue utilisation is posited to be associated with a reduction in the rate at which cognitive resources are consumed, it was hypothesised that higher cue utilisation would be associated with smaller increases in response latency during the sustained attention tasks. Further, based on the proposition that monitoring tasks demand greater effort than process control tasks to sustain attention, it was hypothesised that differences in response latencies over time between participants with higher or lower cue utilisation would be greater in a monitoring task condition compared to a process control task condition.

The results from Study 1 indicated that, as anticipated, in the monitoring condition, participants with higher cue utilisation experienced a smaller increase in response latency during the rail control task, compared to participants with lower cue utilisation. This provides support for the proposition that participants with a greater general capacity for cue utilisation adopted a strategy that reduced the rate at which their cognitive resources were consumed during the period, thereby contributing to the maintenance of a higher level of performance over an extended period. However, no such differences were evident in the process control condition. One explanation for the differences in task-related outcomes is that, to accurately measure the consumption of cognitive resources through changes in response latency, the task must be undertaken over a sufficiently lengthy duration and/or with appropriately high levels of workload to consume cognitive resources, and thereby elicit declines in performance. As the rail control task was specifically designed to reflect a 'low demand' environment, any reduction in cognitive load afforded by the utilisation of cues may not have been sufficiently large to influence response latencies across the different cue utilisation typologies in the process control task. Consequently, it was reasoned that, to assess differences in cognitive load based on cue utilisation, sustained attention tasks of longer duration and/or alternative measures of cognitive load were required.

Study 2 was designed to replicate the response latency findings from Study 1, while increasing the length of the sustained attention task to examine whether differences in performance decrements based on cue utilisation become more apparent during longer process control tasks. To further explore the relationship between cue utilisation and the consumption of cognitive resources, physiological measures were included as correlates to triangulate the outcomes. Based on the proposition that participants with higher cue utilisation would rapidly identify and attend to relevant task features, thereby reducing the demands on cognitive resources, it was hypothesised that these participants would record a relatively greater reduction in visual fixation rates and cerebral oxygenation over the period of the task, compared to participants with lower cue utilisation. During a 45-minute version of the rail control task, lower cue utilisation was associated with a greater increase in response latency over time compared to higher cue utilisation. Further, consistent with expectations, participants with higher cue utilisation demonstrated smaller increases from baseline in cerebral oxygenation in the prefrontal cortex, and a greater decrease in the frequency of fixations throughout the rail control task, compared to participants with lower cue utilisation.

That higher cue utilisation predicted a greater decrease in fixation rates is consistent with the proposition that a greater capacity for cue utilisation is associated with the identification of implicit patterns, and that this enables attention to be directed towards task features of greater relevance, thereby reducing the overall number of features to which operators need attend to within a given period of time (Brouwers et al., 2016; McCormack, Wiggins, Loveday, & Festa, 2014; Williams, Ward, Knowles, & Smeeton, 2002). Further, the differences observed in cerebral oxygenation suggest that a propensity to identify critical cues and rapidly establish feature-event relationships provides an opportunity to reduce the rate at which cognitive resources are consumed. In combination, these findings provide evidence to suggest that differences in performance decrements based on cue utilisation reflect differences in the consumption of cognitive resources.

While the outcomes from Studies 1 and 2 provided evidence indicating that a general capacity for cue utilisation predicts differences in the consumption of cognitive resources during novel sustained attention tasks, these studies relied on cross-task cue utilisation, where cue utilisation in the domain of driving was used to predict performance during a novel rail control task. As participants had no previous exposure to rail control, the differences observed in response latency and cognitive load based on cue utilisation in these studies likely reflected differences in the speed with which participants acquired cue-based relationships during the rail control task. However, as qualified system operators in industrial environments have typically had prior opportunities to acquire cues within their domain of expertise, it remained unclear whether similar effects would be observed amongst qualified operators. Consequently, subsequent studies in the present programme of research were designed to assess the relationship between qualified operators' cue utilisation, sustained attention, and cognitive load.

Due to time constraints and costs, the availability of experienced practitioners in an operational context is often restricted, which limits the utility of longer-duration sustained attention tasks for research within this population. Therefore, to facilitate greater participation rates in future assessments of qualified operators, there was a need to develop a shorter-duration, sustained attention task for process control environments. Study 3 was designed to validate a modified short-duration power control sustained visual search task. As shorter-duration vigilance tasks have demonstrated similar effects to those achieved with longer-duration vigilance tasks (Posner, 1978; Rosvold, Mirsky, Sarason, Bransome, & Beck, 1956), it was hypothesised that performance on the short-duration, sustained visual search task would predict performance on a longer-duration rail control task. The results revealed changes in response latency throughout the sustained visual search task that were positively associated with changes in response latency during a 30-minute monitoring version of the rail control task, a 45-minute monitoring version of the rail control task, and a 45-minute process control version of the rail control task. The findings indicated that the sustained visual search task constitutes a valid alternative to a longer-duration process control task for experimental studies.

Utilising the sustained visual search task validated during Study 3, Study 4 was designed to examine whether qualified operators' cue utilisation differentiated performance during a domain-relevant, sustained attention task. The sustained visual search task allowed context-based assessments of cue utilisation, whereby cue utilisation and task performance are assessed in the same domain. Participants in Study 4 consisted of qualified Distribution Network Service Provider (DNSP) system operators, who were recruited from 12 Australian DNSPs for two experiments. Cue utilisation was assessed using a power distribution version of EXPERTise, while sustained attention was assessed using the sustained visual search task. In both Experiment 1 and Experiment 2, power distribution operators with higher cue utilisation demonstrated shorter mean response latencies during the sustained visual search task, compared to operators with lower cue utilisation. Further, no differences in accuracy based on cue utilisation were observed during the sustained visual search task. The results are consistent with the proposition that power operators with higher cue utilisation retain greater residual cognitive resources during domain-related sustained attention tasks, compared to operators with lower cue utilisation, which likely allows them to rapidly and accurately manage changes in the system state.

These outcomes suggest that differences in cognitive load based on cue utilisation observed amongst novice operators, are also evident amongst qualified operators who have had prior opportunities to acquire relevant cue-based associations. However, as in the rail control task, the sustained visual search task relied on a simplified representation of an operating system, and therefore, likely contained a number of different features to those encountered during operators' regular workdays. Consequently, it was unclear to what extent existing cue-based associations allowed operators to identify relevant features and subsequently reduce their cognitive load, and to what extent differences in response latency reflected differences in the rate of cue acquisition during the task. Therefore, there was a need to establish whether qualified operators' cue utilisation differentiates the consumption of cognitive resources in more complex and dynamic operational environments containing feature-event relationships that are likely to be present during regular operational tasks.

Study 5 was designed to examine whether cue utilisation differentiates qualified operators' cognitive resource consumption during a dynamic sustained attention task. As in Studies 1 and 2, Study 5 was conducted with motor vehicle drivers, who undertook an assessment of cue utilisation using the driving version of EXPERTise 2.0. However, participants in Study 5 also completed a 20-minute simulated driving task, during which eye behaviour metrics and cerebral oxygenation in the prefrontal cortex were recorded. Dynamic tasks, such as driving, contain emerging features which are less predictable than those that are present in repetitious control tasks. Consequently, to reduce cognitive load in less predictable dynamic environments, purposive sampling, whereby operators visually fixate on key features within close proximity (Henderson, 2003; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003), was expected to increase efficiency by reducing the time and effort spent scanning for emerging features (Watson, Brennan, Kingstone, & Enns, 2010). Based on this proposition, it was hypothesised that, during the simulated driving task, participants with higher cue utilisation would demonstrate physiological responses indicative of more efficient search patterns and lower cognitive load, compared to participants with lower cue utilisation.

As hypothesised, the outcomes of Study 5 indicated that participants with higher cue utilisation recorded smaller mean visual saccade amplitudes, smaller mean fixation dispersions, and smaller increases in cerebral oxygenation, while recording fewer missed traffic signals, compared to participants with lower cue utilisation. As in Study 2, these results suggest that greater cue utilisation was associated with more efficient search patterns. Further, in contrast to Study 2, differences in eye behaviour metrics were apparent during the first 5-minute period of the simulated driving task. The association between cue utilisation and differences in eye behaviour from the onset of the experimental task provides support for the proposition that, compared to qualified operators with lower cue utilisation, qualified operators with higher cue utilisation are better able to draw on existing patterns of cues stored in long-term memory to anticipate and respond efficiently to patterns of features in familiar operational environments.

The results arising from Study 5 also indicated that, despite consuming fewer cognitive resources, participants with higher cue utilisation were able to maintain a greater level of performance during the simulated driving task, compared to those participants with lower cue utilisation. This suggests that differences in cognitive resource consumption based on cue utilisation were not due to reductions in effort at the cost of performance. Rather, the results are consistent with models of cue utilisation, which posit that the activation of patterns of cues is a non-conscious process that enables situations to be recognised as familiar, thereby facilitating sound and rapid responses to environmental stimuli (Anderson, 1982; Klein, 1993; Wiggins, 2015).

In combination, the outcomes from Studies 1 to 5 are consistent with the proposition that a greater propensity to identify predictive features in the environment, and appropriately apply cues in response to these environmental features, is associated with more efficient search patterns and a relatively lesser consumption of cognitive resources per unit of exposure. The experimental paradigms used in these studies provided a degree of experimental control, enabling differences in the consumption of cognitive resources to be identified. However, controlled experimental studies necessitate the use of simplified operational environments, which differ from typical operational environments in terms of the complexity of feature-event relationships. Consequently, there was a need to assess the relationship between cue utilisation and cognitive load in real operational environments.

Study 6 was designed to evaluate whether power distribution operators' cue utilisation differentiates cognitive load during regular operational tasks. System operators from four Australian DNSP control rooms participated in the study. During two, 20-minute periods of operators' regular workdays, physiological measures of workload were assessed through changes in cerebral oxygenation in the prefrontal cortex compared to baseline, and through eye behaviour metrics (fixation rates, saccade amplitude, and fixation dispersion). Operators also completed a power distribution version of EXPERTise. The results indicated that there were no statistically significant differences in eye behaviour metrics, based on levels of cue utilisation. However, as hypothesised, during both sessions, controlling for subjective ratings of workload, operators with higher cue utilisation demonstrated relatively smaller increases in cerebral oxygenation in the prefrontal cortex from baseline, compared to operators with lower cue utilisation.

Smaller increases in cerebral oxygenation are consistent with the proposition that operators with higher cue utilisation experience lower cognitive load during process control sustained attention tasks, compared to operators with lower cue utilisation. The results of Study 6 also extend previous research, providing support for the notion that qualified operators' cue utilisation predicts differences in the allocation of cognitive resources during everyday operational tasks. Therefore, even in complex operating environments, in which operators complete a range of different tasks, cue utilisation is associated with lower cognitive load.

Taken together, the results of Studies 1-6 suggest that, during process control tasks, cue utilisation is associated with: (a) more efficient search patterns, (b) a lower consumption of cognitive resources per unit time, and (c) greater sustained attention. Further, these findings indicate that these effects occur in both laboratory and operational settings, and in cross-task and task-related cue utilisation with both novice and qualified operators.

Theoretical contributions

The present programme of research has contributed four theoretical outcomes. The first contribution relates to evidence supporting the proposition that cue utilisation is associated with lower cognitive load during sustained attention tasks. As the automatic activation and retrieval of cues from long-term memory has the advantage of imposing relatively fewer demands on working memory resources, cue utilisation should be associated with lower cognitive load during operational tasks (Chung & Byrne, 2008; Klein et al., 1986; Norman & Shallice, 1986). While there is some empirical evidence to support this proposition, research examining cognitive load and cue utilisation has predominantly relied on inferences derived from mean response latencies, and the use of naïve practitioners engaging in simplified representations of operational environments (Brouwers et al., 2016, 2017; Small et al., 2014). Consequently, there has hitherto been a lack of complementary empirical evidence supporting the relationship between cue utilisation and cognitive load amongst both novice and qualified operators in a range of operational environments, and using alternative measures of cognitive load.

The results from Studies 2 and 5 provide support for the proposition that differences in response latency based on cue utilisation result from differences in the rate at which cognitive resources are consumed. Arguably, previously demonstrated differences in response latencies between participants with higher and lower cue utilisation (e.g., Brouwers et al., 2016, 2017) could be explained by differences in effort or motivation. However, if shorter mean response latencies, demonstrated by participants with higher cue utilisation, are the result of greater effort, these participants should also demonstrate greater increases in cerebral oxygenation, compared to participants with lower cue utilisation.

Instead, the results of Studies 2 and 5 indicated that, despite demonstrating greater performance during the sustained attention tasks, participants with higher cue utilisation also demonstrated smaller increases in cerebral oxygenation from baseline, compared to participants with lower cue utilisation. Similarly, while no performance measures were recorded during in situ testing, the results of Study 6 also indicate that qualified operators' cue utilisation is associated with lower cognitive load during regular operational tasks. Importantly, these findings provide support for the proposition that, across a range of operational environments, cue utilisation reduces cognitive load and the rate at which cognitive resources are consumed.

A further limitation of previous research examining cue utilisation and cognitive resource consumption was the use of cross-task cue utilisation, whereby cue utilisation assessed in one domain (e.g., driving) was used to predict cognitive resource consumption in another domain (e.g., rail control). As cues are developed through experience, cross-task cue utilisation allows researchers to examine the association between participants' general capacity for cue utilisation and performance on a novel task. However, novice operators differ from experienced system operators in that they have had no prior opportunity to acquire cue-based associations through exposure to the operational environment. Consequently, the differences in response latency based on cue utilisation that were evident in past studies (e.g., Brouwers et al., 2016, 2017) reflected differences in the speed with which participants acquired new cue-based relationships during the experimental tasks.

To overcome the limitations of cross-task cue utilisation, Studies 4 and 5 relied on context-based assessments of cue utilisation, whereby cue utilisation and task performance were assessed in the same domain. The results of Studies 4 and 5 indicate that differences in sustained attention and the consumption of cognitive resources based on cue utilisation are also evident amongst qualified operators undertaking familiar tasks. Furthermore, while differences in sustained attention based on cue utilisation emerged over time in Studies 1 and 2, differences in performance and the consumption of cognitive resources were demonstrated throughout the entire sustained attention tasks in Studies 4 and 5. This suggests that, in contrast to novice operators who were required to acquire new cues throughout the experimental tasks, qualified operators with higher cue utilisation were able to utilise previously developed cues stored in long-term memory. The association between cue utilisation and cognitive load demonstrated during in situ testing in Study 6 further supports this proposition, and demonstrates that the association between cue utilisation and the allocation of cognitive resources is not simply a laboratory phenomenon.

The second theoretical contribution is the identification of search pattern efficiency as a potential mediator between cue utilisation and reduced cognitive load. The results of Study 2 indicated that the visual fixation rates of novice operators with higher cue utilisation decreased during sustained attention tasks, whereas fixation rates remained relatively high for those participants with lower cue utilisation. This pattern of results supports the proposition that, during tasks containing implicit patterns, cue utilisation enables attention to be directed towards task features of greater relevance, thereby reducing the overall number of features to which operators attend within a given period of time (Brouwers et al., 2016; McCormack et al., 2014; Williams et al., 2002).

A reduction in fixation rates is indicative of reduced visual scanning. Therefore, during repetitious tasks, where the location of relevant features is predictable, a lower frequency of visual fixations indicates less effort is being invested scanning for features (Watson et al., 2010). However, in more dynamic tasks, such as driving, emerging features are likely to be less predictable, compared to repetitious control tasks. In dynamic environments, reduced fixations in the absence of any other strategies may be inefficient, as this would reduce the number of emerging features to which an operator could attend. Therefore, Study 5 was designed to examine a range of eye behaviour metrics associated with search pattern efficiency during a dynamic driving task. The results confirmed that cue utilisation was associated with eye behaviours, including lower fixation dispersions and saccade amplitudes, which are indicative of more efficient search patterns (Camilli, Terenzi, & Di Nocera, 2007; Di Nocera, Camilli, & Terenzi, 2007).

Study 6 was designed to examine whether similar differences in search patterns based on cue utilisation were evident during regular operational tasks. While no differences in visual search patterns were observed based on cue utilisation, this is likely due to the variance in operational tasks completed by different operators during the in situ testing sessions. Nevertheless, in combination, the results of Studies 2 and 5 suggest that search pattern efficiency may partially mediate the relationship between cue utilisation and reduced cognitive load during sustained attention tasks.

The third theoretical contribution is evidence to support a resource depletion account of the vigilance decrement (Grier et al., 2003; Helton & Russell, 2012; Kahneman, 1973). In Studies 2 and 5, cue utilisation was associated with smaller increases in cerebral oxygenation, which is indicative of lower cognitive load. Further, in these studies, participants with higher cue utilisation demonstrated greater sustained attention, compared to participants with lower cue utilisation. These patterns of results indicate that a lower consumption of cognitive resources per unit of exposure was associated with greater sustained attention. Therefore, these findings provide support for a resource depletion account of vigilance decrement, which posits that performance decrements result from the consumption of cognitive resources over time during sustained attention tasks.

The outcomes of Studies 2 and 5 cannot be easily explained by the underload account of vigilance decrement. Smaller increases in cerebral oxygenation from baseline indicated that participants with higher cue utilisation experienced lower cognitive stimulation during the rail control tasks in Study 2 and the driving task in Study 5. Consequently, if understimulation results in a decline in performance, as predicted by the underload account, participants with higher cue utilisation should demonstrate greater response latencies and/or missed signals during sustained attention tasks, compared to participants with lower cue utilisation. However, despite lower cognitive stimulation, higher cue utilisation was associated with greater sustained attention in these studies.

The findings from Study 5, similarly, cannot be easily explained by the Malleable Attentional Resources Theory (MART), which posits that greater perceived effort is associated with greater availability of cognitive resources (Young & Stanton, 2002). Greater increases in cerebral oxygenation from baseline indicated that participants with lower cue utilisation exerted greater effort during the simulated driving task, compared to participants with higher cue utilisation. If greater effort is associated with the mobilisation of additional resources to meet increased task demands, these participants should have had greater attentional resources available for the detection of critical signals (Young & Stanton, 2002). However, participants with lower cue utilisation recorded a greater frequency of missed signals during the simulated driving task, compared to participants with higher cue utilisation. Consequently, the findings from Study 5 do not support the MART.

The fourth and final contribution made by the present programme of research is evidence to indicate that shorter-duration sustained visual search tasks represent a valid alternative to longer-duration vigilance tasks. The results from Study 3 demonstrated that performance during the short-duration sustained visual search task predicted performance during longer-duration monitoring and process control tasks. This suggests that similar cognitive processes are engaged during demanding visual search tasks and less demanding sustained attention tasks. Further, the outcomes of Study 3 provide evidence to indicate that strategies which reduce the consumption of cognitive resources during long-duration tasks may also be effective in reducing cognitive load during more demanding short-duration tasks. The results of Study 4 provide further evidence to suggest that short-duration sustained attention tasks mimic the behaviour typically observed during long-duration vigilance tasks. Together, the findings from Studies 3 and 4 suggest that short-duration sustained visual search tasks are a valid alternative to longer-duration vigilance tasks for detecting individual differences in sustained attention in process control environments.

In combination, the present programme of research provides empirical evidence to support the proposition that cue utilisation is associated with lower cognitive load during process control tasks, which in turn, is associated with a greater capacity for sustained attention. Further, the present research identifies differences in visual search patterns as a potential mediating factor for the relationship between cue utilisation and cognitive resource consumption during sustained attention tasks. These relationships are depicted visually in Figure 3. This theoretical mediation model is yet to be empirically explored, and therefore, represents a potential avenue for future research.

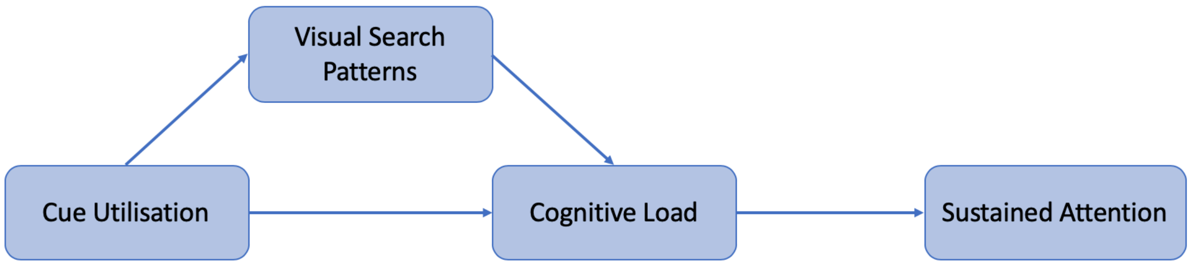

Figure 3. A theoretical model: Cue utilisation increases sustained attention through efficient visual search and decreased cognitive load.

Implications for Applied Environments

The present programme of research extends previous research, indicating that, in addition to improving performance, cue utilisation is associated with lower cognitive load during sustained attention tasks. Consequently, system operators with higher cue utilisation should consume fewer cognitive resources per unit time, resulting in the availability of greater residual cognitive resources, compared to operators with lower cue utilisation. As the availability of residual cognitive resources aids performance, learning, and the management of secondary tasks (Wickens, 2002), these results have implications for the selection, training and management of operators in high-risk industrial environments.

The first implication for applied environments is that assessments of cue utilisation could be used to select operators for roles in high-risk industrial environments. Operators who are able to develop cue-based associations, and utilise these associations in their operational environments, should be able to respond quickly and adaptively to meet the needs of critical situations (Klein, 2008). Further, by consuming fewer cognitive resources during the completion of their primary tasks, operators with higher cue utilisation should retain greater residual cognitive resources, allowing them to better manage the demands of secondary tasks (Wickens, 2002). Consequently, the capacity to identify qualified operators with higher cue utilisation may assist in the selection of job applicants who are better able to sustain attention and maintain performance during demanding situations.

The second implication for applied environments relates to the training of system operators. The results from Studies 1 to 6 consistently demonstrated that cue utilisation typologies within a specific domain were not a function of years of experience in that domain. Consequently, regardless of operators' years of experience, proactive approaches may be beneficial for increasing cue utilisation amongst experienced operators. For instance, qualified operators assessed with relatively lower cue utilisation could be targeted with cue-based training interventions, whereby they are given the opportunity to acquire cues that can be generalised to the broader operational environment (Ivancic & Hesketh, 2000; Klayman, 1988; Scherer et al., 2008; Wiggins, 2015; Wiggins & O'Hare, 2003).

The final applied implication relates to the management of system operators. The relationship between cue utilisation and cognitive load demonstrated throughout the present programme of research suggests that assessments of cue utilisation could provide managers with insights regarding operators' cognitive load during operational tasks, and the associated rate at which cognitive resources are consumed. This insight could assist with the allocation of operators to operational tasks. For instance, operators with higher cue utilisation could be allocated to tasks requiring extended periods of sustained attention, or roles requiring the management of additional secondary tasks. Alternatively, the ability to predict the rate at which operators consume cognitive resources could potentially be used to improve job performance by optimising the length of time between breaks for individual operators.

Limitations and Future Directions

This programme of research provides empirical evidence supporting the proposition that differences in cue utilisation are associated with differences in the rate at which cognitive resources are consumed across a range of operational settings, under a range of operational conditions, and with operators of varying levels of operator experience. However, it is important to note that there are a number of inevitable limitations associated with the outcomes. These limitations and future research recommendations have been summarised under the following two themes: (a) Causal relationships; and (b) Performance in less reliable environments.

Causal relationships

As cue utilisation typologies are quasi-experimental in nature, the present programme of research relied primarily on cross-sectional designs. Therefore, the extent to which causal relationships can be established is limited. Consequently, it is unclear whether a third variable, not controlled for in the present research, better explains the relationships between cue utilisation, cognitive load and sustained attention. For instance, in Study 6, cue utilisation may have been associated with assignments to tasks with greater or lesser demands, or the frequency with which operators opted to take breaks. As these factors likely influence the consumption of cognitive resources, the relationship between cerebral oxygenation and cue utilisation in Study 6 could potentially be explained by environmental features. Similarly, in Studies 2 and

5, the direction of the relationship between cue utilisation and cognitive load has not been empirically established. For example, lower cognitive load could potentially enable greater cue utilisation, rather than cue utilisation leading to a reduction in cognitive load.

To further investigate the nature of the relationship between cue utilisation and cognitive load, longitudinal studies could track cue utilisation and cognitive resource consumption over time and with greater levels of operator experience. A longitudinal approach would help to make clear how the development of cues impacts sustained attention over time, and may be beneficial in establishing the causal relationship between cue utilisation and cognitive load. For example, a longitudinal design may involve assessing power control room operators' cue utilisation and operational performance throughout their careers, to understand the impact of cue utilisation on performance during both early and progressed stages of operators' skill acquisition.

A greater understanding of the causal relationship between cue utilisation and sustained attention is also required to guide the implementation of cue-based training interventions. Multiple-cue judgement approaches have demonstrated that, in situations where the criterion is known, feedback relating to cue validities promotes learning and improves judgement (Balzer et al., 1989; Doherty & Balzer, 1988; Gattie & Bisantz, 2006; Lagnado, Newell, Kahan, & Shanks, 2006; Plessner et al., 2009). These findings reveal a causal relationship between cue-based interventions and performance. However, there is currently no empirical evidence demonstrating that cue acquisition *causes* a decrease in cognitive load during sustained attention tasks.

To investigate whether cue-based interventions cause a decrease in cognitive load, there is a need to conduct experimental studies, in which cue acquisition is manipulated and measures of cognitive load are subsequently recorded during sustained attention tasks. For instance, in situations where the criterion is known, feedback relating to cue validities could be used to manipulate cue utilisation during sustained attention tasks. Alternatively, cue-based interventions, such as cue discovery, whereby the ad-hoc acquisition of cues is enabled through trial and error with simulated operating systems (Klayman, 1988; Wiggins, 2015), could be implemented in situations where the criterion is not known. The outcomes of such experimental designs could help establish the causal link between cue utilisation and cognitive load, which would aid in the development of strategies to improve the performance of operators in high risk industrial environments.

Performance in less reliable environments

Throughout this programme of research, participants with higher cue utilisation demonstrated greater sustained attention, which was assessed in the context of rail control, power control and driving. The rail control simulations used in Studies 1-3 were designed to include implicit patterns in the form of repetitious patterns of train movements. Feature-event relationships in these studies were very reliable, such that specific features were always associated with the same event. Consequently, participants who rapidly acquired task-related patterns and formed associational cues, were able to predict critical events, and thereby record shorter mean response latencies over time. However, outside of the laboratory setting, feature-event relationships are likely to be less reliable. This can result in a salient feature activating an inappropriate association in memory, thereby delaying the accurate recognition of an event (Kahneman & Klein, Gary, 2009; Rowe, Horswill, Kronvall-Parkinson, Poulter, & McKenna, 2009; Wiggins & Loveday, 2015). For instance, Kahneman and Klein (2009) have argued that an overreliance on perceived regularities in the environment can lead to poor judgements and decisions. The activation of an inappropriate association is referred to as miscueing (Brouwers, Wiggins, & Griffin, 2018; Rowe et al., 2009).

During a simulated rail control task, Brouwers et al. (2018) observed that a change in the pattern of feature-event relationships resulted in a greater increase in response latency for participants with higher cue utilisation, compared to those participants with lower cue utilisation. This suggests that participants who are more likely to utilise cues are also more likely to be miscued. Consequently, while Studies 1 and 2 demonstrated an association between cue utilisation and greater performance during tasks containing reliable feature-event relationships, it is unclear whether the advantages afforded by cue utilisation would outweigh the potential negative outcomes associated with miscueing in environments where feature-event relationships are less reliable.

Studies 5 and 6 were designed to examine cue utilisation in more naturalistic environments, which did not contain highly reliable, artificial patterns such as those incorporated within Studies 1 and 2. However, the simulated driving task used in Study 5 was designed as a moderate workload task, and did not contain any intentional or apparent miscues. Further, as operational errors are observed infrequently in power distribution control rooms, there was insufficient performance data available to assess whether operators' performance was negatively influenced by miscueing in Study 6.

While the current programme of research demonstrated that cue utilisation is associated with lower cognitive load during sustained attention tasks, it remains unclear whether cue utilisation is associated with greater sustained attention in operational environments where feature-event relationships are less reliable. Therefore, another area of focus for future research involves the investigation of a range of performance measures in naturalistic operational settings, and the relationship between these performance measures and cue utilisation. The outcomes of such research could help to determine whether cue utilisation is associated with greater sustained attention across a range of operational scenarios.

Conclusions

An understanding of the cognitive processes that underlie sustained attention in high risk operational environments provides a valuable guide for the selection, training and management of system controllers. Differences in sustained attention based on operators' cue utilisation have been evident in a range of operational domains. As the activation of cues in long-term memory has the advantage of reducing the demands on working memory resources, it has been posited that cue utilisation reduces the rate at which cognitive resources are consumed, enabling greater sustained attention. The overall aim of the present programme of research was to examine whether differences in cue utilisation are associated with differences in the allocation of cognitive resources, and the rate at which cognitive resources are consumed, across a range of operational settings, and under a range of operational conditions, using operators of varying levels of operator experience.

Studies 1 and 2 were conducted to establish whether a general capacity for cue utilisation was associated with the consumption of cognitive resources during sustained attention tasks. In Study 1, cue utilisation was associated with smaller increases in mean response latency during a novel rail control simulation. Study 2 replicated and extended these results, demonstrating greater decreases in fixation rates, smaller changes in cerebral oxygenation in the prefrontal cortex, and smaller increases in mean response latency for participants with higher cue utilisation, compared to participants with lower cue utilisation. These results provide support for the proposition that cue utilisation is associated with the consumption of fewer cognitive resources during sustained attention tasks.

Study 3 was designed to validate a newly adapted power control sustained visual search task for process control environments. The outcomes indicated that the sustained visual search task is a valid alternative to a longer-duration process control task for experimental studies. Using the sustained visual search task developed during Study 3, Study 4 examined whether experienced operators' cue utilisation differentiates performance during domain-relevant sustained attention tasks. In two experiments, power distribution operators with higher cue utilisation demonstrated shorter mean response latencies compared to operators with lower cue utilisation. These results provide support for the proposition that experienced operators with higher cue utilisation adopt strategies during operational tasks that reduce their cognitive load, enabling greater sustained attention.

To establish whether differences in cognitive load based on cue utilisation are also evident in more dynamic operational environments, Study 5 examined qualified drivers' cognitive load during a simulated driving task. The results indicated that higher cue utilisation was associated with smaller mean visual saccade amplitudes, smaller mean fixation dispersions, smaller increases in cerebral oxygenation, and fewer missed traffic signals during the simulated driving task. Extending these findings, Study 6 assessed physiological measures of cognitive resource consumption during periods of power distribution operators' regular workdays. Across two testing sessions, higher cue utilisation was associated with smaller increases in cerebral oxygenation in the prefrontal cortex, which is indicative of lower cognitive load. The findings from Studies 5 and 6 are consistent with the broader proposition that experienced operators with higher cue utilisation adopt more efficient search patterns, and experience lower cognitive load, during regular operational tasks, compared to operators with lower cue utilisation.

The present programme of research resulted in four key theoretical contributions which advance an understanding of the processes underpinning the relationship between cue utilisation and sustained attention. These include: (1) support for the proposition that cue utilisation is associated with lower cognitive load, and the consumption of fewer cognitive resources, during sustained attention tasks; (2) the identification of search pattern efficiency as a potential mediator for the relationship between cue utilisation and lower cognitive load; (3) evidence supporting a resource depletion account of the vigilance decrement; and (4) evidence that shorter-duration sustained visual search tasks are a valid alternative to longer-duration vigilance tasks.

Complete Reference List

Amaldi, P., & Leroux, M. (1995). Selecting relevant information in a complex environment: The case of air traffic control. In L. Norros (Ed.), *5th European Conference on Cognitive Science Approaches in Process Control* (Vol. 158, pp. 89–98). doi:10.1097b1818

Anders, N. S., Seijmonsbergen, A. C., & Bouten, W. (2011). Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. *Remote Sensing of Environment*, *115*, 2976–2985. doi:10.1016/j.rse.2011.05.007

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406. doi:10.1037/0033-295X.89.4.369

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem situations. *Psychological Review*, *94*, 192–210. doi:10.1037/0033-295X.94.2.192

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Lawrence Erlbaum.

Baldwin, C. L., & Lewis, B. A. (2017). Positive valence music restores executive control over sustained attention. *PloS One*, *12*, e0186231. doi:10.1371/journal.pone.0186231

Balfe, N., Sharples, S., & Wilson, J. R. (2015). Impact of automation: Measurement of performance, workload and behaviour in a complex control environment. *Applied Ergonomics*, *47*, 52–64. doi:10.1016/j.apergo.2014.08.002

Balfe, N., Wilson, J. R., Sharples, S., & Clarke, T. (2012). Development of design principles for automated systems in transport control. *Ergonomics*, *55*, 37–54. doi:10.1080/00140139.2011.636456

Balzer, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin*, *106*, 410–433. doi:10.1037/0033-2909.106.3.410

Berg, D. J., Boehnke, S. E., Marino, R. A., Munoz, D. P., & Itti, L. (2009). Free viewing of dynamic stimuli by humans and monkeys. *Journal of Vision*, *9*, 19–19. doi:10.1167/9.5.19

Bertram, R., Helle, L., Kaakinen, J. K., & Svedström, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PloS One*, *8*, e66169. doi:10.1371/journal.pone.0066169

Bisantz, A. M., & Pritchett, A. R. (2003). Measuring the fit between human judgments and automated alerting algorithms: A study of collision detection. *Human Factors*, *45*, 266–280. doi:10.1518/hfes.45.2.266.27243

Brehmer, A., & Brehmer, B. (1988). What has been learned about human judgment from thirty years of policy capturing? In B. Brehmer & C. R. B. Joyce (Eds.), *Human judgment: The SJT view*. Amsterdam: North-Holland: Elsevier. doi:10.1016/s0166-4115(08)62171-8

Brehmer, B. (1994). The psychology of linear judgement models. *Acta Psychologica*, *87*, 137–154. doi:10.1016/0001-6918(94)90048-5

Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British Journal of Clinical Psychology*, *21*, 1–16. doi:10.1111/j.2044-8260.1982.tb01421.x

Brouwers, S., Wiggins, M., & Griffin, B. (2018). Operators who readily acquire patterns and cues, risk being miscued in routinized settings. *Journal of Experimental Psychology: Applied*, *24*, 261–274. doi:10.1037/xap0000151

Brouwers, S., Wiggins, M. W., Griffin, B., Helton, W. S., & O’Hare, D. (2017). The role of cue utilisation in reducing the workload in a train control task. *Ergonomics*, *60*, 1–16. doi:10.1080/00140139.2017.1330494

Brouwers, S., Wiggins, M. W., Helton, W., O’Hare, D., & Griffin, B. (2016). Cue utilization and cognitive load in novel task performance. *Frontiers in Psychology*, *7*, 1–12. doi:10.3389/fpsyg.2016.00435

Brown, T. L., & Carr, T. H. (1989). Automaticity in skill acquisition: Mechanisms for reducing interference in concurrent performance. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 686–700. doi:10.1037/0096-1523.15.4.686

Brunswik, E. (1955). Representative Design and Probabilistic Theory in a Functional Psychology. *Psychological Review*, *62*, 193–217. doi:10.1037/h0047470

Caggiano, D. M., & Parasuraman, R. (2004). The role of memory representation in the vigilance decrement, *11*, 932–937. doi:10.3758/bf03196724

Camilli, M., Terenzi, M., & Di Nocera, F. (2007). Concurrent validity of an ocular measure of mental workload. In *Human factors issues in complex system performance. In D. de Waard, G.R.J. Hockey, P. Nickel, and K.A. Brookhuis (Eds.)* (pp. 117–129). Shaker Publishing Maastricht, the Netherlands.

Campitelli, G., & Gobet, F. (2011). Deliberate practice: Necessary but not sufficient. *Current Directions in Psychological Science*, *20*, 280–285. doi:10.1177/0963721411421922

Carvalho, P. V. R., dos Santos, I. L., Gomes, J. O., Borges, M. R. S., & Guerlain, S. (2008). Human factors approach for evaluation and redesign of human-system interfaces of a nuclear power plant simulator. *Displays*, *29*, 273–284. doi:10.1016/j.displa.2007.08.010

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, *27*, 951–964. doi:10.1068/p270951

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 7–75). Hillsdale, NJ: Erlbaum.

Chung, P. H., & Byrne, M. D. (2008). Cue effectiveness in mitigating postcompletion errors in a routine procedural task. *International Journal of Human-Computer Studies*, *66*, 217–232. doi:10.1016/j.ijhcs.2007.09.001

Coderre, S., Mandin, H., Harasym, P. H., & Fick, G. H. (2003). Diagnostic reasoning strategies and diagnostic success. *Medical Education*, *37*, 695–703. doi:10.1046/j.1365-2923.2003.01577.x

Colquhoun, W. P. (1967). Sonar target detection as a decision process. *Journal of Applied Psychology*, *51*, 187. doi:10.1037/h0024343

Cooksey, R. W. (1988). Social judgment theory in education: Current and potential applications. In *Advances in Psychology* (Vol. 54, pp. 273–315). Elsevier. doi:10.1016/s0166-4115(08)62177-9

Cooksey, R. W. (1996). The methodology of social judgement theory. *Thinking & Reasoning*, *2*, 141–174. doi:10.1080/135467896394483

Crane, M. F., Brouwers, S., Wiggins, M. W., Loveday, T., Forrest, K., Giselle, S., … Cyna, A. M. (2018). " Experience Isn ' t Everything ": How Emotion Affects the Relationship Between Experience and Cue Utilization. *Human Factors*, *60*, 685–698. doi:10.1177/0018720818765800

Croskerry, P. (2009). A universal model of diagnostic reasoning. *Academic Medicine*, *84*, 1022–1028. doi:10.1097/ACM.0b013e3181ace703

Crundall, D., Chapman, P., Phelps, N., & Underwood, G. (2003). Eye Movements and Hazard Perception in Police Pursuit and Emergency Eye Movements and Hazard Perception in Police Pursuit and Emergency Response Driving. doi:10.1037/1076-898X.9.3.163

Cuny, X. (1979). Different levels of analysing process control tasks. *Ergonomics*, *22*, 415–425. doi:10.1080/00140137908924626

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, *38*, 105–134. doi:10.1007/s11251-009-9110-0

Dember, W. N., Galinsky, T. L., & Warm, J. S. (1992). The role of choice in vigilance performance. *Bulletin of the Psychonomic Society*, *30*, 201–204. doi:10.3758/BF03330441

Desmond, P. A., & Hoyes, T. W. (1996). Workload variation, intrinsic risk and utility in a simulated air traffic control task: evidence for compensatory effects. *Safety Science*, *22*, 87–101. doi:10.1016/0925-7535(96)00008-2

Dhami, M. K., & Harries, C. (2001). Fast and frugal versus regression models of human judgement. *Thinking & Reasoning*, *7*, 5–27. doi:10.1080/13546780042000019

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A Random Glance at the Flight Deck: Pilots' Scanning Strategies and the Real-Time Assessment of Mental Workload. *Journal of Cognitive Engineering and Decision Making*, *1*, 271–285. doi:10.1518/155534307X255627.

Dimeas, A. L., & Hatziargyriou, N. D. (2005). Operation of a multiagent system for microgrid control. *IEEE Transactions on Power Systems*, *20*, 1447–1455. doi:10.1109/TPWRS.2005.852060

Doherty, M. E., & Balzer, W. K. (1988). Cognitive feedback. *Advances in Psychology*, *54*, 163–197. doi:10.1016/s0166-4115(08)62173-1

Doran, S. M., Van Dongen, H. P. A., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: evidence of state instability. *Archives Italiennes* *de Biologie*, *139*, 253–267.

Dreyfus, H. L., & Dreyfus, S. E. (1986). Why skills cannot be represented by rules. *Advances in Cognitive Science*, *1*, 315–335.

Dunn, N., & Williamson, A. (2011). Monotony in the rail industry: The role of task demand in mitigating monotony-related effects on performance. *Ergonomics Australia–HFESA*, 1–5. doi:10.1201/b12742-87

Durantin, G., Dehais, F., & Delorme, A. (2015). Characterization of mind wandering using fNIRS. *Frontiers in Systems Neuroscience*, *9*, 1–7. doi:10.3389/fnsys.2015.00045

Edkins, G. D., & Pollock, C. M. (1997). The influence of sustained attention on railway accidents. *Accident Analysis & Prevention*, *29*, 533–539. doi:10.1016/s0001-4575(97)00033-x

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgement and choice. *Annual Review of Psychology*, *32*, 53–88. doi:10.1146/annurev.ps.32.020181.000413

Ekkekakis, P. (2009). Illuminating the Black Box: Investigating Prefrontal Cortical Hemodynamics during Exercise with Near-Infrared Spectroscopy. *Journal of Sport and Exercise Psychology*, *31*, 505–553. doi:10.1123/jsep.31.4.505

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*, 462–492. doi:10.1080/001401399185595

Endsley, M. R., & Kiris, E. O. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors and Ergonomics Society*, *37*, 381–394. doi:10.1518/001872095779064555

Epstein, S. (1980). The stability of behavior: II. Implications for psychological research.

American Psychologist, *35*, 790–806. doi:10.1037//0003-066x.35.9.790

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 211–245. doi:10.1037//0033-295x.102.2.211

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305. doi:10.1146/annurev.psych.47.1.273

Ericsson, K. A., & Towne, T. J. (2010). Expertise. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*, 404–416. doi:10.1002/wcs.47

Evans, D. C., & Fendley, M. (2017). A multi-measure approach for connecting cognitive workload and automation. *Journal of Human Computer Studies*, *97*, 182–189. doi:10.1016/j.ijhcs.2016.05.008

Evans, J. S. B. T., Clibbens, J., Cattani, A., Harris, A., & Dennis, I. (2003). Explicit and implicit processes in multicue judgment. *Memory & Cognition*, *31*, 608–618. doi:10.3758/bf03196101

Falkland, E. C., & Wiggins, M. W. (2018). Cross-task cue utilisation and situational awareness in simulated air traffic control. *Applied Ergonomics*, *74*, 24–30. doi:10.1016/j.apergo.2018.07.015

Fallgatter, A. J., & Strik, W. K. (1997). Right frontal activation during the continuous performance test assessed with near-infrared spectroscopy in healthy subjects. *Neuroscience Letters*, *223*, 89–92. doi:10.1016/s0304-3940(97)13416-4

Fallgatter, A. J., & Strik, W. K. (1998). Frontal brain activation during the Wisconsin Card Sorting Test assessed with two-channel near-infrared spectroscopy. *European Archives of Psychiatry and Clinical Neuroscience*, *248*, 245–249. doi:10.1007/s004060050045

Farrington-Darby, T., & Wilson, J. R. (2006). The nature of expertise: A review. *Applied*

Ergonomics, *37*, 17–32. doi:10.1016/j.apergo.2005.09.001

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology*, *69*, 153–166. doi:10.1037//0022-3514.69.1.153

Fishburn, P. C. (1982). *The foundation of expected utility theory. Theory & Decision Library.* Dordrecht, Netherlands: Reidel.

Fu, R., Guo, Y., Chen, Y., Yuan, W., Ma, Y., Peng, J., … Wang, C. (2011). Research on heart rate and eye movement as indicators of drivers' mental workload. In *Road Safety and Simulation* (pp. 1–20).

Gattie, G. J., & Bisantz, A. M. (2006). The effects of integrated cognitive feedback components and task conditions on training in a dental diagnosis task. *International Journal of Industrial Ergonomics*, *36*, 485–497. doi:10.1016/j.ergon.2006.01.011

Ghylin, K. M., Drury, C. G., Batta, R., & Lin, L. (2007). Temporal effects in a security inspection task: Breakdown of performance components. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 51, pp. 93–97). Sage Publications Sage CA: Los Angeles, CA.

Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition*, *4*, 1–21. doi:10.1006/ccog.1995.1001

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482. doi:10.1146/annurev-psych-120709-145346

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the Fast and Frugal Way : Models of Bounded Rationality. *Psychological Review*, *103*, 650–669. doi:10.1037/0033-295x.103.4.650

Gigerenzer, G., Todd, P. M., & Group, A. R. (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory*, *9*, 408–418. doi:10.1101/lm.45202

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, *24*, 631–645. doi:10.1016/S0169-8141(98)00068-7

Goldstein, D. G., & Gigerenzer, G. (2002). Models of Ecological Rationality : The Recognition Heuristic. *Psychological Review*, *109*, 75–90. doi:10.1037//0033-295X.109.1.75

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In KR Boff, L. Kaufman, & JP Thomas (Eds.), Handbook of perception and human performance, Vol. 2. Cognitive processes and performance (pp. 1–49). John Wiley & Sons.

Gratton, G., & Fabiani, M. (2006). Optical imaging of brain function in R. Parasuraman, M. Rizzo (Eds.), Neuroergonomics: The Brain At Work, Oxford University. In *Oxford Series in Human-Technology Interactions* (pp. 65–81). Cambridge, MA.

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

Grubb, P. L., Miller, L. C., Nelson, W. T., Warm, J. S., Dember, W. N., & Davies, D. R. (1994). Cognitive failure and perceived workload in vigilance performance. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 115–121). Erlbaum Hillsdale, NJ.

Grüner, M., & Ansorge, U. (2017). Mobile Eye Tracking During Real-World Night Driving: A Selective Review of Findings and Recommendations for Future Research. *Journal of Eye Movement Research*, *1011*, 1–18. doi:10.16910/jemr.10.2.1

Gunzelmann, G., Gross, J. B., Gluck, K. A., & Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: Integrating mathematical and cognitive modeling. *Cognitive Science*, *33*, 880–910. doi:10.1111/j.1551-6709.2009.01032.x

Gupta, R. A., & Chow, M.-Y. (2010). Networked Control System: Overview and Research Trends. *IEEE Transactions on Industrial Electronics*, *57*, 2527–2535. doi:10.1109/TIE.2009.2035462

Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, *45*, 34–45. doi:10.1016/j.intell.2013.04.001

Hammond, K. R., & Stewart, T. R. (2001). *The essential Brunswik: Beginnings, explications, applications.* New York: Oxford University Press.

Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social-Judgment Theory. In M. Kaplan & S. Schwartz (Eds.), *Human Judgment & Decision Processes* (pp. 271–312). New York: Academic Press Inc.

Hanauer, D. A., Englesbe, M. J., Cowan, J. A., & Campbell, D. A. (2009). Informatics and the American College of Surgeons National Surgical Quality Improvement Program: automated processes could replace manual record review. *Journal of the American College of Surgeons*, *208*, 37–41. doi:10.1016/j.jamcollsurg.2008.08.030

Hancock, P. A., Volante, W. G., & Szalma, J. L. (2016). Defeating the vigilance decrement. *IIE Transactions on Occupational Ergonomics and Human Factors*, *4*, 151–163. doi:10.1080/21577323.2016.1178191

Hancock, P. A., & Warm, J. S. (2003). A Dynamic Model of Stress and Sustained Attention. *Journal of Human Performance in Extreme Environments*, *7*, 519–537. doi:10.7771/2327-2937.1024

Harries, P. A., & Harries, C. (2001). Studying clinical reasoning, part 1: Have we been taking the wrong 'track'? *British Journal of Occupational Therapy*, *64*, 164–168. doi:10.1177/030802260106400402

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (Vol. 52, pp. 139–183). Amsterdam: North-Holland: Elsevier.

Heald, J. E. (1991). Social judgment theory: Applications to educational decision making. *Educational Administration Quarterly*, *27*, 343–357. doi:10.1177/0013161x91027003006

Helton, W. S., Hollander, T. D., Tripp, L. D., Parsons, K., Warm, J. S., Matthews, G., & Dember, W. N. (2007). Cerebral hemodynamics and vigilance performance. *Journal of Clinical and Experimental Neuropsychology*, *29*, 545–552. doi:10.1080/13803390600814757

Helton, W. S., Hollander, T. D., Warm, J. S., Matthews, G., Dember, W. N., Wallaart, M., … Hancock, P. A. (2005). Signal regularity and the mindlessness model of vigilance. *British Journal of Psychology*, *96*, 249–261. doi:10.1348/000712605x38369

Helton, W. S., Hollander, T. D., Warm, J. S., Tripp, L. D., Parsons, K., Matthews, G., … Hancock, P. a. (2007). The abbreviated vigilance task and cerebral hemodynamics. *Journal of Clinical and Experimental Neuropsychology*, *29*, 545–552. doi:10.1080/13803390600814757

Helton, W. S., Kern, R. P., & Walker, D. R. (2009). Conscious thought and the sustained attention to response task. *Consciousness and Cognition*, *18*, 600–607.

doi:10.1016/j.concog.2009.06.002

Helton, W. S., & Russell, P. N. (2011). Working memory load and the vigilance decrement. *Experimental Brain Research*, *212*, 429–437. doi:10.1007/s00221-011-2749-1

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

Helton, W. S., & Russell, P. N. (2013). Visuospatial and verbal working memory load: Effects on visuospatial vigilance. *Experimental Brain Research*, *224*, 429–436. doi:10.1007/s00221-012-3322-2

Helton, W. S., & Russell, P. N. (2015). Rest is best: The role of rest and task interruptions on vigilance. *Cognition*, *134*, 165–173. doi:10.1016/j.cognition.2014.10.001

Helton, W. S., & Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. *Acta Psychologica*, *129*, 18–25. doi:10.1016/j.actpsy.2008.04.002

Helton, W. S., Warm, J. S., Tripp, L. D., Matthews, G., Parasuraman, R., & Hancock, P. A. (2010). Cerebral lateralization of vigilance: A function of task difficulty. *Neuropsychologia*, *48*, 1683–1688. doi:10.1016/j.neuropsychologia.2010.02.014

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*, 498–504. doi:10.1016/j.tics.2003.09.006

Hitchcock, E. M., Warm, J. S., Matthews, G., Dember, W. N., Shear, P. K., Tripp, L. D., … Parasuraman, R. (2003). Automation cueing modulates cerebral blood flow and vigilance in a simulated air traffic control task. *Theoretical Issues in Ergonomics Science*, *4*, 89–112. doi:10.1080/14639220210159726

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological*

Psychology, *45*, 73–93. doi:10.1016/s0301-0511(96)05223-4

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics*, *57*, 787–795. doi:10.3758/bf03206794

Hopkin, V. D. (2017). *Human factors in air traffic control*. CRC Press.

Horowitz, F. D. (1969). Learning, developmental research, and individual differences. *Advances in Child Development and Behavior*, *4*, 83–126. doi:10.1016/s0065-2407(08)60427-9

Howard, G., Chen, J., & Wiggins, M. W. (2003). Simulated Rail Control. Sydney: Unpublished.

Ikehara, C. S., & Crosby, M. E. (2005). Assessing Cognitive Load with Physiological Sensors. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, *00*, 295a–295a. doi:10.1109/HICSS.2005.103

Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence. *Ergonomics*, *43*, 1966–1984. doi:10.1080/00140130050201427

Jere, S., Dauwels, J., Asif, M. T., Vie, N. M., Cichocki, A., & Jaillet, P. (2014). Extracting commuting patterns in railway networks through matrix decompositions. In *13th International Conference on Control Automation Robotics and Vision, ICARCV 2014* (pp. 541–546). doi:10.1109/ICARCV.2014.7064362

Kaber, D. B., & Endsley, M. R. (2004). *The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task*. *Theor. Issues in Ergon. Sci*. doi:10.1080/1463922021000054335

Kaempf, G. L., Klein, Gary, A., Thordsen, M. L., & Wolf, S. (1996). Decision making in complex naval command-and-control environments. *Human Factors*, *38*, 220–231.

doi:10.1518/001872096779047986

Kaempf, G. L., & Klein, G. A. (1994). Aeronautical Decision Making: The next generation. In N. Johnston, N. McDonald, & R. Fuller (Eds.), *Aviation Psychology in Practice* (pp. 223–254). Hants, England: Avebury Technical.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ.

Kahneman, D., & Klein, Gary, A. (2009). Conditions for intuitive expertise: a failure to disagree. *American Psychologist*, *64*, 515–526. doi:10.1518/001872096779047986

Kaplan, R. M., & Saccuzzo, D. P. (2017). *Psychological testing: Principles, applications, and issues*. Nelson Education.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, *15*, 169–182. doi:10.1016/0272-4944(95)90001-2

Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, *134*, 404–426. doi:10.1037/0033-2909.134.3.404

Karlsson, L., Juslin, P., & Olsson, H. (2007). Adaptive changes between cue abstraction and exemplar memory in a multiple-cue judgment task with continuous cues. *Psychonomic Bulletin & Review*, *14*, 648–653. doi:10.3758/bf03193103

Kass, S. J., Vodanovich, S. J., Stanny, C. J., & Taylor, T. M. (2001). Watching the Clock: Boredom and Vigilance Performance. *Perceptual and Motor Skills*, *92*, 969–976. doi:10.1177/003151250109203c01

King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research.* New Jersey: Princeton University Press.

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*, 1–10. doi:10.1016/s0959-4752(01)00014-7

Klayman, J. (1988). Cue discovery in probabilistic environments: Uncertainty and experimentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 317. doi:10.1037//0278-7393.14.2.317

Klein, Gary, A. (1999). Sources of Power. Cambridge, MA: MIT Press.

Klein, G. A. (1989). Recognition-primed decisions (RPD). *Advances in Man-Machine Systems*, *5*, 47–92.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (Ablex, pp. 138–147). Norwood, NJ.

Klein, G. A. (1997). The Recognition-Primed Decision (RPD) model: Looking back, looking forward. In G. A. Klein & C. E. Zsambok (Eds.), *Naturalistic Decision Making* (pp. 285–292). Mahwah, NJ: Lawrence Erlbaum Associates.

Klein, G. A. (1998). *Sources of Power: How people make decisions.* Cambridge, MA: MIT Press.

Klein, G. A. (2003). *Intuition at work.* New York: Currency Books.

Klein, G. A. (2008). Naturalistic Decision Making. *Human Factors*, *50*, 456–460. doi:10.1518/001872008X288385.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 30, pp. 576–580). doi:10.1177/154193128603000616

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. In *Journal of Cognitive Engineering* *and Decision Making* (Vol. 4, pp. 186–209). SAGE Publications Sage CA: Los Angeles, CA. doi:10.1518/155534310x12844000801203

Klein, G. A., & Klinger, D. (1991). Naturalistic Decision Making Human Systems. *IAC Gateway*, *11*, 16–19. doi:10.1518/001872008X288385

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (1993). *Decision making in action: models and methods.* Norwood, CT: Ablex.

Lagnado, D. A., Newell, B. R., Kahan, S., & Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *Journal of Experimental Psychology: General*, *135*, 162–183. doi:10.1037/0096-3445.135.2.162

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, *38*, 963–974. doi:10.2307/2529876

Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin*, *139*, 870–900. doi:10.1037/a0030694

Lansdale, M., Underwood, G., & Davies, C. (2010). Something Overlooked? How experts in change detection use visual saliency. *Applied Cognitive Psychology*, *24*, 213–225. doi:10.1002/acp.1552

Lee, S. J., Kim, J., & Jang, S. C. (2011). Human error mode identification for NPP main control room operations using soft controls. *Journal of Nuclear Science and Technology*, *48*, 902–910. doi:10.1080/18811248.2011.9711776

Lieberman, H. R., Niro, P., Tharion, W. J., Nindl, B. C., Castellani, J. W., & Montain, S. J. (2006). Cognition during sustained operations: comparison of a laboratory simulation to field studies. *Aviation, Space, and Environmental Medicine*, *77*, 929–935.

Lim, J., & Dinges, D. F. (2008). Sleep Deprivation and Vigilant Attention. *Annals of the New*

York Academy of Sciences, *1129*, 305–322. doi:10.1196/annals.1417.002

Lin, C. J., Yenn, T.-C., & Yang, C.-W. (2010). Automation design in advanced control rooms of the modernized nuclear power plants. *Safety Science*, *48*, 63–71. doi:10.1016/j.ssci.2009.05.005

Lipshitz, R. (1993). Converging themes in the study of decision making in realistic settings. In A. Klein, Gary, J. M. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 103–137). Norwood, New Jersey: Ablex Publishing Corporation.

Lipshitz, R., Klein, Gary, A., & Orasanu, J. (2001). Focus Article : Taking Stock of Naturalistic Decision Making, *352*, 331–352. doi:10.1002/bdm.381

Loft, S., Neal, A., & Humphreys, M. S. (2007). The development of a general associative learning account of skill acquisition in a relative arrival-time judgment task. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 938–959. doi:10.1037/0096-1523.33.4.938

Loft, S., Sanderson, P., Neal, A., & Mooij, M. (2007). Modeling and predicting mental workload in en route air traffic control: Critical review and broader implications. *Human Factors*, *49*, 376–399. doi:10.1518/001872007x197017

Loveday, T., Wiggins, M. W., Festa, M., Schell, D., & Twigg, D. (2013). Pattern recognition as an indicator of diagnostic expertise. In P. L. Carmona, J. S. Sanchez, & A. L. N. Fred (Eds.), *Pattern Recognition - Application and Methods* (Vol. 204, pp. 1–11). Berlin: Springer. doi:10.1007/978-3-642-36530-0

Loveday, T., Wiggins, M. W., Harris, J. M., O'Hare, D., & Smith, N. (2013). An objective approach to identifying diagnostic expertise among power system controllers. *Human Factors*, *55*, 90–107. doi:10.1177/0018720812450911

Loveday, T., Wiggins, M. W., & Searle, B. J. (2013). Cue utilization and broad indicators of workplace expertise. *Journal of Cognitive Engineering and Decision Making*, *8*, 98–113. doi:10.1177/1555343413497019

Loveday, T., Wiggins, M. W., Searle, B. J., Festa, M., & Schell, D. (2013). The capability of static and dynamic features to distinguish competent from genuinely expert practitioners in pediatric diagnosis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*, 125–137. doi:10.1177/0018720812448475

Luck, S. J., Hillyard, S. a, Mouloua, M., & Hawkins, H. L. (1996). Mechanisms of visual-spatial attention: resource allocation or uncertainty reduction? *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 725–737. doi:10.1037/0096-1523.22.3.725

Mackworth, N. H. (1948). The breakdown of vigilance durning prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*, 6–21. doi:10.1080/17470214808416738

MacLean, K. A., Aichele, S. R., Bridwell, D. A., Mangun, G. R., Wojciulik, E., & Saron, C. D. (2009). Interactions between endogenous and exogenous attention during vigilance. *Attention, Perception, & Psychophysics*, *71*, 1042–1058. doi:10.3758/app.71.5.1042

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia*, *37*, 661–670. doi:10.1016/S0028-3932(98)00127-4

Mann, D. T. Y., Williams, A. M., Ward, P., & Janelle, C. M. (2007). Perceptual-cognitive expertise in sport: A meta-analysis. *Journal of Sport and Exercise Psychology*, *29*, 457–478. doi:10.1123/jsep.29.4.457

Matthews, G., & Desmond, P. A. (1997). Underload and performance impairment: evidence from studies of stress and simulated driving. In D. Harris (Ed.), *Engineering Psychology* *and Cognitive Ergonomics* (pp. 355–361). Ashgate: Aldershot.

Matthews, G., Warm, J. S., Reinerman-Jones, L. E., Langheim, L. K., Washburn, D. A., & Tripp, L. D. (2010). Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. *Journal of Experimental Psychology: Applied*, *16*, 187–203. doi:10.1037/a0019572

Matthews, G., Warm, J. S., Shaw, T. H., & Finomore, V. S. (2010). A multivariate test battery for predicting vigilance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *54*, 1072–1076. doi:10.1177/154193121005401405

McCalley, J., Asgarpoor, S., Bertling, L., Billinion, R., Chao, H., Chen, J., … Grigg, C. (2004). Probabilistic security assessment for power system operations. In *Power Engineering Society General Meeting, 2004. IEEE* (pp. 212–220). IEEE.

McCammon, I., & Hägeli, P. (2007). An evaluation of rule-based decision tools for travel in avalanche terrain. *Cold Regions Science and Technology*, *47*, 193–206. doi:10.1016/j.coldregions.2006.08.007

McCormack, C., Wiggins, M. W., Loveday, T., & Festa, M. (2014). Expert and competent non-expert visual cues during simulated diagnosis in intensive care. *Frontiers in Psychology*, *5*, 1–7. doi:10.3389/fpsyg.2014.00949

McVay, J. C., & Kane, M. J. (2012). Drifting from slow to “d’oh!”: Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 525. doi:10.1037/a0025896

Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological Science*, *21*, 914–919. doi:10.1177/0956797610373933

Metzger, M., & Polakow, G. (2011). A survey on applications of agent technology in industrial process control. *IEEE Transactions on Industrial Informatics*, *7*, 570–581. doi:10.1109/tii.2011.2166781

Molloy, R., & Parasuraman, R. (1996). Monitoring an automated system for a single failure: Vigilance and task complexity effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *38*, 311–322. doi:10.1177/001872089606380211

Morrison, B. W., & Morrison, N. (2015). Diagnostic Cues in Major Crime Investigation. In M. W. Wiggins & T. Loveday (Eds.), *Diagnostic Expertise in Organizational Environments* (pp. 91–98). Surrey, England: Ashgate Publishing Limited, Surrey, England.

Morrison, B. W., Wiggins, M. W., Bond, N. W., & Tyler, M. D. (2013). Measuring relative cue strength as a means of validating an inventory of expert offender profiling cues. *Journal of Cognitive Engineering and Decision Making*, *7*, 211–226. doi:10.1177/1555343412459192

Mourant, R. R., & Rockwell, T. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors*, *14*, 325–335. doi:10.1177/001872087201400405

Müller, S., Abernethy, B., & Farrow, D. (2006). How do world-class cricket batsmen anticipate a bowler's intention? *The Quarterly Journal of Experimental Psychology*, *59*, 2162–2186. doi:10.1080/02643290600576595

Mumaw, R. J., Roth, E. M., Vicente, K. J., & Burns, C. M. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors*, *42*, 36–55.

Myers, J. L., Well, A., & Lorch, R. F. (2010). *Research design and statistical analysis*. Routledge.

Nachreiner, F., & Nickel, P. (2006). Human factors in process control systems:The design of human–machine interfaces. *Safety Science*, *44*, 5–26. doi:10.1016/j.ssci.2005.09.003

Nachreiner, F., Nickel, P., & Meyer, I. (2006). Human factors in process control systems: The design of human-machine interfaces. *Safety Science*, *44*, 5–26. doi:10.1016/j.ssci.2005.09.003

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*, 214–255. doi:10.1037//0033-295x.86.3.214

Navon, R. (2005). Automated project performance control of construction projects. *Automation in Construction*, *14*, 467–476. doi:10.1016/j.autcon.2004.09.006

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Norman, D. A. (2015). The human side of automation. In *Road Vehicle Automation 2* (pp. 73–79). Springer.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research* (pp. 1–17). New York: Plenum.

Noyes, J., & Bransby, M. (2001). *People in control: human factors in control room design*. London, UK: IET.

Nuechterlein, K. H., Parasuraman, R., & Jiang, Q. (1983). Visual Sustained Attention : Image Degradation Produces Rapid Sensitivity Decrement over Time. *Science*, *220*, 327–329. doi:10.1126/science.6836276

O'Hara, J. M., & Hall, R. E. (1992). Advanced control rooms and crew performance issues: implications for human reliability. *IEEE Transactions on Nuclear Science*, *39*, 919–923. doi:10.1109/23.159733

Ong, M., Russell, P. N., & Helton, W. S. (2013). Frontal cerebral oxygen response as an indicator of initial attention effort during perceptual learning. *Experimental Brain Research*, *229*, 571–578. doi:10.1007/s00221-013-3619-9

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science*, *205*, 924–927. doi:10.1126/science.472714

Parasuraman, R., de Visser, E., Clarke, E., McGarry, W. R., Hussey, E., Shaw, T. H., & Thompson, J. C. (2009). Detecting threat-related intentional actions of others: Effects of image quality, response mode, and target cuing on vigilance. *Journal of Experimental Psychology: Applied*, *15*, 275–290. doi:10.1037/a0017132

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation induced "complacency." *The International Journal of Aviation Psychology*. doi:10.1207/s15327108ijap0301_1

Parasuraman, R., Mouloua, M., Molloy, R., & Hilburn, B. (1996). Monitoring of automated systems. In Mustapha Mouloua (Ed.), *Automation and human performance: Theory and applications* (pp. 91–115). Mahwah, NJ: Lawrence Erlbaum Associates.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*, 230–253. doi:10.1518/001872097778543886

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans : A Publication of the IEEE Systems, Man, and Cybernetics Society*, *30*, 286–297. doi:10.1109/3468.844354

Parasuraman, R., Warm, J. S., & Dember, W. N. (1987). Vigilance: Taxonomy and utility. In *Ergonomics and Human Factors: Recent Research* (pp. 11–32). doi:10.1007/978-1-4612-4756-2_2

Pashler, H., & Wagenmakers, E. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science*, *7*, 528–530. doi:10.1177/1745691612465253

Pauley, K., O'Hare, D., & Wiggins, M. W. (2009). Measuring expertise in weather-related aeronautical risk perception: The validity of the Cochran–Weiss–Shanteau (CWS) Index. *The International Journal of Aviation Psychology*, *19*, 201–216. doi:10.1080/10508410902979993

Perry, N. C., Wiggins, M. W., Childs, M., & Fogarty, G. (2012). Can reduced processing decision support interfaces improve the decision-making of less-experienced incident commanders? *Decision Support Systems*, *52*, 497–504. doi:10.1016/j.dss.2011.10.010

Perry, N. C., Wiggins, M. W., Childs, M., & Fogarty, G. (2013). The Application of Reduced-Processing Decision Support Systems to Facilitate the Acquisition of Decision-Making Skills. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*, 535–544. doi:10.1177/0018720812467367

Pigeau, R. A., Angus, R. G., O'Neill, P., & Mack, I. (1995). Vigilance latencies to aircraft detection among NORAD surveillance operators. *Human Factors*, *37*, 622–634. doi:10.1518/001872095779049291

Plessner, H., Schweizer, G., Brand, R., & O'Hare, D. (2009). A multiple-cue learning approach as the basis for understanding and improving soccer referees' decision making. *Progress in Brain Research*, *174*, 151–158. doi:10.1016/s0079-6123(09)01313-2

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903. doi:10.1037/0021-9010.88.5.879

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui

(Ed.), *Encyclopedia of Human Computer Interaction* (pp. 211–219). Hershey, PA: IGI Publishing. doi:10.4018/978-1-59140-562-7.ch034

Posner, M. I. (1978). *Chronometric explorations of mind.* Lawrence Erlbaum.

Power, D. J., & Sharda, R. (2009). Decision support systems. In *Springer handbook of automation* (pp. 1539–1548). Springer.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Section 4: The Advanced Progressive Matrices. Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Reason, J. (1990). *Human error*. Cambridge, UK: Cambridge university press.

Reason, J. (2000). Human error: Models and management. *BMJ*, *320*, 768–770. doi:10.1136/bmj.320.7237.768

Reinerman-Jones, L. E., Matthews, G., Langheim, L. K., & Warm, J. S. (2011). Selection for vigilance assignments: A review and proposed new direction. *Theoretical Issues in Ergonomics Science*, *12*, 273–296. doi:10.1080/14639221003622620

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*, 747–758. doi:10.1016/s0028-3932(97)00015-8

Robinson, D. A. (1965). The mechanics of human smooth pursuit eye movement. *The Journal of Physiology*, *180*, 569–591. doi:10.1113/jphysiol.1965.sp007718

Rodger, D. (1983). The psychology of vigilance. *Journal of Occupational Psychology*, *56*, 265–266.

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology*, *20*, 343–350. doi:10.1037/h0043220

Rowe, R., Horswill, M. S., Kronvall-Parkinson, M., Poulter, D. R., & McKenna, F. P. (2009). The effect of disguise on novice and expert tennis players' anticipation ability. *Journal of Applied Sport Psychology*, *21*, 178–185. doi:10.1080/10413200902785811

Salthouse, T. A. (1991). Expertise as the circumvention of human processing limitations. In *Toward a general theory of expertise: Prospects and limits* (pp. 286–300). Cambridge, NY: Cambridge University Press.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Eye Tracking Research & Applications (ETRA) Symposium* (pp. 71–78). Palm Beach Gardens, FL: ACM.

Scerbo, M. W. (1998). What's so boring about vigilance? In R. R. Hoffman, M. F. Sherrick, & J. S. Warm (Eds.), Viewing psychology as a whole: The integrative science of William N. Dember (pp. 145–166). Washington DC.

Scherer, R., Lee, F., Schlogl, A., Leeb, R., Bischof, H., & Pfurtscheller, G. (2008). Toward self-paced brain–computer communication: navigation through virtual worlds. *IEEE Transactions on Biomedical Engineering*, *55*, 675–682. doi:10.1109/tbme.2007.903709

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, *5*, 205–221. doi:10.1007/BF01323044

Schoemaker, P. J. H. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, 529–563.

Schooler, J. W. (2014). Metascience could rescue the 'replication crisis.' *Nature News*, *515*, 9–9. doi:10.1038/515009a

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*, 319–326. doi:10.1016/j.tics.2011.05.006

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50*, 864–878. doi:10.1037/e578082012-006

Schvaneveldt, R. W., Beringer, D. B., & Lamonica, J. A. (2001). Priority and organization of information accessed by pilots in various phases of flight. *The International Journal of Aviation Psychology*, *11*, 253–280. doi:10.1037/e406222004-001

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, *134*, 207. doi:10.1037/0033-2909.134.2.207

Shaw, T. H., Emfield, A., Garcia, A., de Visser, E., Miller, C., Parasuraman, R., & Fern, L. (2012). Evaluating the Benefits and Potential Costs of Automation Delegation for Supervisory Control of Multiple UAVs. In *Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, pp. 1498–1502). doi:10.1177/154193121005401930

Shaw, T. H., Funke, M. E., Dillard, M., Funke, G. J., Warm, J. S., & Parasuraman, R. (2013). Event-related cerebral hemodynamics reveal target-specific resource allocation for both “go” and “no-go” response-based vigilance tasks. *Brain and Cognition*, *82*, 265–273. doi:10.1016/j.bandc.2013.05.003

Shaw, T. H., Matthews, G., Warm, J. S., Finomore, V. S., Silverman, L., & Costa, P. T. (2010). Individual differences in vigilance: Personality, ability and states of stress. *Journal of Research in Personality*, *44*, 297–308. doi:10.1016/j.jrp.2010.02.007

Shaw, T. H., Satterfield, K., Ramirez, R., & Finomore, V. (2013). Using cerebral hemovelocity to measure workload during a spatialised auditory vigilance task in novice and experienced observers. *Ergonomics*, *56*, 1251–1263. doi:10.1080/00140139.2013.809154

Shaw, T. H., Warm, J. S., Finomore, V., Tripp, L., Matthews, G., Weiler, E., & Parasuraman, R. (2009). Effects of sensory modality on cerebral blood flow velocity during vigilance.

Neuroscience Letters, *461*, 207–211. doi:10.1016/j.neulet.2009.06.008

Sherbino, J., Dore, K. L., Wood, T. J., Young, M. E., Gaissmaier, W., Kreuger, S., & Norman, G. R. (2012). The relationship between response time and diagnostic accuracy. *Academic Medicine*, *87*, 785–791. doi:10.1097/ACM.0b013e318253acbd

Sheridan, T. B. (1970). On how often the supervisor should sample. *IEEE Transactions on Systems Science and Cybernetics*, *6*, 140–145. doi:10.1109/TSSC.1970.300289

Simon, H. A. (1978). Rationality as process and as product of thought. *The American Economic Review*, *68*, 1–16. doi:10.1017/cbo9780511598951.005

Simon, H. A. (1992). What is an "explanation" of behavior? *Psychological Science*, *3*, 150–161. doi:10.1111/j.1467-9280.1992.tb00017.x

Singer, R. N., Cauraugh, J. H., Chen, D., Steinberg, G. M., & Frehlich, S. G. (1996). Visual search, anticipation, and reactive comparisons between highly-skilled and beginning tennis players. *Journal of Applied Sport Psychology*, *8*, 9–26. doi:10.1080/10413209608406305

Skånér, Y., Strender, L.-E., & Bring, J. (1998). How do GPs use clinical information in their judgements of heart failure?: A Clinical Judgement Analysis study. *Scandinavian Journal of Primary Health Care*, *16*, 95–100. doi:10.1080/028134398750003241

Small, A. J., Wiggins, M. W., & Loveday, T. (2014). Cue-based processing capacity, cognitive load and the completion of simulated short-duration vigilance tasks in power transmission control. *Applied Cognitive Psychology*, *28*, 481–487. doi:10.1002/acp.3016

Smallwood, J. (2010). Why the global availability of mind wandering necessitates resource competition: Reply to McVay and Kane (2010). *Psychological Bulletin*, *136*. doi:10.1037/a0018673

Smallwood, J. (2011). Mind-wandering while reading: Attentional decoupling, mindless reading and the cascade model of inattention. *Language and Linguistics Compass*, *5*, 63–77. doi:10.1111/j.1749-818x.2010.00263.x

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience*, *20*, 458–469. doi:10.1162/jocn.2008.20.3.458

Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., Connor, R. O., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition*, *13*, 657–690. doi:10.1016/j.concog.2004.06.003

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*, 946–958. doi:10.1037/0033-2909.132.6.946

Smit, A. S., Eling, P. A. T. ., & Coenen, A. M. . (2004). Mental effort causes vigilance decrease due to resource depletion. *Acta Psychologica*, *115*, 35–42. doi:10.1016/j.actpsy.2003.11.001

Smith, D. G., Brody, D. S., & Wigton, R. (1986). Pleasing patients and the decision to use antibiotics for sore throats. *Medical Decision Making*, *6*, 126–139.

Smith, L., Gilhooly, K., & Walker, A. (2002). Factors influencing prescribing decisions in the treatment of depression: A social judgement theory approach. *Applied Cognitive Psychology*, *17*, 51–63. doi:10.1002/acp.844

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*, 137–149. doi:10.3758/bf03207704

Stanton, N. A., Salmon, P., Walker, G. H., & Jenkins, D. (2009). *Human factors in the design and evaluation of central control room operations*. CRC Press.

doi:10.1201/9781439809921

Stanton, N. A., Young, M. S., & Walker, G. H. (2007). The psychology of driving automation: a discussion with Professor Don Norman. *International Journal of Vehicle Design*, *45*, 289–306. doi:10.1504/ijvd.2007.014906

Steinborn, M. B., Flehmig, H. C., Westhoff, K., & Langner, R. (2009). Differential effects of prolonged work on performance measures in self-paced speed tests. *Advances in Cognitive Psychology*, *5*, 105–113. doi:10.2478/v10053-008-0070-8

Stewart, T. R. (1988). Judgment analysis: procedures. *Human Judgment: The SJT View*, 41–74. doi:10.1016/s0166-4115(08)62170-6

Stewart, T. R., Roebber, P. J., & Bosart, L. F. (1997). The importance of the task in analyzing expert judgment. *Organizational Behavior and Human Decision Processes*, *69*, 205–219. doi:10.1006/obhd.1997.2682

Sturman, D., Wiggins, M. W., Auton, J. C., & Loft, S. (2019). Cue utilization differentiates resource allocation during sustained attention simulated rail control tasks. *Journal of Experimental Psychology: Applied*, Advance online publication. doi:10.1037/xap0000204

Stuss, D. T., Shallice, T., Alexander, M. P., & Picton, T. W. (1995). A multidisciplinary approach to anterior attentional functionsa. *Annals of the New York Academy of Sciences*, *769*, 191–212. doi:10.1111/j.1749-6632.1995.tb38140.x

Tai, R. H., Loehr, J. F., & Brigham, F. J. (2006). An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments. *International Journal of Research & Method in Education*, *29*, 185–208. doi:10.1080/17437270600891614

Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors*, *16*, 339–352. doi:10.1177/001872087401600402

Temple, J. G., Warm, J. S., Dember, W. N., Jones, K. S., LaGrange, C. M., & Matthews, G. (2000). The effects of signal salience and caffeine on performance, workload, and stress in an abbreviated vigilance task. *Human Factors*, *42*, 183–194. doi:10.1518/001872000779656480

Thackray, R. I., Bailey, J. P., & Touchstone, R. M. (1977). Physiological, subjective, and performance correlates of reported boredom and monotony while performing a simulated radar control task. In *Vigilance* (pp. 203–215). Springer. doi:10.1007/978-1-4684-2529-1_12

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: A simulator study. *Accident Analysis and Prevention*, *35*, 381–391. doi:10.1016/S0001-4575(02)00014-3

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*, 82–96. doi:10.1177/1745691614556681

Underwood, G., Chapman, P., Bowden, K., & Crundall, D. (2002). Visual search while driving : skill and awareness during inspection of the scene, *5*, 87–97.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, *46*, 629–646. doi:10.1080/0014013031000090116

Underwood, G., Crundall, D., & Chapman, P. (2011). Driving simulator validation with hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour*, *14*, 435–446. doi:10.1016/j.trf.2011.04.008

Vicente, K. J., Roth, E. M., & Mumaw, R. J. (2001). How do operators monitor a complex, dynamic work domain? The impact of control room technology. *International Journal of Human-Computer Studies*, *54*, 831–856. doi:10.1006/ijhc.2001.0463

Visser, T., Tichon, J. G., & Diver, P. (2012). Reducing the dangers of operator distraction through simulation training. *Asia-Pacific Simulation and Training Conference and Exhibition*, 18–21.

Walker, N., & Catrambone, R. (1993). Aggregation bias and the use of regression in evaluating models of human performance. *Human Factors*, *35*, 397–411. doi:10.1177/001872089303500302

Waller, M. J., Gupta, N., & Giambatista, R. C. (2004). Effects of Adaptive Behaviors and Shared Mental Models on Control Crew Performance. *Management Science*, *50*, 1534–1544. doi:10.1287/mnsc.1040.0210

Warm, J. S. (1984). *Sustained attention in human performance*. New York: Wiley.

Warm, J. S., Matthews, G., & Parasuraman, R. (2009). Cerebral Hemodynamics and Vigilance Performance. *Military Psychology*, *21*, 75–100. doi:10.1080/08995600802554706

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

Watson, M. R., Brennan, A. A., Kingstone, A., & Enns, J. T. (2010). Looking versus seeing: Strategies alter eye movements during visual search. *Psychonomic Bulletin and Review*, *17*, 543–549. doi:10.3758/PBR.17.4.543

Weiss, D. J., & Shanteau, J. (2003). Empirical assessment of expertise. *Human Factors*, *45*, 104–116. doi:10.1518/hfes.45.1.104.27233

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickers (Ed.), *Attention and Performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*, 159–177. doi:10.1080/14639220210123806

Wickens, C. D. (2008). Multiple Resources and Mental Workload. *Human Factors*, *3*, 449–455. doi:10.1518/001872008X288394

Wickens, C. D., & Flach, J. M. (1988). Information processing. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in Aviation* (pp. 111–155). San Diego, CA: Academic Press.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance (3rd ed.).* Upper Saddle River, New Jersey: Prentice Hall.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. Psychology Press.

Wiggins, M. W. (2011). Vigilance decrement during a simulated general aviation flight. *Applied Cognitive Psychology*, *25*, 229–235. doi:10.1002/acp.1668

Wiggins, M. W. (2012). The role of cue utilisation and adaptive interface design in the management of skilled performance in operations control. *Theoretical Issues in Ergonomics Science*, *15*, 283–292. doi:10.1080/1463922X.2012.724725

Wiggins, M. W. (2014). Differences in situation assessments and prospective diagnoses of simulated weather radar returns amongst experienced pilots. *International Journal of Industrial Ergonomics*, *44*, 18–23. doi:10.1016/j.ergon.2013.08.006

Wiggins, M. W. (2015a). Diagnosis and Instructional Systems Design. In M. W. Wiggins & T. Loveday (Eds.), *Diagnostic Expertise in Organizational Environments* (pp. 69–80). Surrey, England: Ashgate Publishing Limited, Surrey, England.

Wiggins, M. W. (2015b). *Diagnostic Expertise in Organizational Environments.* Surrey: Ashgate.

Wiggins, M. W., Azar, D., Hawken, J., Loveday, T., & Newman, D. (2014). Cue-utilisation typologies and pilots' pre-flight and in-flight weather decision-making. *Safety Science*, *65*, 118–124. doi:10.1016/j.ssci.2014.01.006

Wiggins, M. W., Brouwers, S., Davies, J., & Loveday, T. (2014). Trait-based cue utilization and initial skill acquisition: Implications for models of the progression to expertise. *Frontiers in Psychology*, *5*, 1–8. doi:10.3389/fpsyg.2014.00541

Wiggins, M. W., & Loveday, T. (2015). *Diagnostic expertise in organizational environments.* Surrey, England: Ashgate Publishing, Ltd.

Wiggins, M. W., Loveday, T., & Auton, J. C. (2015). *EXPERT Intensive Skills Evaluation (EXPERTise) Test*. Sydney: Macquarie University.

Wiggins, M. W., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, *1*, 305–320. doi:10.1037/1076-898X.1.4.305

Wiggins, M. W., & O'Hare, D. (2003a). Expert and novice pilot perceptions of static in-flight images of weather. *International Journal of Aviation Psychology*, *13*, 173–187. doi:10.1207/S15327108IJAP1302_05

Wiggins, M. W., & O'Hare, D. (2003b). Weatherwise: Evaluation of a Cue-Based Training Approach for the Recognition of Deteriorating Weather Conditions during Flight. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*, 337–345. doi:10.1518/hfes.45.2.337.27246

Wiggins, M. W., Stevens, C., Howard, A., Henley, I., & O'Hare, D. (2002). Expert, intermediate and novice performance during simulated pre-flight decision-making. *Australian Journal of Psychology*, *54*, 162–167. doi:10.1080/00049530412331312744

Wiggins, M. W., Whincup, E., & Auton, J. C. (2018). Cue utilisation reduces effort but increases arousal during a process control task. *Applied Ergonomics*, *69*, 120–127. doi:10.1016/j.apergo.2018.01.012

Wigton, R. S. (1988). Use of linear models to analyze physicians' decisions. *Medical Decision Making*, *8*, 241–252. doi:10.1177/0272989x8800800404

Wigton, R. S. (1996). Social Judgement Theory and Medical Judgement. *Thinking & Reasoning*, *2*, 175–190. doi:10.1080/135467896394492

Wigton, R. S., Patil, K. D., & Hoellerich, V. L. (1986). The effect of feedback in learning clinical diagnosis. *Journal of Medical Education*, *61*, 816–822. doi:10.1177/0272989X8800800404

Williams, A. M., Ward, P., Knowles, J. M., & Smeeton, N. J. (2002). Anticipation skill in a real-world task: Measurement, training, and transfer in tennis. *Journal of Experimental Psychology: Applied*, *8*, 259–270. doi:10.1037/1076-898X.8.4.259

Young, M. S., Brookhuis, K. A., Wickens, C. D., & Hancock, P. A. (2015). State of science: Mental workload in ergonomics. *Ergonomics*, *58*, 1–17. doi:10.1080/00140139.2014.956151

Young, M. S., Robinson, S., & Alberts, P. (2009). Students pay attention! Combating the vigilance decrement to improve learning during lectures. *Active Learning in Higher Education*, *10*, 41–55. doi:10.1177/1469787408100194

Young, M. S., & Stanton, N. A. (2002). Malleable Attentional Resources Theory: A new explanation for the effects of mental underload on performance. *Human Factors*, *44*, 365–375. doi:10.1518/0018720024497709

Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, *43*, 303–311. doi:10.1016/j.jesp.2006.02.004

Appendix: Ethics Approval Letters

Appendix (Ethics approval letters) of this thesis has been removed as it may contain sensitive/confidential content