Remotely Piloted Aircraft: The impact of audiovisual feedback and workload on operator performance

Author:
Dunn, Matthew

Publication Date:
2023

DOI:
https://doi.org/10.26190/unsworks/24887

Remotely Piloted Aircraft: The impact of audiovisual feedback and workload on operator performance

Matthew J. M. Dunn

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Aviation

Faculty of Science

November 2022

This page intentionally blank

THESIS / DISSERTATION SHEET

Surname/Family Name	:	DUNN
Given Name/s	:	MATTHEW JOSEPH
Abbreviation for degree as given in the University calendar	:	PhD
Faculty	:	SCIENCE
School	:	AVIATION
Thesis Title	:	REMOTELY PILOTED AIRCRAFT: THE IMPACT OF AUDIOVISUAL FEEDBACK AND WORKLOAD ON OPERATOR PERFORMANCE

Abstract

Remotely piloted aircraft systems (RPAS) offer new possibilities to a growing civilian field. However, unlike pilots of conventionally piloted aircraft, the remote pilot operates in a sensory deprived environment. A reduction in available sensory cues present unique human factors challenges, hence the aim of the present study was to understand the impact of real-time auditory feedback from the RPA on operator performance. Experiment 1 tested conventional pilots manually flying a small multi-rotor RPA under two visual operating conditions. Experiment 2 introduced a non-pilot participant group in an automated configuration. Experiment 3 retained the methodology of the second, but presented an adaptable auditory feedback component, with incremental volumes above and below a comfortable sound level. Key findings revealed transitional instances of auditory feedback being perceived as either sound (i.e., useful information, and/or arousing) or noise (i.e., sound that is unwanted), with results broadly aligned with predicted values associated with behavioural models of performance such as the Maximal Adaptability Model. In addition, no statistical significant differences in task performance (automated flight conditions) between the pilot and non-pilot participants were evident. Together, these findings suggest the ability to include or remove the availability of sensory cueing for remote pilots should be dependent on the stage of flight and associated workload. In addition, they raise questions about the restrictions imposed on who is permitted to operate an RPA. Practically, this gives credence to the inclusion of adaptable sensory cueing in future systems. Furthermore, consideration should be given to licensing operators without conventional flying experience for more complex, automated RPAS operations.

Thesis submission for the degree of Doctor of Philosophy

Thesis Title and Abstract | Declarations | Inclusion of Publications Statement | Corrected Thesis and Responses

ORIGINALITY STATEMENT

☑ I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

COPYRIGHT STATEMENT

AUTHENTICITY STATEMENT

☑ I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.

Thesis submission for the degree of Doctor of Philosophy

Thesis Title and Abstract | Declarations | **Inclusion of Publications Statement** | Corrected Thesis and Responses

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in the candidate's thesis in lieu of a Chapter provided:

- The candidate contributed **greater than 50%** of the content in the publication and are the "primary author", i.e. they were responsible primarily for the planning, execution and preparation of the work for publication.
- The candidate has obtained approval to include the publication in their thesis in lieu of a Chapter from their Supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis.

☑ The candidate has declared that **some of the work described in their thesis has been published and has been documented in the relevant Chapters with acknowledgement**.

A short statement on where this work appears in the thesis and how this work is acknowledged within chapter/s:

> Some of the work described in my thesis (literature review and experiment chapters) has been published. The results from two experiments, comprising Chapter 4 and Chapter 5 were published in the peer-reviewed journal Ergonomics. The applicable citation is included for reference at the commencement of these chapters.

Candidate's Declaration

I declare that I have complied with the Thesis Examination Procedure.

RESEARCH SUPPORT ACKNOWLEDGEMENT

‘I acknowledge the support and assistance provided by the Australian Government Research Training Program Scholarship and the UNSW Scientia PhD Scholarship in producing this thesis.’

ABSTRACT

Remotely piloted aircraft systems (RPAS) offer new possibilities to a growing civilian field. However, unlike pilots of conventionally piloted aircraft, the remote pilot operates in a sensory deprived environment. A reduction in available sensory cues present unique human factors challenges, hence the aim of the present study was to understand the impact of real-time auditory feedback from the RPA on operator performance. Experiment 1 tested conventional pilots manually flying a small multi-rotor RPA under two visual operating conditions. Experiment 2 introduced a non-pilot participant group in an automated configuration. Experiment 3 retained the methodology of the second, but presented an adaptable auditory feedback component, with incremental volumes above and below a comfortable sound level. Key findings revealed transitional instances of auditory feedback being perceived as either sound (i.e., useful information, and/or arousing) or noise (i.e., sound that is unwanted), with results broadly aligned with predicted values associated with behavioural models of performance such as the Maximal Adaptability Model. In addition, no statistical significant differences in task performance (automated flight conditions) between the pilot and non-pilot participants were evident. Together, these findings suggest the ability to include or remove the availability of sensory cueing for remote pilots should be dependent on the stage of flight and associated workload. In addition, they raise questions about the restrictions imposed on who is permitted to operate an RPA. Practically, this gives credence to the inclusion of adaptable sensory cueing in future systems. Furthermore, consideration should be given to licensing operators without conventional flying experience for more complex, automated RPAS operations.

ACKNOWLEDGEMENTS

There are parallels between the PhD journey and operating remotely piloted aircraft: despite often being conducted out of sight, they both require the coordination of a great deal of individuals and teams working together behind-the-scenes to enable success. I was fortunate to be awarded the Scientia PhD Scholarship and have benefited from it immensely during my candidature, both in career development activities and educational support. The scholarship afforded me many wonderful experiences which I would not have engaged in otherwise.

There are many people and groups I would like to acknowledge who have supported me in this endeavour and it would not be possible without them.

To my extended family thank you for your well wishes over the years. To my immediate family - my brother Sean, mother Carmel and late father Robert, thank you for your love, support and encouragement throughout my life.

To the doctors and nurses who have provided care to me as an acromegaly patient for more than two decades, you’ve enabled me to pursue a normal life.

To my colleagues at BOM who have permitted me to complete this research while flexibly maintaining part-time work, thank you for your understanding.

To staff in other areas of UNSW including the Scientia PhD team, and the DVCR UNSW Drones team, thank you for facilitating the career development activities during my candidature and enabling the research to proceed.

To the academic and professional staff within the School of Aviation, thank you for creating a welcoming and collegial work environment that has smoothed my candidature experience, and allowed the research to take place.

To the students within the School, past and present, thank you for the support and advice you have provided to me over the years, and for the social engagements I've been invited to.

To the participants who volunteered their time to take part in the research, it is very much appreciated. The findings reported in this thesis would not be possible without them.

To my co-supervisor, Professor Gabriel Lodewijks, thank you for your mentorship, guidance and insight, which has been appreciated. To my co-supervisor Dr. Tay Koo, thank you for first supporting my application to commence this research journey and for inclusion in the peripheral projects along the way, particularly the work with Dhimmuru.

Finally to my primary supervisor, Professor Brett Molesworth. Your ongoing support, feedback and advice has been greatly appreciated. From the start of the journey in 2017 and to now, I thank you for your time and the motivation you have given me over the years.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY

The International Civil Aviation Organization (ICAO, 2018) defines the following terms which will be used in the chapters of this thesis:

Remote Pilot: A person charged by the operator with duties essential to the operation of a remotely piloted aircraft and who manipulates the flight controls, as appropriate during flight time.

Remotely Piloted Aircraft (RPA): An unmanned aircraft which is piloted from a remote pilot station.

Remotely Piloted Aircraft System (RPAS): A remotely piloted aircraft, its associated remote pilot station(s), the required command and control links and any other components as specified in the type design.

LIST OF ABBREVIATIONS

AAIB: Air Accidents Investigation Branch

AGL: Above Ground Level

ANOVA: Analysis of Variance

AT: Arousal Theory

ATSB: Australian Transport Safety Bureau

BEA: Bureau d'Enquêtes et d'Analyses

BVLOS: Beyond Visual Line of Sight

BVLOS-M: BVLOS Monitor

BVLOS-G: BVLOS Goggles

CAA: Civil Aviation Authority

CASA: Civil Aviation Safety Authority

CASR: Civil Aviation Safety Regulations

CFIT: Controlled Flight Into Terrain

dBA: A-weighted Decibel

DJI: Dà-Jiāng Innovations

EEG: Electroencephalography

ER: Extended Range

EVLOS: Extended Visual Line of Sight

FAA: Federal Aviation Administration

FOV: Field-of-View

FPV: First-Person-View

GCS: Ground Control Station

GPS: Global Positioning System

HP: Hewlett Packard

ID: Identification

IFR: Instrument Flight Rule

ICAO: International Civil Aviation Organization

LED: Light-Emitting Diode

M: Mean

MAM: Maximal Adaptability Model

MP: Mishap Pilot

MRPA: Mishap Remotely Piloted Aircraft

MOT: Ministry of Transport (Israel)

NASA: National Aeronautics and Space Administration

NASA-TLX: NASA Task Load Index

NTSB: National Transportation Safety Board

RMSE: Rating Scale for Mental Effort

RPA: Remotely Piloted Aircraft

RPAS: Remotely Piloted Aircraft System

SD: Standard Deviation

SHEL: Software, Hardware, Environment, and Liveware

STSB: Swiss Transportation Safety Investigation Board

SWAT: Subjective Workload Assessment Technique

TSB: Transportation Safety Board of Canada

UK: United Kingdom

UNSW: University of New South Wales

USAF: United States Air Force

VLOS: Visual Line of Sight

VPS: Vision Positioning System

VTOL: Vertical Take-off and Landing

WP: Workload Profile

CHAPTER 1 – INTRODUCTION

The aim of this study is to examine the effects of additional workload and sensory stressors, namely auditory feedback on the task performance of remote pilots. Distinguishable from conventionally piloted aircraft, remote pilots operate within a sensory deprived environment, and typically rely upon visual sensory feedback exclusively (Tvaryanas et al., 2006; Hobbs, 2017). They also work under various operational conditions, some of which are cognitively taxing. How workload and the absence of auditory feedback affects remote pilots has received little attention in the literature, and hence is the primary aim of the present study.

Possibly not surprising to some, the reduction or absence of sensory cueing on its own, as well as combined with fluctuating workload has been attributed to a number of remotely piloted aircraft accidents. Remotely Piloted Aircraft Systems (RPAS) accident rates are notably higher compared to conventional aircraft (Williams, 2004; ATSB, 2018; ATSB, 2020), with some estimates reporting their occurrences as much as ten to fifty times greater (Weibel & Hansman, 2005; Jaussi & Hoffmann, 2018; Renshaw & Wiggins, 2020). In an attempt to address this high accident rate, professionals including aviation safety professionals often draw on knowledge and information about manned aircraft. While such information can be useful, it has its limitations, primarily because of several key fundamental differences, the most obvious being the lack of sensory auditory information. Therefore, understanding how both workload and auditory information affects remotely piloted operators is important if improvements in RPAS safety are desired.

RPA are presently equipped with a number of sensory tools, possibly the most important being a camera (with the exception of some highly automated military RPA; Bergmann, 2017). Cameras relay real-time monocular visual information back to the remote pilot via various means, including a tablet, monitor or goggles display. However, they are void of any substantive depth perception, thus limiting available visual stimuli (Smolyanskiy & Gonzalez-Franco, 2017; Luo et al., 2021). Visually, RPAS can be operated under Visual Line of Sight (VLOS), which requires the remote pilot to physically sight the RPA unaided. If permitted, RPA can also be operated Extended/Beyond Visual Line of Sight (E/BVLOS), which means beyond the view of the remote pilot (theoretically unlimited distance). Contingent on the camera field-of-view and resolution, imagery transmitted at present, even with the most sophisticated RPAS are inferior to that received by conventional pilots.

Possibly the next most important sensory information for RPAS operators is auditory. Auditory information from an RPAS and conventional aircraft contains key information about its state relative to the environment, and includes information about engine power settings, meteorological conditions and/or mechanical failures (Walker, 1997). This information often complements visual information, providing critical cues for changing aircraft state. However, a lack of either piece of information increases the likelihood of an incident, as illustrated through the following two examples.

In 2007 at the Kandahar Air Base during a landing of an MQ-1B Predator military RPA, the RPA crashed nose-first into the runway (USAF, 2007b). Utilising the transmitted real-time visual feed, the remote pilot misjudged the height of their aircraft above the runway, and interpreted the physical touchdown as a response to their controlled flare input. Conceivably, the expectation of a similar visual scene associated with a flare control input, versus an inadvertent touchdown, was reinforced to the remote pilot. With only a limited visual field-of-view, the remote pilot was oblivious to the aircraft bouncing due to subsequent incorrect control input responses. That is, at a critical stage of flight, a dynamic scene was visually ambiguous, and the RPA's touchdown on the runway was not apparent to the remote pilot (USAF, 2007b; Arrabito et al., 2011). Access to additional sensory information may have resolved this ambiguity, such as screeching tyres on touchdown, or variations in airflow.

Perhaps obvious to some, and in particular to a conventional pilot, a reduction of visual, auditory, somatic, vestibular, proprioceptive, and olfactory feedback received by remote pilots can be problematic (Drury & Scott, 2008; Hobbs & Lyall, 2016; Blumer et al., 2018). Take for example auditory information. With an aircraft, such information can be forthcoming from the engine/power plant (i.e., RPA propeller/s) and airframe. Auditory information from engines reflect associated changes in response to power commands, variations in wind/airflow, and engine state (Stearman, 1997; Hobbs, 2017). For example, an increase in power to climb, or a decrease in power to descend will be associated with changes to the frequency, tone, and sound intensity of the engine/propeller (Baklanov, 2016; Schlüter, 2016; Tuccio et al., 2017; Moshkov, 2020). While technological advancements have enabled quieter cockpit environments in modern commercial aircraft (i.e., new generation engines, sound insulation, noise-cancelling headphones; Leylekian et al., 2014; Burgess & Molesworth, 2016; Bravo et al., 2022; Lee et al., 2022), contrary to remote pilots operating under E/BVLOS, the intensity of auditory information experienced by conventional pilots has not reduced to zero. The importance of such information is best illustrated in the below accident which occurred in 2016.

At the Nevada Test and Training Range in 2016, a U.S. Air Force MQ-9 was lost, attributed in part to a failure by the remote pilot to recognise the reduced energy state of their RPA (USAF, 2017). During a changeover between remote crews, a climb was inadvertently initiated. Upon realising, the remote pilot changed to a manual configuration, reducing power from 100 to 30 percent for the descent, but not accounting for the aircraft pitch angle which remained positive. The aircraft quickly lost airspeed and lift, resulting in an aerodynamic stall (USAF, 2017; Jaussi & Hoffman, 2018). While visual and audible alerts for the stall warning occurred, the remote pilot failed to observe these, possibly as a result of the sensory deprivation encountered by the remote pilots; a key distinction from other pilots (Jaussi & Hoffman, 2018). Furthermore, as noted in the investigation report, the remote pilot did not prioritise the handover checklist and a ‘loss of situational awareness was a substantially contributing factor to the mishap’ (USAF, 2017, p. 12). Had real-time auditory feedback been made available, perhaps critical auditory information about the changing aerodynamic state of the RPA could have prompted a different and more timely response.

Combining sensory information sources is arguably a necessary method to increase the value and meaning of environmental/operational settings. However, not all sensory information is positive or desired. When there are limited cognitive resources available, additional sensory information has the potential to increase mental workload beyond capacity. In response to such conditions, adaptations can be made by the human to accommodate surplus or redundant information, by shedding or ignoring such information.

However, sometimes useful sensory information can be perceived as a stressor, known to adversely affect performance. For remote pilots with variable workloads, the requirements to maintain spatial orientation, selective attention, auditory attention, situational awareness, visual perception, spatial processing, attention to detail, and conscientiousness (Howse, 2011) could be considered additional stressor sources. Combined with sensory cueing as a further stressor source, these can be cognitively taxing, and hence positive or detrimental to performance.

One way to alleviate the effect of stressors on cognitive resource consumption is through the use of automation (Prinzel III et al., 2002; Freeman et al., 2004). Automation can be engaged to assist in managing workload in different operational conditions, allowing for better focus on a primary task (Cummings et al., 2014; Dixon et al., 2005; Dixon & Wickens, 2006; McKendrick et al., 2014; Mekdeci & Cummings, 2009). While RPA can be programmed or flown manually, the abilities of the remote pilot to conduct complex flying tasks are greatly aided by an automated system (Stevenson et al., 2015a). However, an overreliance on automation can also be burdensome. The requirement to monitor automation can have a detrimental effect on performance through under arousal or boredom (Cummings et al., 2013). Hence, using automation to moderate workload, coupled with additional sensory cues, does not necessarily lead to a simple, dichotomous outcome. Moreover, there is some evidence to suggest that as RPAS technology continues to develop with increased automation, human remote pilots could place too much trust in the human-machine system (Liu et al., 2015), and thus become further removed from it. Therefore, understanding the relationship between variables such as workload and sensory cueing on the task performance of remote operators is important, and thus is one of the aims of the present study.

In pursuit of this aim, the thesis is structured as follows. Chapter 2 will discuss auditory, visual, and multimodal sensory cueing as they apply to remote pilots. Chapter 3 will present the relevant literature for workload and behavioural models relating to performance, arousal, and adaptability, including as they apply to piloting and RPAS. Following these chapters, three experiments will be presented in Chapter 4 through 6. In Experiment 1 (Chapter 4), the effect of real-time auditory feedback on remote pilot task performance will be examined. However, this effect will only be examined for pilots with previous fixed-wing conventional flying experience. Their manual flying performance will be measured using a small RPA, under two workload levels, and utilising three methods of visual feedback. Experiment 2 (Chapter 5) will be expanded to include non-pilots, as well as pilots, under automated operating conditions. Workload will be extended to three levels, and will measure participant performance in separate perception and decision-making tasks. Experiment 3 (Chapter 6) will maintain the methods of Experiment 2, but will measure non-pilots only. The strategies used by participants to deal with the auditory feedback will also be assessed, with two additional volume intensities introduced. Common to all three experiments is the inclusion and removal of real-time auditory feedback from the RPA, and the mapping of the results to the Maximal Adaptability Model. Finally, a discussion of the key findings will be presented before a conclusion in Chapter 7.

CHAPTER 2 – AUDITORY, VISUAL, AND MULTIMODAL CUEING

2.1 Overview

This chapter will describe the theory relating to auditory (i.e., nonverbal, as distinguished from speech) and visual information (including visual display types), and multi-modal feedback (i.e., auditory-visual pairing) important to remote pilots. As will be discussed, auditory and visual feedback provide important and useful information, essential for safe aircraft operations. However, both benefits and detriments of single and combined sensory information sources exist, potentially affecting successful task completion. Because remote pilots operate within a sensory deprived environment, it is therefore pertinent to discuss these elements for a better understanding and prediction of their performance.

2.2 Auditory Cueing

2.2.1 The Auditory Cortex

When a physical object vibrates sufficiently enough to change the pressure of its surrounding medium (e.g. air), sound waves are created (Moore, 2012). The human auditory system senses these pressure waves when captured by the outer ears. From the outer ear, they are directed to the inner ear, and transduced into neural code, before transferring via the auditory nerve to the primary auditory cortex in the brain for analysis and processing (Møller, 1983; Belin et al., 1998; Pickles, 2015; Schacter et al., 2011). The physical positioning of the two outer ears allows for the rapid analysis of sound waves by the auditory cortex, distinguishing between the perceived differential pitch, loudness, timbre (i.e., the quality of sound allowing for differentiation between two sources of identical pitch) and spatial localisation of the sound (Bizley & Cohen, 2013; Joris et al., 2004; Schacter et al., 2011; Zatorre et al., 2002; Zaehle et al., 2004). It is through this system that humans perceive and experience their surroundings auditorily (Litovsky, 2015).

2.2.2 Auditory Information

For humans perceiving sound (i.e., auditory information), its intensity or loudness (i.e., sound level) is defined and measured through the unit A-weighted decibel (dBA; Chepesiuk, 2005; Litovsky, 2015; Molesworth, 2016). Sound sources include both verbal (e.g., human speech) and nonverbal (i.e., artificial or natural). In aviation, sources of sound can be generated by the powerplant output (i.e., engine), aerodynamic contact between the outside air and the aircraft/helicopter fuselage or control surfaces, and verbal communications via radio (Antuñano & Spanyers, 2006). While the loudness of sound is subjectively perceived by individuals (Litovsky, 2015), some examples of common sound intensities measured in dBA are given by Chepesiuk (2005) including: a quiet room (28 – 33 dBA), normal conversation (40 - 65 dBA), using a vacuum cleaner (62 – 85 dBA), or operating a hair dryer (59 – 90 dBA). In comparison, sound intensities within conventional aircraft can range from 60 dBA – 88 dBA for commercial jets (Antuñano & Spanyers, 2006; Ozcan & Nemlioglu, 2006), to 80 dBA – 102 dBA for small helicopters (Antuñano & Spanyers, 2006).

For conventional pilots, sound pressures experienced through combined ambient noise and communications in the cockpit have been estimated at 85 dBA in commercial airliners (Müller & Schneider, 2017), 95 dBA in smaller fixed-wing aircraft, and 100 dBA in helicopters (Gasaway, 1986). Although differing in intensity, each source of sound has the potential to induce physiological and/or psychological responses in humans (Jerison, 1959; Poulton, 1979; Szalma & Hancock, 2011; Molesworth, 2016). For tasks requiring higher cognitive abilities, the relationship between auditory information and task performance can be variable (Banbury et al., 2001). For example, depending on the type of auditory information being processed (e.g., speech-like versus white noise), its effect on cognitive tasks like memory recall, can be potentially distracting (Banbury et al., 2001; Sörqvist, 2010). Hence, it is important to distinguish between the information type, and in reference to the task category being performed.

To make sense of most auditory information, auditory scene analysis occurs whereby multiple, simultaneous sounds are separated by the auditory system, allowing for their identification and organisation (Bregman, 1994; Baldwin, 2012). As described by Baldwin (2012), two combined processes are involved in auditory cognition: a bottom-up process for the sensing and attention given to a sound, and a top-down process for interpreting and making sense of it (Baldwin, 2012; Guastavino, 2018). However, auditory information is not always presented in isolation, and as such, must be attended to and processed using finite available resources (Baldwin, 2012). In the case of multiple sound sources presented concurrently, individuals have a limited capacity to process this information (Moray, 1967).

The processes involved with selection and attendance to a sound are described by several different theories. Broadbent's Early Filter Model (Broadbent, 1958) proposes that due to a 'bottleneck' of sensory overload (e.g., multiple sound sources), a filtering process occurs whereby auditory information is selected for further processing, or rejected (Eysenck, 1994; Lachter et al., 2004). Treisman counters this theory by suggesting multiple sources of auditory information, could instead be stored for later processing, rather than outrightly rejected (Treisman, 1960; Treisman, 1964). Late selection and limited-capacity models are also used to explain how auditory information processing occurs, in combination with other systems including working and long-term memory (Baldwin, 2012). Irrespective of the mechanisms involved, because auditory information demands a level of processing, there will also be a corresponding impact on cognitive performance.

Auditory information processing can also be susceptible to the kind of sound generated, and the type of cognitive tasks involved. Banbury and colleagues (2001) reviewed a number of studies that measured cognitive performance and the interfering effect of irrelevant sound, defined as sound that can be distracting to complex mental tasks (e.g., serial recall, comprehension). The authors noted several key findings. Firstly, acoustic variations (i.e., pitch, timbre, tempo) to the auditory information are more determinant of the disruption caused to task performance, compared to monotonous or repetitive sounds. Furthermore, these variations to the nonverbal sound can be as similarly distracting as verbal sounds (i.e., speech; Banbury et al., 2001). Importantly, the authors found irrelevant sounds do not need to be presented simultaneously with the task being undertaken, for a similarly detrimental effect on performance to be observed (Banbury et al., 2001). That is, it can be detrimental, through a temporally lasting effect. This seemingly has implications to an applied setting like an aircraft cockpit, regardless of it being conventionally or remotely located, where the completion of complex tasks is intermittently required, and where an exposure to (and processing of) irrelevant sound also exists.

Edwards (2007) further elaborated on the involvement of top-down processing (i.e., based on prior knowledge and expectation; Dennett, 1995) by the cognitive system when attending to degraded auditory information quality. If a deteriorated perception of the auditory information exists (e.g., signal degradation or physiological), the cognitive load placed on the individual could be at the expense of other resources (Edwards, 2007). Similarly, Stenfelt and Rönnberg (2009) expanded on the role of distortion in auditory cognition. The authors argued the level of distortion present during the transmission of auditory information, can negatively influence an individual's cognitive processing abilities (Stenfelt & Rönnberg, 2009).

In contrast to degraded auditory information, when it is transmitted without any signal distortion (e.g., no other concurrent sounds, or masking), simple bottom-up cognitive processing is involved. This can occur when the sound is heard and listened to, comprehended, and then used to formulate a reaction (Stenfelt & Rönnberg, 2009). That is, any additional impact to one's cognitive load is limited. However, the cognitive system reverses the direction of the processing when a distortion to the quality of information is introduced, by changing from a bottom-up to top-down method (Stenfelt & Rönnberg, 2009). A reversal of the processing direction and its impact on cognitive load is not necessarily disadvantageous. Due to the perception of an unfamiliar or unexpected sound, an arousal response can potentially be induced. As will be discussed in the next chapter, if combined with an appropriate level of workload, performance could be raised to an optimal level.

Griffiths and Warren (2004) explain the perceptual theory of auditory objects, comprising the acoustic source of the information, and the temporal event that follows. The auditory system transforms information from an acoustic stimulus, to perceived auditory objects (Bizley & Cohen, 2013), the flow of which has been described as an auditory stream (Bregman, 1994; Näätänen et al., 2001). For humans to make sense of these streams, higher cognitive processes in the brain detect, organise, and extract auditory information into distinct auditory objects, including the identification of familiar patterns and anomalous sounds, or concurrent sounds such as musical instruments playing simultaneously within an orchestra (Bizley & Cohen, 2013; Leaver & Rauschecker, 2010; Näätänen et al., 2001). By interacting with the multi-modal cortex and higher cortices, sounds initially processed by the auditory cortex form the individual behavioural reaction to auditory stimuli (Griffiths & Warren, 2004). Hence, auditory information can benefit individuals by utilising it to form a response to stimuli, and act upon task requirements. A review of the literature in this area pertaining to RPAS operations failed to reveal any literature. Therefore, a key objective of the first two experiments presented in this thesis is to compare task performance between an ecologically normal condition (i.e., without auditory feedback) and conditions where real-time auditory information, directly from the RPA is introduced.

Outside of RPAS operations, the positive role of non-verbal auditory information on task performance is evidenced in studies involving individuals’ exposure to music. In one example, Kiger (1989) identified nuances in the role of background music on student comprehension performance. The type of music was separated in two: high information load and low information low (i.e., related to a variable or consistent rhythm, and tonal dissonance), and compared with a silent condition. The comprehension of a reading task was found to be superior under the music condition with a low information load, compared to the high information load and silent conditions. The author offered two explanations for the performance differences: an interference mechanism to attention and concentration, due to competing cognitive resources under high information loads, or an over/under-arousal effect (Kiger, 1989). For similar tasks, Sörqvist (2010) also suggests other mechanisms could be possible, including working memory capacity, the changing-state effect (e.g., versus semantic auditory distraction), and individual differences.

Another study involving nonverbal auditory information was conducted by Fox and Embrey (1972). The authors conducted a series of experiments involving repetitive, visual quality control inspections in both laboratory and factory settings, under a variety of auditory conditions, including with a music component. Improved visual detection performance (i.e., identifying faulty parts on an assembly line) was found when music was introduced, against both silent and background machine-generated noise conditions. As noted by the authors, the duration of the music (i.e., short was determined as better than continuous), and the rhythm (i.e., lively was found to be better than slow/steady), positively influenced performance the most, and could likely be attributed to a stimulatory effect (i.e., arousing; Fox & Embrey, 1972). As explained by Molesworth (2016), this effect can also be seen in tasks that may be boring or repetitive, and which utilise fewer cognitive resources to complete.

Dosseville et al. (2012) also identified an enhancive function of classical music in changing students' perception of a learning environment. A notable shortcoming was admitted by the authors in not including other forms of sound, such as individually preferred music (i.e., non-preferred music has previously been shown to negatively affect comprehension performance compared to silent conditions; Johansson et al., 2012), other music genres or white noise, to determine their effects on learning. As a result of this, no detailed explanation for the mechanism involved in producing a positive effect on academic performance was provided by the authors (Dosseville et al., 2012). Nevertheless, it adds to the evidence suggesting the relationship between nonverbal auditory information and performance, relating directly to the type of activity being undertaken and the cognitive facility required (e.g., repetitive visual monitoring/monotonous versus learning).

The effect of nonverbal auditory information on cognitive performance was further explored by Cockerton et al. (1997). The authors measured participants' task performance on a cognitive test in the presence of background, software-created music, that was considered free-flowing and harmonious (Cockerton et al., 1997). While acknowledging the type of music (e.g., classical versus contemporary, or unknown versus familiar; Cockerton et al., 1997; Schellenberg et al., 2007) may contribute to individual performance differences (Rauscher et al., 1993; Stough et al., 1994), the music used for their study was generated in real-time, and hence novel/unfamiliar. Participants exposed to the music answered more questions, with higher accuracy than those in the no music condition. While these findings were dependent on the task, they support the notion of an arousing or stimulatory effect of auditory information, and its relationship with cognitive performance. Although task and source dependent, this potentially positive attribute of auditory information, and its effect on human performance, will be discussed more broadly in the research chapters.

Other forms of auditory information can have a positive influence on task performance, and additional spatial and temporal information about a remote environment can be delivered to the human user, directly from a source (Günther et al., 2018; Yang et al., 2020). This information has been shown to assist performance, as demonstrated by Akamatsu et al. (1995) in a target selection task. Similarly, Naujoks et al. (2014) found a positive effect of auditory feedback on driver performance when presented with a visual alert during a take-over request in a highly automated vehicle. Under moderate to difficult task conditions, the reaction time and driving accuracy of participants were both improved when an auditory cue was added to the alert system.

The inclusion of additional auditory information in virtual environments also appears to be positive. Where navigation is required, the provision of spatialised auditory feedback has been found to increase the feeling of presence within the virtual location (Hendrix & Barfield, 1996). This finding, in the context of remotely piloted aircraft is important. As Simpson et al. (2013) describe, there is the potential for spatial awareness of remotely located RPA operators to be enhanced using spatial auditory systems when combined with visual display information.

While the introduction or inclusion of auditory information has been demonstrated to enhance performance completion in some tasks (Rausch et al., 2014), the removal of auditory information (i.e., silence/quiet) can also provoke a similarly positive response in some individuals (Furnham & Bradley, 1999; Kämpfe et al., 2011; Johansson et al., 2012; Avila et al., 2012). This was shown by Cassidy and MacDonald (2007) who measured the cognitive performance of forty student introverts and extroverts, exposed to four auditory conditions: positive low arousal music, negative high arousal music, background noise, and silence. The low and high arousal music examples used (i.e., related to structure, timbre and lyrical message; labelled 'relaxing' and 'aggressive' respectively) were chosen through a selection process, subjectively rated by a separate cohort of participants. Overall participant performance measured in the memory and processing (i.e., Stroop test; Stroop, 1935) tasks were found to be superior under the silent auditory condition. As the authors explain, the noise and high arousal music conditions could have been perceived and processed in a similar way, hence their effect on performance being similar (Cassidy & MacDonald, 2007). This was further evident for introverts, who performed poorly on the Stroop test when noise or high arousal music was introduced. The authors suggest for tasks that consume cognitive processing capacity, the addition of auditory information could overwhelm resources, leading to a distraction effect (Cassidy & MacDonald, 2007).

Cassidy and MacDonald's result parallels the findings of Furnham and Bradley (1999), where introverts performed to a lower standard in comprehension and memory tasks, in the presence of background music, compared to extroverts, and to a silent control condition. The results from these studies contrast with the previously described positive and arousing effect of auditory information. With respect to enhanced task performance, a binary effect on performance, manipulated via the crude addition or removal of auditory information to a task, appears to be too simplistic in terms of predicting task performance. That is, other non-auditory variables, such as the context in which the sounds are presented, also appear to interact with and influence performance.

A further source of auditory information can be found via warnings or alerts that are used to attract the attention of individuals (Baldwin et al., 2012). The primary intentions of auditory warnings or alerts are to communicate information with urgency. However, the potential also exists for confusion and/or incorrect responses to be made by the human operator if the quantity of alerts are excessive (Meredith & Edworthy, 1995; Edworthy & Hellier, 2000), or when there is detection difficulty (e.g., in noisy environments; Edworthy & Hellier, 2000). Hence, it is apparent the context in which auditory information is presented (or removed) to individuals, also determines its ability to be detrimental (i.e., noise) or meaningful (i.e., feedback).

2.2.3 Auditory Feedback

Auditory feedback is described as the sound '*produced in response to user actions'* (Frid et al., 2019). In contrast to auditory information which may only exist in the background (e.g., broadband noise), this description of feedback implies an interactive component between the source of the sound, and the human operator. Hence, for the operation of plant, and/or teleoperated equipment, this description of feedback could extend to the produced engine/machine sounds, in response to programmed inputs by the human operator in the case of an automated setting, or from direct manual control. However, if an absence or reduction of auditory feedback exists, task performance may be also affected (McLane & Wierwille, 1975; Hellier et al., 2011).

One function of the rising prevalence of electric automobiles, is a reduction of associated engine sound (Denjean et al., 2012). The reduction in auditory feedback via the removal of sound associated with new engine technology, and its effect on driving performance was studied by Denjean and colleagues (2012) using a car simulator. Following a training period, participants were tasked with accelerating to a target speed, without the assistance of the speedometer, under three auditory conditions, including no sound. The authors found that without auditory feedback from the engine, participants accelerated for a longer period of time, and underestimated their speed (i.e., their performance deteriorated). These findings are corroborated by similar studies involving the removal of auditory feedback and a decline in automobile speed perception (McLane & Wierwille, 1975; Hellier et al., 2011; Denjean et al., 2012), and perceived quality/realism in virtual environments (Rojas et al., 2012; Rojas et al., 2014).

The removal of auditory feedback, however, can also have a beneficial effect on performance. As described by Cowan et al. (2015), and supported by previous studies, (Recanzone; 2003; Mastoropoulou et al., 2005) the effect of concurrently presented sound can detrimentally change visual perception via a distracting effect on visual cognitive processing. Recanzone (2003) elaborated further on this by discussing the modality appropriateness hypothesis, as described by Welch and Warren (1980), and suggesting the sensory system most dominant for perception will depend on the type of task being performed. For spatial tasks, the visual system will be dominant, while for temporal rate discrimination tasks, the auditory system will be dominant. For tasks or workplaces that may include a combination (e.g., remote pilots and teleoperated equipment), determining the shift in dominant perception systems is important. The close association and effects of auditory feedback, combined with visual cues, will be further discussed in relation to multimodal cueing later in this chapter.

In addition to perception, the inclusion of auditory feedback can also be assistive to task completion. Compared to visual-only tasks, Frid et al. (2019) found a positive effect to performance on a virtual throwing task (i.e., visual computer-based, error rate) when an auditory and haptic feedback component were included. Similarly, improved task performance (i.e., duration to complete the task) was also demonstrated by Cowan et al. (2015) in a virtual cognitive skills training environment for surgeons. Relevant auditory feedback (i.e., operating room ambience and associated sounds) was included and considered beneficial, while silent and white noise were perceived as distracting (Cowan et al., 2015).

The transmission of auditory feedback from remote environments can provide further information about systems/equipment, and can positively assist with performance (Akamatsu et al., 1995; Günther et al., 2018). For remote pilots, auditory feedback from the engines or propellers could provide valuable information about their state, as well as any changes to the external environment (e.g., turbulent wind conditions via aerodynamic sound; Howe, 2017). The value of this information has been demonstrated in previous military RPA accidents, which were partially attributed to an absence of auditory feedback (USAF, 2007a; USAF, 2013). Beyond the half-way point of a twenty-hour reconnaissance mission, an MQ-1B Predator fixed-wing RPA experienced a sudden and short-duration (i.e., two seconds) decrease in engine RPM (revolutions per minute). Over the next fifteen minutes and without the remote pilot's awareness, the engine began to fail, evidenced by changes in the oil pressure, turbo oil temperature, RPM and propeller pitch (i.e., via post-crash data logger; USAF, 2007a). Because the engine parameters remained within the range that would not trigger automated alerts or warnings, the remote pilot continued the flight, while oblivious to the changing state of the engine, prior to it seizing. It is, however, likely this failure would have generated an audible change in the powerplant, which would have been noticeable if perceived from inside the cockpit. Hence, having real-time auditory from the RPA in this instance could have alerted the remote pilot to the changing situation.

A different example of the consequences of lacking real-time auditory feedback can be seen in an accident involving an MQ-9 Reaper (i.e., fixed-wing RPA; USAF, 2013). Following an automated segment of the flight, and while navigating to a test range, the remote pilot disengaged the altitude hold mode. An incorrectly configured throttle setting resulted in a reverse thrust engine condition whenever the throttle was not set to full. While the remote pilot initiated a return-to-base command due to a perceived engine problem, the reverse thrust setting was not noticed, and hence resulted in an aerodynamic stall (USAF, 2013). This example further highlights the importance of receiving real-time auditory feedback to convey valuable information about the current state of the remotely located aircraft. As with the MQ-1B Predator fixed-wing RPA accident, had acoustic information been transmitted, it is possible the change in engine sound from the normal operating condition to the reverse thrust configuration could have provided a cue as to the unusual aircraft state.

Despite extreme examples of the consequences relating to the lack of auditory feedback, there is conditional support both for and against its inclusion, which appears to be largely task-dependent. Not all auditory information is equal, and as previously discussed, can induce arousing/stimulatory or distracting effects, evidenced by the ability, or not, to complete tasks successfully. The human operator can be exposed to different forms of auditory information and feedback, sometimes simultaneously, and as such, may not always be interpreted as meaningful. Furthermore, and like with the removal of auditory feedback from a task, its effect on human performance (e.g., if perceived as noise) can also be nuanced.

2.2.4 Noise

When auditory information is perceived as unhelpful or unwanted, it can be categorised as noise (Wagner et al., 1996). Noise can be either intermittent (i.e., short durations) or continuous (i.e., prolonged). Intermittent noise is generally found to be more detrimental to performance than continuous noise (Muzammil & Hasan, 2004; Szalma & Hancock, 2011). Some of these negative effects of noise include reduced perception, memory and attention, slower reaction times (Hygge et al., 2002; Molesworth et al., 2014; Shimizu et al., 2002; Stevenson et al., 2013), increased fatigue (Landström, 1990; Melamed & Bruhis, 1996; Kjellberg et al., 1998), decreased detection performance, and annoyance (Haines et al., 2001; Saeki et al., 2004; Marquis-Favre et al., 2005). Furthermore, masking effects (i.e., concurrently presented auditory information) of noise can also reduce performance when auditory discrimination is required (Molesworth, 2016).

From a human factors research perspective, the potentially negative effects of noise are significant. Distinguishing between auditory feedback and noise is therefore important, as the type of information presented to the individual can influence their task performance through altered cognitive processes (Smith, 1985; Tremblay et al., 2000). Therefore, further investigation into the nuances, and the way auditory information is perceived and processed, and how these differences contribute to human task performance within applied RPAS settings is necessary.

In a study investigating the noise effect with RPAS operations, Dunn and colleagues (2020) found nuanced examples of auditory feedback effects on remote pilot manual flying performance under visual line of sight and BVLOS operating conditions. Specifically, in operating conditions perceived to be reflective of a high level of workload, auditory information presented at a comfortable level (i.e., not too loud, or too soft; individual preference) was found to assist the human operator navigation accuracy in a spotting/perception task. Conversely, in operating conditions perceived to be reflective of a low level of workload, real-time auditory feedback from the RPA hindered task performance. While the sound intensity remained consistent, the differing counterintuitive task performance is a potential example of changing form, from feedback to noise and vice versa.

As noted prior, the intensity of noise experienced by human operators can also have varying effects on performance. Mehta et al. (2012) conducted five studies and determined a moderate level (70 dBA) of ambient background noise versus low (50 dBA) and high (85 dBA) levels had a positive effect on creativity. The authors suggested at moderate intensity levels, processing difficulty and construal increased, requiring a higher level of abstract processing. At the higher intensity levels, the loud noise became detrimental by overwhelming the cognitive processing ability (Mehta et al., 2012).

In addition to intensity, the duration of exposure to noise can also influence performance. Short durational noise (combined with loud noise) has been found to be the most cognitively taxing, and thus detrimental to humans in performing tasks (Becker et al., 1995; Molesworth et al., 2015). This was further illustrated by Szalma and Hancock (2011), who found exposure to intermittent noise to be more detrimental to human performance than continuous noise. The authors posit this primarily to be caused due to individuals' limited ability to develop a response to the changing sounds (i.e., insufficient time; Szalma & Hancock, 2011). Hence, internal psychological or physiological adaptations of the individual to external stressors, can help in moderating performance. This will be further described in the next chapter.

2.3 Visual Cueing

2.3.1 Vision and the Visual Cortex

A visual stimulus is perceived by the human eye as electromagnetic energy (Wickens et al., 2004). For human perception, memory, and decision-making, vision is considered to be the dominant sensory cue involved (Posner et al., 1976; Koppen & Spence, 2007a). In determining if light can be adequately perceived, intensity, luminance and illumination are used to describe the quality of light. While contrast refers to the brightness of a target to its background or surround, illuminance involves the decline of reflected light from the luminous source depending on the background colour, absorption properties of the object, and the physical distance from where it is being viewed (Wickens et al., 2004). Changes to the wavelength energy also allow for the perception of colour by the visual cortex (Hawkins & Orlady, 2017). One of the main components of the eye is the iris, which controls and adjusts the pupil to accommodate the amount of light entering the eye (Hughes, 2004; Wickens et al., 2004; Snell & Lemp, 2013; Hawkins & Orlady, 2017).

To adjust focus to the incoming light, the ciliary muscle moves the eye lens, while the retina, as a layer comprises receptor cells (rods and cones) at the back of the eye, receives this light (Snell & Lemp, 2013). These receptor cells are utilised under different light conditions to provide visual acuity (i.e., rods in low light, cones in adequate light) and transform the electromagnetic energy in the form of light, into neural energy. From here, the optic nerve provides a passage to the brain which processes the neural energy into information, commonly referred to as vision. Binocular vergence changes the convergence of each eye to see objects clearly and at different distances, allowing for depth perception, and importantly facilitating the perception of 3-Dimensional physical space (Wickens et al., 2004; Orban et al., 2006; Parker, 2007; Hawkins & Orlady, 2017).

Visual perception can be influenced by combining other senses (i.e., bottom-up processing), utilising the eyes, the vestibular system for balance, and the brain, however it can also be influenced by past experiences, expectations, and emotions (i.e., top-down processing; Hawkins & Orlady, 2017). In perceiving depth and distance, visual cueing involving binocular vision, perspective, apparent movement of objects, relative size, and texture gradients, are utilised. Inside a conventional aircraft cockpit, pilots extract information visually from flight instruments, and when available, utilise binocular vision and FOV to perceive depth and motion outside their aircraft (Howe, 2017). However, when the latter is not possible (e.g., night or under instrument meteorological conditions), some aircraft are fitted with advanced instrumentation that attempts to reproduce such information. For remote pilots, limited visual feedback under BVLOS operating conditions can produce perceptually deceiving imagery (i.e., visual illusion), in terms of control alignment, which are most dangerous during the take-off and landing stages of flight

(Shively et al., 2015). Visual illusions are defined by the American Psychological Associated as "the misperception of external visual stimuli that occurs as result of a misinterpretation of the stimuli" (VandenBos, 2015, Vv section). Some examples of visual illusions include erroneous depth perception or spatial disorientation at night (Gibb, 2007; Gibb et al., 2008), and as suggested by their definition, can remain undetected (Hawkins & Orlady, 2017).

2.3.2 Visual Stimuli, Illusions and Display Types

Inside a traditional aircraft cockpit, the pilots extract information visually from the flight instruments and/or technology onboard, and when possible, utilise binocular vision and FOV to perceive depth and motion in the space outside the aircraft (Howe, 2017). For RPAS, the latter is not possible. Despite access to this information, incidents and accidents continue to occur. The reduction or degradation of visual sensory cueing can be a significant problem for conventional pilots (e.g., at night or in low visibility). This problem is likely to extend to remote pilots also.

Grosz et al. (1995) describe the importance of certain visual information sources used by pilots when flying. When no clouds are present, optic flow of the expanding horizon (with forward movement of the observer inside the aircraft), and perceived relative movement of peripheral ground objects, can provide valuable information about the trajectory of the aircraft, particularly during the approach to a runway (Grosz et al., 1995). However, when the ground is featureless and at night with low-density lighting, a misperception of height is possible (Calvert, 1954; Gibb, 2007; Gibb et al., 2008).

As stated by Tredici (1980), a lack of texture of a background or object can provide an illusory effect of false height, due to a lack of vertical references (i.e., appearing to be higher above the ground than actual). From the air, such occurrences are common when flying over calm water, flat deserts, snow covered terrain and at night with low density ground illumination features (Tredici, 1980). Previc (2004) further detailed this by describing the additional effect of aircraft speed, likening high and fast, and low and slow flights, as appearing visually similar to a pilot onboard. Similarly, water or reflections on windshields, sloping terrain, and adverse weather conditions can also adversely affect depth perception (Hawkins & Orlady, 2017).

Missing or reduced visual stimuli can also induce illusions. Visual illusions can be induced via stationary and/or moving stimuli (Kawabata, 1976), and include misperceptions of depth, contrast or movement. Pinna (2013) notes a visual illusion involves a mismatch, where "the necessary condition for the occurrence and perception of an illusion is the discovery of this mismatch" (p. 318). As described by Coren and Girgus (1978), the word illusion comes from a Latin word meaning "to mock" (p. 2). Because vision has been previously described as the dominant sensory system in humans, this is an appropriate description, particularly because illusions can be misleading, deceptive and/or confusing (Mukerji, 1957; Pinna, 2013).

Within the context of conventionally piloted aircraft, some incidents and accidents have been attributed, in part to visual illusions. As stated by Tredici (1980), pilots are trained to overcome visual illusions by relying on their instruments, however "sometimes during stressful situations this is forgotten, and the pilot erroneously goes back to his basic physiologic sensors, with at times disastrous results" (p. B5-4). Such illusions can appear during the approach stage of flight, before the landing task (Kraft, 1978; Tredici, 1980). Commonly, these can occur when the pilot dismisses correct information from cockpit instrumentation, and instead relies on their own visual judgement. Other illusions relevant to all pilots can occur when visual and additional systems interact (Hawkins & Orlady, 2017).

For remote pilots, visual stimuli are presented via the utilisation of different hardware configurations in terms of visual display type. Like conventional pilots, visual stimuli can also be lacking (i.e., at night, over water). Under BVLOS conditions, the remote pilot cannot directly see their remotely located aircraft and hence are reliant on receiving visual information via on-board camera/s. Presently, there are two available visual options: a monitor/tablet display (i.e., like those used for a personal computer) and first-person-view (FPV) goggles. The FOV of the remote pilot is restricted to that of the RPA onboard camera, and may reduce their awareness to surrounding airspace. However, Balog et al. (2017) suggest there are benefits to be gained by utilising this technology. Expanding the visual sources could alleviate the burden of currently limited visual feedback for BVLOS operations.

Stevenson et al. (2015b) assessed alternate visual display methods for small RPA comparing the manual control accuracy of landing a fixed-wing RPA under three visual conditions: FPV using goggles, FPV using a monitor display, and remotely viewed under VLOS. While a significant limitation of the experiment was the very small sample size (n=2; a possible reflection of the regulatory restrictions, and associated impact on participant recruitment), the results led the authors to suggest a dissonance in performance depending on previous flying experience. The conventional pilot with flight simulator experience performed better under the FPV goggles condition. Conversely, accuracy performance under the FPV goggles condition was reduced for the remote pilot who had no previous traditional flying experience or experience using flight simulators. It is difficult to determine conclusively whether previous flying experience or visual display type had the greater influence on task performance (i.e., accuracy). However, it is a valuable early comparison of alternate display types, and will be measured more robustly, in addition to different participant groups, in the three experiments to be presented.

One advantage of FPV goggles is the potential to eliminate surrounding visual distractions (Ruiz et al., 2015a). For example, situational awareness in an inspection environment, where an operator on the ground can assess construction sites or infrastructure in real time and without distraction can be improved (Paes et al., 2017). This is in comparison to a tablet device or monitor display showing the same FPV live image projected onto a 2D screen, where the visual environment around the display is also visible, possibly containing distracting or conflicting visual information. Hence, the stage of flight, the type of task, the operating environment and the visual information presented, will determine the usefulness, or not, of FPV goggles as a sufficient display type.

Some limitations do exist for FPV goggles, however. They may negatively influence task performance if anatomical movement of the remote pilot including control inputs are not replicated virtually. While this could be considered a secondary source of visual cueing, the importance for an individual to visually distinguish their own biomechanical response to movement was demonstrated by Prablanc et al. (1979). Participant accuracy, who were tasked with pointing their index finger towards randomly projected objects on a screen was measured. Accuracy reduced when the visual cue of seeing one's own finger, and its directional position relative to the projected object was removed. In reference to remote pilots, the absence of sighting physical body movements may be a further limitation presently associated with FPV goggles and BVLOS operations, particularly where manual control inputs are required.

2.3.3 Visual Information and RPAS

Remote pilots must rely on cameras and sensors relaying the visual information back to the visual display, to make sense of the scene. The limited visual cues available under BVLOS operating conditions can produce perceptual illusions in terms of control alignment, which are most dangerous during the take-off and landing stages of flight (Shively et al., 2015). As has been previously described, these visual operating conditions may also contribute to spatial disorientation at night and over water (Self et al., 2006). Some implications of such visual illusions include: failure to sight traffic, failure to notice changes in weather, and failure to notice terrain and obstructions (Howe, 2017). Extremes of such cases were seen in 2010 when a military RPA crashed in Kandahar, Afghanistan. The accident was attributed to operator error. However, in part due to the limited visual cueing available, the operator did not realise she had been flying the aircraft upside down, which lead to the crash (Whitlock, 2014a; Whitlock, 2014b).

The importance of camera view as a source of visual feedback to remote pilots has been further emphasised in several accidents involving other military RPAS. The United States Air Force (USAF) lost an MQ-1B Predator in 2012 following a crash during take-off, when the crew relied upon the same camera source of transmission (USAF, 2012b).

The remote pilot and sensor operator were each using the camera to view the runway during the take-off roll. However, at a speed of 61 knots, the single camera relaying the visual information rotated upwards 20 degrees (i.e., automatically, without remote pilot input), resulting in a loss of forward vision. The crew initiated an aborted take-off, reducing the throttle and disabling the link, an act which subsequently resulted in a ‘lost-link’ failsafe action between the GCS and the RPA, hence overriding the remote pilot’s throttle and brake inputs. The RPA automatically increased the throttle setting and pitched up, but without sufficient airspeed, and hit a barrier before lifting off (USAF, 2012b). While there several factors contributed to this accident, the reliance of the operating crew upon a single visual source, that when failed triggered a series of negative events, could have been mitigated by a redundant system.

In another accident, also involving an USAF MQ-1B Predator in 2015, a failure by the remote pilot to correctly visually identify a convective weather system led to a loss in aircraft control following a suspected lightning strike (USAF, 2016). As noted by the USAF report,

> A factor substantially contributing to the mishap was the Mishap Pilot’s (MP) misidentification of adverse weather. The MP misidentified a layer of clouds directly ahead and above the MRPA’s (i.e., Mishap RPA) altitude as haze and continue flight toward the area of developing weather. As the MRPA flew under the cloud layer, the MP did not notice further indications of developing weather, such as increasing cloud cover, virga (precipitation that evaporates before reaching the ground) or lightning, and continued flight below the base of the developing

thunderstorm, where the MRPA was subsequently affected by lightning. (USAF, 2016, Executive Summary section)

Although the forecast weather, obtained almost seven hours prior to the accident indicated favourable conditions, there were several visual cues signifying the deteriorating weather conditions. Despite this, the report indicated there was “no evidence to suggest human factors were a factor in the mishap” (USAF, 2016, p. 9). A pilot onboard may have experienced other sensory cues related to the expected vicinity turbulence associated with thunderstorms. Furthermore, it is highly probable the visual display used by the remote pilot to obtain 2-Dimensional visual information transmitted from the onboard camera, was inadequate to provide depth perception, a necessary perspective for accurately determining the volumetric characteristics of cloud type, height, and extent (Narasimhan & Nayar, 2002).

A further limitation associated with onboard cameras used to relay visual information from the RPA, is the FOV available to the remote pilot, and its relationship with spatial awareness (AAIB, 2020b). Like modern digital cameras, a zoom function is also available for utilisation by the remote pilot on the ground, on some RPAS models. While this can enable the magnification of the visual scene, without the need of the RPA to change its physical proximity, it can also limit the visual information received. At a construction site in 2020 in the United Kingdom (UK), a multi-rotor RPA collided with a tower structure while conducting an aerial filming operation. The camera mode was changed to obtain higher quality imagery; however this also changed the FOV relayed to the remote pilot. The AAIB determined the accident was caused by the changing FOV, and subsequent incorrectly perceived visual information by the remote pilot, including real distances between the RPA and other obstacles. Despite being alerted by a ground observer to the potential of a collision, the remote pilot ignored the threat. As concluded by the AAIB, the remote pilot "disregarded this information because it did not conform to his mental model of the situation" (AAIB, 2020b, p. 107).

A further example of visual misperception was demonstrated in 2017, when a civil-operated, fixed-wing RPA was performing an automated aerial survey using pre-programmed waypoints (AAIB, 2017). Unlike some conventionally piloted aircraft with robust autopilot navigation systems, the automation available in this flight did not account for changing wind speed and direction. As such, several overshoots of downwind waypoints were performed which brought the RPA into close proximity of peripheral obstacles in an adjacent construction site (i.e., tower cranes). When interviewed by the AAIB (UK) investigators, the remote pilot stated that no manual intervention to the flight was performed prior to the RPA collision with the cranes, due to an underappreciation of the significance of two visual aspects: the proximity of the RPA to the cranes could not be easily ascertained, and the height of the cranes was difficult to determine from their vantage point (AAIB, 2017). The problems associated with judging obstacle height visually from the ground was also shown in a collision between a multi-rotor RPA and a wind turbine blade (AAIB, 2020a). The remote pilot was aware of the obstacles and planned to fly at the maximum 400 feet above ground level (AGL) to maintain an adequate separation with other airspace users. However, the remote pilot incorrectly assumed the blades to be almost 100 feet AGL lower than they were, having utilised an internet search rather than referencing the correct aeronautical chart (AAIB, 2020a). Again, the visual feedback relied upon and the remote pilot's mental model of the operating area, were seemingly incongruent.

The detrimental effects of degraded visual cueing available to remote pilots is also illustrated by accidents involving mid-air collisions between RPA and conventionally piloted aircraft and helicopters (AAIB, 2015; AAIB, 2016a; NTSB, 2017; MOT, 2018; STSB, 2019; TSB, 2021). In 2017, a remote pilot operating a small multi-rotor RPA recreationally (and with no prior aviation experience) at a beach in Staten Island, New York, was involved in a mid-air collision with a helicopter flying at 300 feet AGL (NTSB, 2017). The remote pilot was operating the RPA below the maximum legal height for civil RPAS of 400 feet AGL at the time of the collision, but illegally under BVLOS conditions, and at night. Following an investigation by the NTSB (National Transportation Safety Board), it was further determined the pilot was unaware of the temporary flight restrictions in the area. The pilot was unaware a helicopter could be flying at a height below 400 feet AGL, and solely relied on the manufacturer's software for determining if the airspace was safe to fly in (NTSB, 2017). In addition to this, the NTSB suggested relying solely on the FPV camera view and manufacturer's app display (i.e., DJI GO4), was insufficient in maintaining a see-and-avoid watch of other airspace users, a common principle in traditional aviation (NTSB, 2017).

These examples highlight some of the limitations associated with the current level of visual imagery presented to remote pilots via onboard cameras. While the quality in terms of visual definition, FOV and depth perception can be improved in line with technological progression, and quantity can be attained by simply adding extra visual sensors at redundant locations, the remote pilot remains in a sensory deprived environment.

Furthermore, and like their conventional counterparts, remote pilots remain susceptible to visual illusions. A seemingly obvious way to overcome some of these concerns would be to provide additional sensory cueing like real-time auditory feedback, however this approach is not without complexities, as will be next be discussed.

2.4 Multimodal Cueing

The integration of multiple sensory signals to estimate and experience the surrounding world, is an important function of the human brain (Stein, 1998; Hillis et al., 2002). The sensing process of external stimuli is described by Clark and Yuille (1990) as the method for obtaining measurements of "structures in the world or environment and their configuration" (p. 3). By perceiving multiple sources of sensory information simultaneously, useful meaning about one's environment can be drawn (Clark & Yuille, 1990). As defined by Maragos et al. (2008), the perception process involves several stages where the filtering of information occurs, and an understanding of the experienced sensation is formed. Following the perception and filtering of multiple sensations, cognitive processes facilitate the performance of critical tasks including comprehension, learning, memory, and decision-making (Maragos et al., 2008).

As opposed to unimodal cueing, such as single auditory or visual stimuli presented in isolation, multimodal cueing involves two or more sensory cues. These can be provided intramodally (e.g. two or more visual cues) or intermodally (e.g. combined auditory and visual cueing; Hillis et al., 2002). Human interactions with their environment requires the processing of multiple senses simultaneously, however, where the sensory of information is incongruous (e.g., temporally and/or spatially disparate; Alais et al., 2010), initial perception can be adversely affected (McGurk & MacDonald, 1976). Furthermore, multiple sensory sources could become problematic or overwhelming when a defined level of task performance is required under higher workload conditions, or where the finite cognitive resources available to do so are preoccupied (Wickens, 2002). The uncertainty that may be associated with separate cues requires a degree of inference to be made by the individual (i.e., from previous experience), with equal or different weightings applied to each sense, and dependent on the context (Ernst & Bülthoff, 2004; Deroy et al., 2016). Hence, determining the source of an auditory signal to a co-located or independent visual source, requires different weightings placed on each sensory input (e.g., finding the location of a bird from its call at night versus daytime; weighted towards the auditory signal at night and visual signal in the day; Deroy et al., 2016).

Previously, Clark and Yuille (1990) described the separate modules involved with sensory processing, applying to each of the senses (e.g., auditory or visual system), and the potential benefits of receiving multiple cues. The authors argued that although each sense is processed separately (i.e., modular), the sensory information is 'fused' during the processing stage, and can be helpful in providing a level of redundancy. For example, perceived visual illusions due to an incorrect assumption by the visual processing system could be corrected via auditory or vestibular systems (Clark & Yuille, 1990; White & Hancock, 2020). A common example of this is described by Ernst and Bülthoff (2004) and the ambiguity associated with the perceived motion of another moving object, relative to the stationary observer (e.g., a stationary train being passed by another moving train). By utilising and combining additional senses (e.g., vestibular), the brain collects this information and uses it to resolve the ambiguity (Ernst & Bülthoff, 2004).

More recently, Alais et al. (2010) explained the transition in view of separate unisensory or modular processing of single senses, to a multisensory interaction between the senses, occurring earlier in the processing stage. Similarly, Van der Stoep et al. (2017) explain a multisensory integration mechanism that combines and processes multiple senses simultaneously. Importantly, the authors suggest that not all sensory information sources, provided together, are perceived or experienced equally (or at all). However, a temporal and spatial alignment of the stimuli is necessary for the information to become meaningful (Van der Stoep et al., 2017). Hence, when exposed to multisensory stimuli in a sensory deprived environment such as those involving teleoperations, consideration needs to be given to the relevance of additional sensory cues being added or removed, and what their effect on performance may be.

So far in this chapter, auditory and visual information/feedback have each been discussed separately. While additional sources of sensory information exist in addition to those already described (i.e., vestibular, olfactory, etc.), and their combinations can also be referred to as multisensory or multimodal, these other pairings are beyond the scope of this research. Hence, for the remaining sections of this and subsequent chapters, unless specified, all references to multimodal/multisensory cueing will refer to the auditory and visual pairing.

2.4.1 Auditory and Visual Cueing

In aviation and other complex working environments, multimodal cueing is used to convey important messages or warnings to the human about systems, and the state of the operating environment (Baldwin et al., 2012). The addition of auditory cueing to a visual search task in dynamic environments has shown to be assistive (McIntire et al., 2010; Hancock et al., 2013; Sun et al., 2022). It has also been shown to reduce visual distraction (i.e., fewer errors) in robotically-controlled movement, and thought to facilitate parallel processing through the provision of an alternate source of contextually similar information (Secoli et al., 2011). However, if poorly implemented, competing attentional and information-processing resources can affect the response of the human operator (Wickens, 1984; Wickens, 2002). In a driving simulation study by Ferris et al. (2006), participants were tasked to avoid roadside mines while driving a military vehicle, based on a visual warning (i.e., red and green LED lights). The findings revealed participants performed with a higher level of accuracy when an auditory cue was combined to the visual warning, with the authors concluding the importance of crossmodal (i.e., intermodal) cueing in the design of future vehicles and aircraft flight decks (Ferris et al., 2006). Interestingly, participants also subjectively rated the inclusion of the auditory component as annoying, despite their superior task performance in its presence. As will be discussed in the next chapter (Chapter 3), a link also exists between subjective workload and multimodal cueing, which can add further complexity to predicting task performance.

As specified by Wickens (2002), there are advantages associated with visual-auditory cueing. By negating intra-modal cueing such as visual-visual or auditory-auditory combinations, disadvantages (e.g., masking) may be overcome. However, the combination of both visual and auditory information presented together is not always positive. The Colavita Effect (Colavita, 1974; Spence, 2009) is a well-known bias towards visual information when audio-visual stimuli are present. Where the two cues exist together, visual perception is commonly enhanced by auditory information, but only when they are aligned in temporal and spatial occurrences (i.e., synchronous; Frassinetti et al., 2002). In cases where an aircraft pilot is spatially disoriented, the additional auditory feedback can be disadvantageous to task performance, by increasing the cognitive workload of the operator (Lewkowicz et al., 2018). When auditory stimulus is perceived to be presented before the visual stimulus, or when a delayed offset is perceived between the two cues, the Colavita Effect can be minimised or eliminated (Koppen & Spence, 2007b). Similarly, when more than two different types of cues are presented concurrently, visual dominance can also be eliminated (Hecht & Reiner, 2009).

In a study by Wada et al. (2003), the authors described the effects and interactions between auditory and visual cues. They measured the abilities of twelve participants to determine whether auditory cues were increasing or decreasing, visual cues were increasing or decreasing, and used combinations as well as ambiguous settings. The auditory and visual cues were each found to be dominant depending on the condition. These findings parallel the Sound-Induced Flash illusion, whereby a single flash of light accompanied by a series of auditory beeps, can induce the perception of multiple visual flashes (Shams et al., 2005; Nava & Pavani, 2013). Similarly, Hecht and Reiner (2009) tested participants under different combinations of auditory, visual, and haptic stimuli; sometimes two combinations at a time, or sometimes three. The authors elaborated on the Colavita Effect, describing a dominant bias towards the visual sense when auditory and visual stimuli were presented together. The authors concluded visual dominance to be usually found in bi-sensory tasks involving either auditory or haptic cues.

As suggested by Koppen and Spence (2007c), the Colavita effect can be overcome by spatially locating the auditory cue away from the visual cue. Within audio-visual interactions, the effects appear to be one-way; that is, a visual stimulus can be influenced by auditory cues, but not in the reciprocal direction. Conversely, Roach et al. (2006) proposed that this one-way direction is possible to change and occur in both directions when the relative sensitivities of each cue are also matched. This would seem to have implications in very specific circumstances, and caution should be taken with these results due to a very small sample size. Nevertheless, these examples present obvious consequences to the field of RPAS and remote pilots, where the availability or not of some sensory cues are further complicated by the physical separation between the aircraft and human operator. Perceptual errors could be problematic for remote pilots if they have not been sufficiently exposed to, and/or experienced the visual flying conditions in person, or from inside a conventional aircraft cockpit.

Temporal and spatial interactions are described by Alais et al. (2010) as factors influencing the perception of audiovisual information. In terms of spatial colocation or separation, the ventriloquist effect is used to demonstrate the dominance of visual versus auditory feedback, and when the signal-to-noise ratio is low (Choe et al., 1975; Slutsky & Recanzone, 2001; Alais et al., 2010). The ventriloquist effect takes its name from ventriloquism, whereby to an observer, the location of speech sound comes from an apparent visual source (i.e., the non-speaking puppet). It can be also measured with simple visual and auditory stimuli when presented at disparate temporal and spatial locations (Jack & Thurlow, 1973; Slutsky & Recanzone, 2001). As demonstrated by Alais and Burr (2004), although previous studies of the effect (Pick et al., 1969; Mateeff et al., 1985) suggested the visual system to dominate the auditory, the reverse can also occur when the visual stimuli is blurred (i.e., increased signal noise). Moreover, when the visual stimulus is partially degraded (e.g., slightly blurry), the authors argue the processing that occurs is not dominated by one sense over the other, but rather equally, via an optimal pairing (Alais & Burr, 2004). It is therefore apparent, that perceptual errors are also contingent on the quality of the sensory source of information.

In a meta-analysis by Burke et al. (2006), forty-three studies of similar nature were described relating to the effects of visual-auditory and visual-tactile feedback in the form of force feedback or vibrations, on user performance as compared to visual-only stimulus. The authors explained the potentially negative effects of technology in terms of extra information and the potential for cognitive overload, with poor decision-making outcomes and a degraded level of situational awareness. However, when the extra information was provided in the form of additional cueing, advantages of adding the extra cues could be seen, including improved reaction time and overall performance scores of participants. Furthermore, there appeared to be no effect on error rate regardless of the additional cues being available or not. Audio-visual feedback improved performance when workload was low, but for high workload it was determined to have the opposite effect due to the two modalities being cognitively related (Burke et al., 2006; Sigrist et al., 2013). This conclusion is in contrast to other studies involving navigation in flight or driving simulators (Sigrist et al., 2013).

In a visual search and tracking task, Bronkhorst et al. (1996) found improved performance (i.e., total search time), and lower subjective workload ratings by participants, when a 3-Dimensional auditory alert was added to a visual radar representation of a target aircraft. As concluded by the authors, the synchronised auditory feedback, unburdened the visual processing channel (Bronkhorst et al., 1996). A similar study by Tannen et al. (2004) produced comparable results for workload in a flight simulator visual search task. Notably, head motion data was also captured, with the authors finding the additional auditory cues assisted the visual search strategy. This was substantiated by fewer head movements by participants (Tannen et al., 2004). Further evidence of the benefits of utilising multimodal displays (i.e., auditory and visual) on driving performance was found by Liu (2001) in a driving simulator. Participants using the multimodal display in a navigation task were found to perform more accurately, with faster response times, and lower subjective workload ratings, compared to the visual-only displays. Liu (2001) explained the visual-only display having a negative effect on participants' attentional resources, leading to less safe driving overall. These findings are further supportive of the usefulness of multimodality.

In an investigation of human performance limitations and user interface design in remotely operated robots by Chen et al. (2007), FOV was found to contribute to erroneous speed judgements by the human operators. Additionally, in the case of multiple cameras being used on the one device, attention switching (i.e., having to shift visual focus between different camera views), and change blindness (i.e., interpreting changes in different visual scenes) were also listed as negative effects. Depending on the camera and imagery transmission to the operator, depth perception, time delays and imagery frame rates were found to lead to degraded motion perception, as well as underestimation or overestimation of distances and sizes. However, auditory information can be usefully incorporated to address cognitive tunnelling associated with teleoperated and/or augmented visual displays (Olmos et al., 2000; Chen et al., 2007; Wickens et al., 2009; Dicke et al., 2013). In the context of RPAS, combined auditory-visual feedback could be beneficial under current ecological BVLOS operating conditions.

During time-constrained periods with high workload, access to multiple sources of sensory information could also be advantageous by providing redundant cues or feedback, and better facilitating cognitive processing (Oskarsson et al., 2012; Haas & van Erp, 2014; Huang et al., 2019; Marucci et al., 2021). Furthermore, multiple sources of sensory information provided by two or more different cue types can be beneficial in overcoming illusions (Brill et al., 2015; Cunio et al., 2019). However, the addition of auditory information to visual stimuli may not always be optimal, or useful (Deatherage, 1972; Ho & Spence, 2008). Similarly, if redundant visual cues (i.e., two or more, unimodal) are paired with auditory cues, performance can also be negatively affected (Lees et al., 2012). This is also possible if auditory alerts are communicated in the presence of background noise, where the background noise masks the alert (Baldwin et al., 2012; Cunio et al., 2019).

The weightings applied to interpreting equivalently presented multimodal stimuli has been previously described, however the threshold of intensities for two or more sensory information sources is not always of equal levels (Manjarrez et al., 2007). Stochastic resonance is a counterintuitive phenomenon (Söderlund et al., 2007), and is described as the application of a moderate intensity of acoustic noise to a weak signal (e.g., visual), in order to amplify its detection, and enhance information processing (Rausch et al., 2014; Van der Groen & Wenderoth, 2016). In a separate study by Manjarrez et al. (2007) a weak visual signal was enhanced by an acoustic noise; an effect that deteriorated when the visual signal was strong. This parallels the "inverse effectiveness" (Ross et al., 2007, p.1,147) of multisensory integration, in which visual stimuli enhanced the perception and comprehension of speech in a noisy environment. These examples reflect some of the nuances of multimodal stimuli on performance, and have an association with 'just noticeable differences' as will be discussed below.

2.4.2 Just Noticeable Differences

Weber's Law describes the detection of small changes in stimuli, where the degree of change in background stimuli divided by the background stimuli will be a constant, most notably, for the perceptual changes of sound and light (Shen, 2003). Subjectively observed changes have been tested by several studies, including visual changes in velocity, weight, auditorily perceived changes in time, changes in pressure, and differences in smell (Ekman, 1959; Ekman et al., 1961; Solomons, 1900; Stone & Bosley, 1965). These relative changes are sometimes referred to as the "just-noticeable-difference" (Shen, 2003, p. 241-242). In the context of RPAS, and the limited sensory cueing environment available to remote pilots relative to conventional pilots, being able to identify small changes in stimuli may be critical. Hence, the addition of other sensory cueing, including the combination of visual feedback with real-time auditory feedback, could provide enough information to the operator to perform tasks when the available stimuli is degraded or non-existent. A related example of this was found by McLane and Wierwille (1975) who measured driving performance in a driving simulator utilising auditory (vehicle sound) and haptic (motion cues, roll, yaw) feedback. Participant performance was compared under combinations of cueing, with either all cueing or only some removed, and considered previous examples involving aircraft simulators. Although no statistically significant findings were found from the auditory cueing, participants' abilities to maintain a certain speed reduced when the auditory cue was removed, suggesting the importance of this information in assessing correct equipment function.

In a separate study by Hulusic et al. (2008), the authors investigated the effect of related and unrelated sounds in perceiving the rendering quality of a visual, virtual scene. Related sounds were determined to induce an attention response in participants to the scene, where slight degradations in scene quality were more noticeable, compared with scenes that included unrelated sounds (Hulusic et al., 2008). It was suggested the unrelated sound was distracting (i.e., perceived as noise), where no noticeable changes to the visual rendering quality were observed. Hence, it was an effect that could be exploited in future computational processing requirements. Nevertheless, in terms of audio-visual cueing, for correct perception of the visual scene, the two senses need to be aligned and congruous. This effect is similarly reported by McGurk and McDonald (1976) in terms of speech perception, where the auditory component of syllables was incorrectly dubbed over the visual component of lip movement (McGurk & McDonald, 1976).

Real-time visual information from the RPA is typically the only sensory cue available to remote pilots at present. However, auditory information in the form of alerts or warnings that are received at the GCS, as opposed to real-time auditory feedback directly from the RPA motors or engines, are also common. While historically, real-time visual cueing has been provided to remote pilots, the possibility of introducing concurrent auditory stimuli (i.e., multimodal cueing) is dependent upon technological advances. There are examples of both advantageous and disadvantageous outcomes of including the two sensory information types, depending on the associated workload and the cognitive facilitation required. While the context of the research portrayed in this chapter has also included other fields such as automobiles and general remotely operated equipment, there are important crossovers relevant to RPAS. Furthermore, the effect of additional stressors like workload, its interaction with multisensory information, and how this influences human performance, will be discussed in the following chapter (Chapter 3).

2.5 Summary

This chapter has provided an overview of the theory relating to auditory, visual and multimodal feedback, relevant to remote pilots. The induced human physiological and psychological reactions in response to auditory and visual information and feedback have been widely studied, including the interactions between more than one cue type and their effects on task performance. From an aviation human factors research perspective, with more than a century of data and a collaborative international approach through ICAO, systems have been designed where the human is placed centrally in the aircraft system. The proliferation of cheap, accessible RPAS to civil operators presents new challenges given the scale and speed at which new systems are being designed, including the operations they are being used for. At present, the remote pilot must adapt to the system. This includes adapting to the limited or degraded sensory environment. In this context, the variability of the task, the equipment available, and the presence or not of auditory sensory cueing in addition to standard (i.e., ecological) visual feedback, present challenges to remote pilots.

The effects of added or removed sensory cueing on human performance, particularly for civil remote pilots remains an area of limited applied research. This is particularly the case in terms of the availability of real-time auditory feedback transmitted from the remotely located aircraft, directly to the remote pilot. Specifically, there are knowledge gaps which remain in terms of understanding how remote pilots can operate effectively within a sensory deprived environment. Moreover, there is limited applied research in how the effects of auditory and visual cueing can influence the performance of remote pilots under different workload levels.

In the following chapter (Chapter 3), an overview of mental workload will be presented. Furthermore, some relevant behavioural models of performance will be discussed, including the Yerkes-Dodson Law, Arousal Theory, and the Maximal Adaptability Model. These concepts will further contextualise the three experiment chapters, with application to remote pilots.

CHAPTER 3 – MENTAL WORKLOAD AND PERFORMANCE

3.1 Overview

This chapter will discuss mental workload, the ways it can be measured, and its relationship with arousal, adaptability, and performance applicable to remote pilots. To better predict task performance influenced by mental workload, behavioural models are considered. In terms of task performance, remote pilots typically rely on automation to assist mental workload regulation. While automation can facilitate cognitive resources to be more appropriately allocated to perception and decision-making tasks, negative effects associated with under arousal and an overreliance on the technology also exist, all of which will be discussed in this chapter.

3.2 Mental Workload

The definition of 'mental workload' appears to be nebulous (also referred to as cognitive workload, hereon referred to as workload interchangeably; Volante et al., 2016), and has been described by authors in different ways. Perhaps too succinctly, Estes (2015) defines workload as simply “the work done by the mental system*”* (p. 1,176). Wickens and Tsang (2015) define workload as the “relationship between the resources required to carry out a task and the resources available to, and hence supplied by, the operator” (p. 277). Nygren (1991) defines workload as a reflection of “the interaction between a particular individual and the demands imposed by a specific task“ (p. 18). Common to these definitions is the limited cognitive resources available. Hence, the following definition of mental workload is adopted for this thesis: the demand on cognitive resources while performing a task.

What is less ambiguous with workload is its effect on performance. Bainbridge and Dorneich (2009) explain stressors to be any source that leads to a deterioration in performance, and can include noise and workload (Gaillard, 1993). As a stressor, workload can induce stress (MacDonald, 2003). Importantly, while related, stress and stressors are distinguishable. Gaillard (1993) acknowledges the 'broadness' of stress as a concept and describes its meaning in relation to environmental demands, behavioural reactions under strain, a state of feeling, and the process in response to strain over time. As such, workload as a stressor, and even the smallest amount of it, can be seen as a source of stress.

Importantly, stressors do not affect individuals equally (Gaillard, 1993). Hence, adverse performance is not an automatic response. According to Hart and Staveland (1988), mental, physical, and temporal demands of a task can vary, and are often dependent on experience and training. Within an automated system, Hooey et al. (2017) contends the following four elements drive workload: environmental factors, the task itself, the equipment being operated, and the operator. This parallels the SHEL model (System, Hardware, Environment, Liveware) first proposed by Edwards (1972), acknowledging automation as a function used to assist the operator in performing the required task. This is particularly the case during periods of low and high workloads (Bruggen, 2015; Volante et al., 2016).

Task performance during transitions from low to high and high to low workloads, has been shown to diminish (Cox-Fuenzalida, 2007). High mental workload can potentially impair performance via a fatigue response (Fan & Smith, 2017). The expected changes in performance induced by workload, either low or high (Brookhuis & de Waard, 2010), can also relate to behavioural mechanisms such as boredom (Smith, 1981), and hypo/hyper-arousal (Dehais et al., 2020). Borghini et al. (2014) elaborate further on the relationship between mental workload and performance, by describing an additional factor of situational awareness, defined by Endsley (1988) and Stanton et al. (2001) as the perception, comprehension and projection of a dynamic environment. As mental workload increases, situational awareness decreases, further decreasing performance (Borghini et al., 2014).

A closely associated construct to mental workload, is working memory, defined by Arlinger and colleagues (2009) as "the ability to simultaneously store and process information over a short period of time" (p. 373). It requires short term memory utilising sensory subsystems (e.g., auditory and visual; Baddeley & Hitch, 1974) and the storage and processing of information simultaneously (Baddeley, 1992). It assists the relationship between sensory inputs, information processing, cognitive load and task performance (Norman, 2013). When mental workload is high, the functioning of these processes can become impeded, leading to performance detriments (Hockey, 1997; Cox-Fuenzalida, 2007; Groemer et al., 2010), as evidenced by previous aviation accidents such as discussed below.

During an approach to Blackbushe Airport, UK, the pilot of an Embraer EMB-505 Phenom 300 (i.e., light business jet aircraft) encountered thirty-six traffic collision alerts (audible, automated voice) in a two-minute period during the downwind leg. The pilot had to manoeuvre to avoid a microlight and other light aircraft operating in the vicinity (AAIB, 2016b). Changes in trajectory requiring a climb in altitude occurred at a mentally demanding stage of flight, positioning the aircraft higher and faster than required for the final approach. Despite ineffectively deploying the speed brakes to slow the aircraft, it touched down 700m beyond the threshold, overrunning the runway, resulting in four fatalities. As determined by the investigation:

> Towards the end of the flight, a number of factors came together to create a very high workload situation for the pilot, to the extent that his mental capacity could have become saturated. His ability to take on new and critical information, and adapt his situational awareness, would have been impeded. In conjunction with audio overload and the mental stressors this can invoke, this may have lead (sic) him to become fixated on continuing the approach towards a short runway. (AAIB, 2016b, p. 45)

The fixation described in this accident could be described as a tunnelling effect, with attention limited to one channel of information (Li et al., 2022). This effect can be explained by insufficient residual cognitive resources induced by information overload (Wickens, 2005; Wickens & Alexander, 2009). The consequences of such an overload were further witnessed in the loss of Air France Flight 447. When the aircraft entered a dense region of convective weather, the accumulation of ice crystals blocked the pitot tube, thereby disabling the airspeed indicator. A series of events followed (e.g., instrument failures, autopilot and autothrust disconnection, audible alerts/warnings, and aerodynamic stall), from which the pilots were unable to recover. While degraded manual flying skills contributed to the accident (Haslbeck & Hoermann, 2016), increased mental workload shifted the pilots' attention away from responding to the aerodynamic stall. Seemingly, they did not notice they were even stalling (BEA, 2012). According to the final report:

> The crew never referred either to the stall warning or the buffet that they had likely felt. This prompts the question of whether the two co-pilots were aware that the aeroplane was in a stall situation. In fact the situation, with a high workload and multiple visual prompts, corresponds to a threshold in terms of being able to take into account an unusual aural warning. In an aural environment that was already saturated by the C-chord warning, the possibility that the crew did not identify the stall warning cannot be ruled out. (BEA, 2012, p. 179)

Recall in the previous chapter, noise was described as unwanted auditory information that can negatively affect performance. However, in this accident example, due to an overwhelming situation and apparent saturation of cognitive resources, the audible stall warning (i.e., wanted information) could not be attended to. The failure to perceive/detect the stimuli is again representative of overwhelming mental workload, via attentional tunnelling.

RPAS operators are not immune from the above described effects. In 2012, a USAF MQ-9A in the Republic of Seychelles (USAF, 2012a) crashed shortly after take-off. An audible alert and visual warning message indicating an incorrect flap setting, led the remote pilot to incorrectly shut down the fuel supply to the engine. Assuming an engine failure, the remote pilots initiated an emergency landing and returned to the airport. However, given the urgency of the manoeuvre and associated high workload, the landing gear was not extended, crashing in a gear-up configuration. In the investigation report (USAF, 2012a), a contributing factor was determined to be "channelized attention" (p. 18), whereby "an overly tight focus of attention leading to the exclusion of comprehensive situational awareness" (p. 18). This was further supported by one of the remote pilot's testimony who explained they were only focused on their altitude, airspeed and range (USAF, 2012a).

Channelised attention was also found to be a contributing factor for another crash involving an MQ-9A in Nevada (USAF, 2015) during high workload stages of flight. An instructor and student remote pilot were performing a training mission comprising a series of take-offs and landings. During the mission, the student remote pilot failed to maintain the required ground track, air speed and altitude, forcing the instructor pilot to make corrections. On the tenth approach, after the student pilot initiated a descent, the instructor also directed a descent, doubling the rate. Because of the increased descent rate, the RPA became obscured to the GCS by terrain, degrading the visual signal quality. By the time a climb was initiated, the RPA impacted with the ground (i.e., CFIT; USAF, 2015).

Despite many definitions of mental workload, its importance in safety-critical operating environments is evident by previous accidents. A challenge of measuring the effects of mental workload in applied settings relates to the ethics and consequences of creating conditions that may result in accidents (e.g., inducing fatigue, or high workload while driving on public roads; Brookhuis & de Waard, 2010). A consequence of the many definitions are the various ways in which mental workload can be measured objectively and subjectively (Brookhuis & de Waard, 2010; Tjolleng et al., 2017; Gao & Wang, 2020). The next section will discuss how workload is measured, focusing on objective and subjective methods.

3.2.1 Mental Workload Measurement

Objective measurement of mental workload can be performed using physiological responses and/or by proxy via task performance (Young et al., 2015). Physiological methods are advantageous by enabling real-time measurement, and as stated by Young et al. (2015), “are a natural type of workload index since work demands physiological activity by definition” (p. 4). Some measures include heart rate, respiration, skin resistivity, eye movement, and brain electrical activity via electroencephalogram (also referred to as ‘EEG’; Tao et al., 2019). For example, in response to workload changes, heart rate has been shown to vary (Jorna, 1993; Henelius et al., 2009; Mansikka et al., 2016).

In a driving study by Hidalgo-Muñoz et al. (2019) differences in breathing rate and heart rate were also found in response to varying workloads. Importantly, the effectiveness of the measures were relative to the task itself: heart rate was more sensitive to the combined workload-driving task, while breathing rate was sensitive to different workloads only. As Charles and Nixon (2019) explain, no one physiological parameter can be used to provide an overall measure of mental workload. Other potential shortcomings include their intrusiveness and cost of instrumentation (Azman et al., 2010; Knoll et al., 2011; Krigolson et al., 2017).

Objective workload assessment can also be determined via the performance of secondary tasks (Ogden et al., 1979). The type of tasks can require reaction/response time to a stimulus, memory, tracking, monitoring and mental arithmetic (Huddleston & Wilson, 1971; Owen, 1991). The *n*-back is one commonly used example (Kirchener, 1958), requiring working memory to remember a stimulus *n* times back (e.g., visually displayed numerals, appearing individually in sequence; Jaeggi et al., 2010). Another technique includes the Serial 3s and Serial 7s and involves a verbal count backwards from a pre-defined starting number, then a continued subtraction of 3 or 7 respectively from each previous number. While performance in secondary tasks can be used as an assessment of workload, some limitations do exist, such as practice or learning effects, individual intelligence quotient and/or previous education (Williams et al., 1996). Hence, workload can also be evaluated using alternate methods like subjective measures, which in many cases offer a feasible and more practical substitute.

Subjective ratings of mental workload comprise surveys and questionnaires (Young & Stanton, 2004; Xiao et al., 2005; Park & Jung, 2006). In experimental settings, these ratings are said to assess workload changes in response to manipulated independent variables (e.g., external stressors). However, unlike objective physiological measures which can assess responses to workload in real-time, subjective measures rely on measurement and analysis, post-event. Despite the potential shortcomings associated with this, and provided an acknowledgement is made, subjective measures remain a practical, cost-effective, and useful tool for mental workload assessment.

An example of a univariate subjective measure includes the Cooper-Harper Scale (also referred to as ‘modified’ i.e., from the ‘original’ Cooper-Harper Scale; Cooper & Harper, 1969; Mansikka et al., 2019). It is a useful measure for assessing overall mental workload, and presents a flow-chart/decision-tree to the user for determining one of ten possible ratings (Wierwille & Casali, 1983; Kilmer et al., 1988; Cummings et al., 2006). Another univariate measure is the Rating Scale for Mental Effort (RSME), consisting of a vertically-aligned scale with nine possible descriptors (i.e., from ‘no effort’ to ‘extreme effort’), presented at spaced, unequal intervals (Verwey & Veltman, 1996; Widyanti et al., 2013a). Its advantages are simplicity, and short duration to complete (Ghanbary Sartang et al., 2016). However, a limitation of the RSME addressed by alternative multivariate measures, is the ability to determine subfactors contributing to mental workload (Wierwille & Eggemeier, 1993).

An alternate to the RMSE is the Subjective Workload Assessment Technique (SWAT; Reid & Nygren, 1988). SWAT is a multidimensional measure, rating the loadings of time, mental effort, and psychological stress from 1 (low) to 3 (high) (Dey & Mann, 2010). It’s validity has been questioned by Boyd (1983) who argued the dimensions used should be further developed, and queried peoples’ ability to accurately assess their own workload. However, others have since found the test to be reliable and valid (Rubio et al., 2004; Xiao et al., 2005). A limitation of this method is the requirement for memory recall, and the time it takes to complete and score (i.e., estimated to exceed 45-minutes; Widyanti et al., 2013a).

A further example is the Workload Profile (WP) measure (Tsang & Velazquez, 1996) which compares the mental workload between tasks of different difficulties. The assessment measures eight dimensions of workload: perceptual/central processing, response, manual output, speech output, and spatial, verbal, visual and auditory processing (Tsang & Velazquez, 1996; Rubio et al, 2004). It is considered a reliable test with moderate face and convergent validity (Longo & Orru, 2019), and concurrent validity with task performance (Tsang & Velazquez, 1996). Despite the disadvantage of increased complexity, multivariate methods are advantageous for exploring workload multidimensionality.

Another widely used assessment measure is the NASA (i.e., National Aeronautics and Space Administration) Task Load Index (NASA-TLX, see Appendix 2; Hart & Staveland, 1988). This questionnaire measures perceived workload via six subcategories: mental demand (i.e., how mentally demanding the task was), physical demand (i.e., how physically demanding the task was), temporal demand (i.e., how hurried or rushed the pace of the task was), performance (i.e., how successfully the task was thought to be completed), effort (i.e., how hard the work was required to maintain a level of performance), and frustration (i.e., how insecure, discouraged, irritated, stressed or annoyed the participant was in completing the task). Possible scores for each category range from 'Very Low' being equal to 0 and 'Very High' being equal to 100, and is considered a useful method for predicting individual task performance (Rubio et al., 2004). It has been shown to have high correlation with performance, and also a very high convergent validity compared with the SWAT and WP measures (Rubio et al., 2004). Moreover, Hart and Staveland (1988), determined the NASA-TLX to be most valid for measuring subjective workload among participants, and having the highest user acceptance of four questionnaires reviewed.

The high user acceptance is also demonstrated by other studies utilising the NASA-TLX. Byers et al. (1988) assessed four subjective workload questionnaires administered to four crews operating within a simulated RPAS Ground Control Station (GCS). The authors found the NASA-TLX to have the highest validity and best user acceptance compared with the Modified Cooper-Harper, Overall Workload and SWAT assessments. Other studies incorporating the multivariate NASA-TLX, include tasks involving the supervision of multiple RPAS (Ruiz et al., 2015b), RPAS flight training (De la Torre et al., 2016; Albeaino et al., 2021), and virtual reality environments (Dell'Agnola et al., 2020).

One argument against subjective measures, however, is that humans estimating their own performance or ability can be inadequate (i.e., Dunning-Kruger effect; Kruger & Dunning, 1999; McIntosh et al., 2019; Wang et al., 2022). Furthermore, there is a requirement for memory recall post event, which does not compare with objective measurements, like physiological data, obtained in real-time. Subjective scales are also vulnerable to interpretation variances (Widyanti et al., 2013b), and have susceptibility to context bias (Colle & Reid, 1998). While these challenges/problems apply equally to workload scales, previous research has demonstrated correlation between objective (e.g., EEG) and subjective (e.g., survey) measures relating to workload (Zeier, 1994; Murata, 2005; Lin & Cai, 2009; Zhou et al., 2014). This includes within aviation settings (e.g., the assessment of air traffic controllers; Aricò et al., 2016), and in other fields such as healthcare (Mazur et al., 2013; Dye et al., 2017).

3.2.2 Workload and Automation

Automation is defined by Parasuraman and Riley (1997) "as the execution by a machine agent (usually a computer) of a function that was previously carried out by a human" (p. 231). Sheridan and Verplank (1978) originally described ten different levels of automation, in the context of teleoperated undersea equipment, relating to the degree of human interaction with the system. Kaber and Endsley (2004) similarly describe these levels for multi-task environments. Others have defined computer/machine involvement in the processing and filtering of information from multiple sources, decision-making, and task response or action (Sheridan, 2018; Kaber, 2018; Wickens, 2018). In multi-task environments, workload can be regulated by using automated systems. By removing some manual tasks, working memory capacity can be increased, allowing cognitive resources to be allocated to other tasks (Bainbridge & Dorneich, 2009).

When humans interact with complex systems, a mental model is developed, based on the available information (Borgman, 1986; Staggers & Norcio, 1993). However, if there is too much information to contend with, cognitive resources can become overwhelmed (Baddeley, 1986, as cited in Bainbridge & Dorneich, 2009). Hence, automation can be useful for improving efficiency, enhancing safety, and reducing mental demand, and the perceived workload associated with complex tasks (Lee & Seppelt, 2012; Wickens et al., 2021). While these benefits appear to offer a positive contribution to improving human performance, Harris et al. (1995) explain this logic as too simplistic. Put simply, "switching to automation becomes a source of demand. Automated tasks cannot be forgotten, but are metamorphosed into additional monitoring tasks. Even automated tasks have to be monitored, which can lead to an infinite regress" (Harris et al., 1995, p. 183). This criticism was exemplified in a monitoring task by Greenlee et al. (2018), whereby a decline in vigilance occurred via a failure to detect road hazards while operating automated vehicles. As the authors propose, sustained monitoring is demanding, increasing perceived workload and stress (Greenlee et al., 2018). Seemingly, the negative consequences of automation are also duration-dependent.

In modern aircraft, automation has enabled more efficient flight control, navigation and fuel management (Kantowitz & Campbell, 1996; Landry, 2009). However, for pilots, a number of negative issues were reported by Funk et al. (1999) including: inadequate understanding of the automated systems, unnoticed operation or function, overconfidence in the automation, inadequate training, poor design, negative changes to workload, and reduced job satisfaction. Furthermore, a reduction in situational awareness is possible (Dorneich et al., 2006; Onnasch et al., 2013). From an aviation perspective, these shortcomings are particularly relevant.

In 2009, a Bombardier DHC-8-400 twin-engine turboprop aircraft crashed five nautical miles short of the runway following an aerodynamic stall while flying an automated instrument approach (NTSB, 2010). In response, Geiselman et al. (2013) describe the automated system function caught the pilots by surprise, causing an incorrect manual control response. At a critical stage of the flight, the automation could have communicated the changing state of the aircraft with more deliberate information. Instead, it "stopped flying the aircraft, startled the crew, and did not afford them adequate time to respond confidently" (Geiselman et al., 2013, p. 23).

Another pertinent example occurred in 2013 when a Boeing 777-200ER operated by Asiana Airlines struck a seawall prior to landing at San Francisco International Airport. Although many contributing safety factors were recognised, a lack of automation understanding was acknowledged by the investigation. According to the NTSB (2014), the pilots failed to recognise how the autothrottle and autopilot modes interacted on approach; an incorrect response which led to the low energy state of the aircraft, and landing well short of the target zone. Furthermore, the airline's policy of full automation for line flying enabled an overreliance on the auto-systems. Had the pilot "been provided with more opportunity to manually fly the 777 during training, he would most likely have better used pitch trim, recognized that the airspeed was decaying, and taken the appropriate corrective action of adding power" (NTSB, 2014, p. xiii).

This recommendation is similarly provided by Casner et al. (2013) who tested pilot response to abnormal events (e.g., engine loss during take-off), and described the benefits towards recognising the changing situation, as opposed to relying on automated alerts only. An overreliance on automation shifts the human's central location, by decreasing the importance of their role, a situation highlighted by these accidents (Miller & Holley, 2017). Furthermore, the regulation of mental workload by automation can change states of perceived under/overwork.

From an RPAS perspective, automation assists remote pilots by managing their level of workload in different operational conditions, and allowing them to better focus on a primary task (Cummings et al., 2014; Dixon et al., 2005; Dixon & Wickens, 2006; Mekdeci & Cummings, 2009). However, unexpected situations and how these contribute to overall task performance is an evolving area of research (Cummings et al., 2014). Automation is relied upon for different missions, including navigation between waypoints. Hence, for large portions of RPAS flights, the remote pilot will have a monitoring role only. When automation increases workload by requiring vigilance from the human operator monitoring the system (Szalma et al., 2004), the situation is described as a paradox by Warm et al. (1996). As they suggest, "although automation is designed to reduce the workload of operators, it may place them at a functional disadvantage through understimulation" (Warm et al., 1996, p. 185). Furthermore, long periods of low workload requiring monitoring of highly automated systems can lead to low arousal, boredom and distraction (Cummings et al., 2013). Stimulation induced via arousal, and its potential effect on task performance will be further discussed in the next section.

As RPAS technology continues to develop with increased automation, human remote pilots could place too much trust in the human-machine system (Liu et al., 2015), and thus become further removed from it. Equally, this over reliance could be detrimental under high time pressure conditions. For operations involving variable workload, as described by Harris et al. (1995), "where workload periodically exceeds operator capacity, it should be possible to maintain operator supervision of all tasks during all operational phases by having operators invoke automation during high taskload periods" (p. 183). Hence, the use of automation should be task and operator dependent, reflective of the workload required (Harris et al., 1995). The use or not of automation is likely to change mental workload, and consequently performance. Behavioural responses to this variability via arousal and adaptation are also likely.

3.3 Workload and Performance

In this section, the Yerkes-Dodson Law (Yerkes & Dodson, 1908), Arousal Theory (Broadbent, 1978), and the Maximal Adaptability Model (Hancock & Warm, 1989) will be explained. While several models/theories exist, these theories are considered to be the leading ones that describe task performance via arousal or behavioural adaptations in response to workload and/or exposure to external stressors.

3.3.1 Yerkes-Dodson Law

The Yerkes-Dodson Law (Yerkes & Dodson, 1908) originated from habit-formation and stimulus strength experiments, involving mice learning a task in response to receiving increments of electric shock. It has been described as an inverted U-shaped curve of the relationship between arousal (i.e., response) and performance efficiency. As argued by Teigen (1994), the meaning of the law has evolved over decades, and has since been used to explain independent and dependent variable combinations including: stimulus strength and habit-formation, stress and learning, intensity of motivation and problem-solving, and level of arousal and quality of performance. A criticism of the law maintains no universal agreement on its definition, and that "its vicissitudes can be regarded as a demonstration of the mutability and fuzziness of allegedly 'basic' psychological concepts" (Teigen, 1994, p. 526).

Despite this criticism, its importance remains as a foundation for other since formulated theories, relating to human behaviour and predicted performance. In the context of mental workload, optimal performance is expected with moderate stimuli, decreasing as the stimuli becomes too low or too high. Furthermore, a relationship also exists between performance and other stressors, with responses differentiated between cognitively taxing and simple tasks.

3.3.2 Arousal and Performance

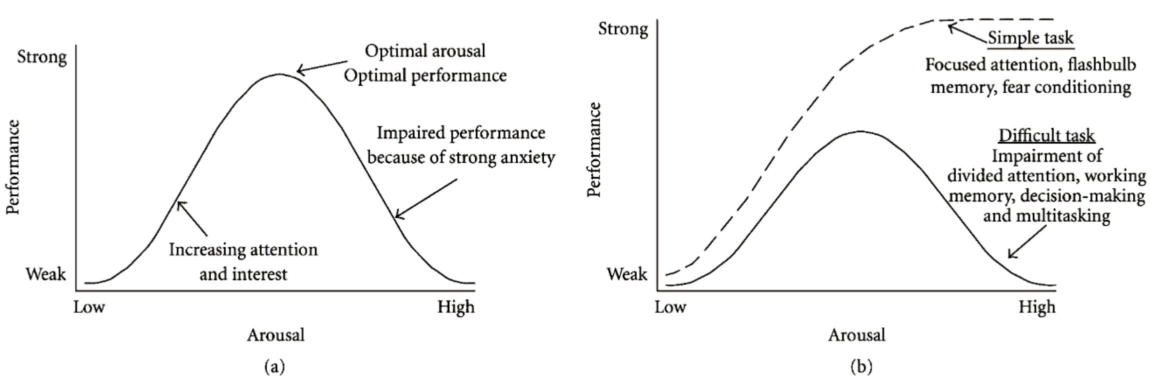
The construct of arousal has been described as activation, energy mobilisation and excitation, manifesting psychologically and physiologically (Duffy, 1957; Thayer, 1978; Neiss, 1988). Physiological arousal is differentiated from wakefulness, being “a state of heightened alertness/responsiveness to sensory inputs mediated by the ascending pathways, which is accompanied by an increase of physiological activity (postural tone, cardiac and respiratory rhythms)” (Adamantidis & De Lecea, 2008, p. 1,480). Arousal exists on a continuum ranging from low to high (Sanbonmatsu & Kardes, 1988), and like workload, can be measured objectively (e.g., biochemically, heart rate, skin conductivity, respiration rate; Lacey, 1959; Baum et al., 1982; Mehler et al., 2009; Reimer et al., 2009; Reimer & Mehler, 2011) or subjectively (Robertson et al., 2007; Rosebrock et al., 2017). The type of task can also elicit different physiological responses, depending on an acceptance or rejection of external environmental parameters, described by Lacey (1959) as an environmental intake or environmental rejection (Cacioppo et al. 1991). For example, visual detection tasks demand environmental intake while mental arithmetic or memory tasks require a degree of environmental rejection.

Tang and Harvey (2004) refer to the multidimensional aspect of arousal, and distinguish physiological (e.g., heart rate, blood pressure, respiration, perspiration) from psychological arousal (also referred to as ‘cognitive arousal’) relating to the activation of the mind. The effect of arousal on emotion and performance are described by Reeve (2009), whereby low and high arousal can lead to negative emotions (i.e., boredom and stress, respectively), each corresponding with lower performance.

In modelling the association between arousal and performance, the U-shaped and inverted Hebb curve can be used (Hebb, 1955). However, as Diamond et al. (2007) discuss, a hybridisation of this with the Yerkes-Dodson law has occurred, through misunderstanding or misrepresentation over many decades. Accordingly, a ‘Hebbian’ version of the original Yerkes-Dodson curve has evolved, further refining the original linear/curvilinear relationship between arousal and performance, by different settings: a simple task, or difficult/complex task (Broadhurst, 1959; Teigen, 1994; Diamond et al., 2007; Engineer et al., 2012). During a simple task, performance increases more slowly from low to moderate arousal levels, and declines more quickly when arousal increases from moderate to high. A comparison of both the original Yerkes-Dodson and Hebbian curves is shown in Figure 1.

Figure 1

A Comparison of the Hebbian Curve (a) and Yerkes-Dodson Curves (b)



Note. Figure reproduced with permission (Diamond et al., 2007).

This relationship was demonstrated by Arent and Landers (2003) for a simple reaction time task, where physiological arousal was manipulated while riding a cycle ergometer.

Conversely for tasks with a higher level of complexity, performance rapidly increases from a state of low arousal. Small incremental increases in arousal translate to higher performance more quickly than for easy tasks. Similarly, performance reaches an optimum point, although slightly lower than in a simple task, before gradually declining as arousal increases (Kahneman, 1973; Diamond et al., 2007). For difficult tasks, performance is predicted to be at an optimum level well before a moderate level of arousal is induced. This signifies for complex tasks, a moderate level of workload will not automatically deliver optimal results, hence the introduction of an additional stressor may be detrimental.

Like Hebb (1955), Broadbent (1978) describes Arousal Theory (AT), being linked with attention, and predicts when arousal is either low or high, the corresponding level of task performance is at a minimum. Hence, optimum performance exists between the two extremes of arousal. A stress-induced arousal response also impacts on workload, and can be used for motivation or effort to complete a task optimally. Arousal induced by workload has been linked to performance, and as exposure to external stressors increase, performance also increases to an optimum level before diminishing (Broadbent, 1978; Loewen & Suedfeld, 1992). For example, low arousal associated with driving a vehicle on a monotonous, low-demanding road, minimises performance ability (Thiffault & Bergeron, 2003). As arousal increases, the efficiency to perform a task increases to an optimal point, before retreating as arousal continues to rise. High arousal could manifest as a state of disorganisation, impairing processing ability, and becoming detrimental to performance (Dang & Tapus, 2013; Ünal et al., 2013). Hence, in a simplistic way, if the difficulty of a task can be modelled, the resulting performance of the task can be better predicted or anticipated.

In the context of sensory information, arousal can also be used to explain the effect of noise exposure on task performance (e.g., accuracy, speed), via induced attentional narrowing under cognitively demanding tasks. Arousal has also been described as one mechanism responsible for a beneficial effect of noise on visual information processing, particularly when the visual stimuli quality is degraded (Sanders, 1983; Keuss et al., 1990). For example, as described by Keuss et al. (1990), a driver interpreting a traffic sign in poor visibility may do so more effectively if noise (i.e., low to moderate intensity) is also present.

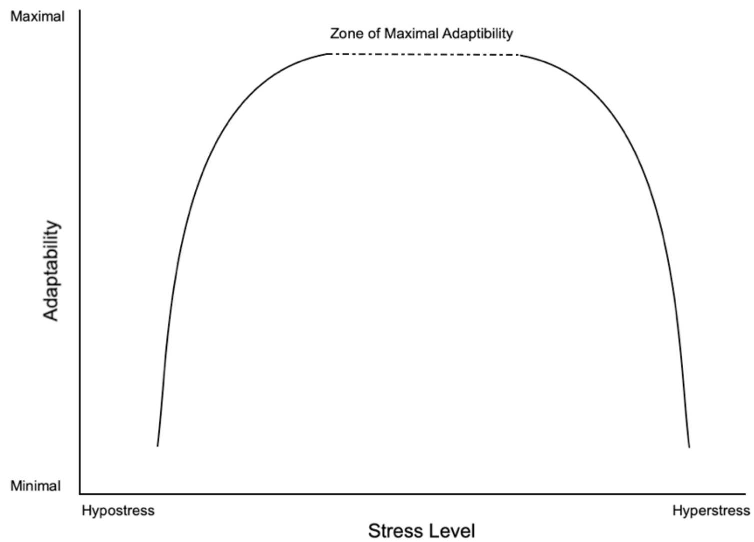
Alternate to Broadbent's (1978) AT, Poulton (1979) describes a composite theory involving arousal and masking. For auditory information, noise can mask inner speech and degrade task performance. Poulton notes that short duration noise should promote arousal and facilitate performance, while intermittent noise becomes more disruptive compared to a continuous source. Like AT, Poulton's theory suggests task accuracy to be more impaired than speed when exposed to noise (Poulton, 1978; Poulton, 1979). In addition to performance influenced by arousal, adaptational responses can also be elicited.

3.3.3 Adaptability and Performance

The Maximal Adaptability Model (MAM; also described as the 'Dynamic Adaptive Theory'; Hancock & Warm, 1989; Dehais et al., 2020) can be used to predict and explain performance. It maintains consistency with the previously described theories relating to arousal by attempting to define the location of optimal performance, explaining how other stressors impact on this function. The MAM describes how individuals generate an adaptational response to stressors via a physiological or psychological response (i.e., behavioural, also referred to as “attentional resource capacity”; Hancock, 2022, p. 4). At the extreme regions of under or overstress, the adaptation can break down due to limitations associated with cognitive processing capacity. These regions are described as unstable (i.e., zones of dynamic instability), but importantly the task itself can be considered a source of stress (Hancock, 1987; Hancock & Szalma, 2008). Also represented as an inverted U-shaped curve (see Figure 2), the MAM explains how an individual adapting to external stressors is at a minimum when these stressors are either low or high (i.e., hypostress or hyperstress). Adaptation ability is at a maximum in the plateau region between the two extremes, and considered as stable (i.e., the ‘comfort zone’; Hancock & Szalma, 2008).

Figure 2

The Maximal Adaptability Model Curve



Note. Adapted from the original maximal adaptability model of stress proposed by Hancock and Warm (1989).

The original and Hebbian Yerkes-Dodson curves, and the MAM are similarly shaped. Because of this, it is possible previous studies have misattributed findings to the various models (Hancock & Ganey, 2003; Doohan et al; 2022; Hancock, 2022). In response to one criticism of the model, Hancock (2022) argues:

> Experience with a stress, in general, provides some degree of defense against those self-same stress effects in relation to on-going task performance. Experience with a

> task (skill), in general, provides some degree of defense against task failure, especially under stress. These tendencies are both expression of adaptation of differing response systems of the body. The ability of the individual, either with respect to the external stress or the task they are undertaking, are critical dimensions which determine the degree of performance change witnessed. (p. 5)

One key difference between the MAM and other arousal models are its complexity, described as an ‘extended-U’ curve (i.e., differentiated from simple ‘inverted-U’), accommodating both physiological and psychological responses to stress (Hancock & Ganey, 2003; Hancock & Warm, 2003).

An example of a relevant external stressor includes the auditory cueing described in Chapter 2. Szalma and Hancock (2011)’s meta-analysis of studies involving noise and human performance found the MAM (Hancock & Warm, 1989) to best explain the adaptive response made by individuals exposed to external stressors, like noise. Hancock and Warm (1989) propose noise can overload the processing capacity via masking, distortion of the auditory task-relevant information, and/or a vocal response. Similarly, under zero absolute dBA conditions (i.e., the absence of noise), the removal of a sensory cue can be equally disrupting, by limiting the ability to generate an adaptational response to no cue. Behavioural models like the MAM can be considered as further means of explaining task performance results under multistressor conditions (Yerkes & Dodson, 1908; Broadbent, 1978; Poulton, 1979; Hancock & Warm, 1989; Szalma & Hancock, 2011). Recall in Chapter 2, noise as a stressor negatively influenced task performance via masking and/or attentional narrowing (Broadbent, 1978; Britton & Delay, 1989; Ke et al., 2021). While an arousal response due to an external stressor may increase, where performance also increases to an optimum level before diminishing (Broadbent 1978; Loewen & Suedfeld, 1992), the MAM outlines an alternate view between the relationship of noise and performance. Here, individual behavioural adaptations to external stressors are made.

The process of experiencing and responding to auditory stimuli from a MAM viewpoint, follows three stages: at the input stage (e.g., where exposure to acoustic noise is associated with the task), the adaptation stage, (e.g., where a psychological response to noise occurs and an adjustment to task performance takes place), and at the output stage (e.g., the response speed or accuracy of a given task; Hancock & Warm, 1989). Szalma and Hancock (2011) identified several features associated with the type of acoustic noise experienced, and its effect on task performance. Intermittent noise was found to be more damaging to performance than continuous noise. This seems logical as it was previously theorised that people can find ways to adapt to longer duration noise, and develop the appropriate responses (e.g., habituation; Namba & Kuwano, 1988; Banbury & Berry, 1997), as opposed to limited temporal exposure. Despite the potential for adaptation and/or habituation, irrelevant noise has also been shown to disrupt performance in a variety of tasks, irrespective of the meaningfulness of the stimuli presented (Banbury & Berry, 1998; Jahncke, 2012; Klatte et al., 2013). While short duration noise exposure could limit an adaptational response, it has been observed that an arousal response can also exist (Szalma & Hancock, 2011). Hence, a temporary increase to the information-processing resources available for completing a task via arousal rather than adaptation, may still remain possible.

In responding to stressors, the type of task being performed is also relevant for the MAM. For example, negative effects of noise are more pronounced for cognitive tasks than perceptual tasks, and for those requiring performance accuracy as opposed to reaction time or speed (Szalma & Hancock, 2011). During vigilance tasks (e.g., perception tasks), noise can facilitate performance via an increase in arousal, by increasing stimulation to an unstimulating task and/or environment. Conversely for cognitively demanding tasks, a reduction of resource capacity due to adaptive compensatory responses, may be incurred. If the intensity, duration, and content of the noise are dynamic in addition to the type of task, a resource-based approach could more effectively explain the effect of noise on performance. Namely, exposure to stress (i.e., a stressor from external stimuli, e.g., auditory feedback), and the cognitive demand of the task, determine the ability of the individual to form an adaptive response (Szalma & Hancock, 2011).

With respect to RPAS operations, the type of task being undertaken could be benefited by an adaptational response to variable sensory cueing, like dynamic real-time auditory feedback. Because exposure to limited sensory cues (e.g., audio-visual) are thought to result in increased workload (Scerra & Brill, 2012; Hancock et al., 2013), task performance can be affected (Dunn et al., 2020). Consequently, exposure to dynamic stressors like auditory feedback, combined with stage of flight, affects the ability to make an adaptational response (Lee & Liu, 2003; Di Nocera et al., 2007; Crognale & Krebs, 2011; Liu et al., 2016). Hence, because auditory information can be considered both noise (i.e., distraction/stressor; Kjellberg, 1990) or as a stimulant/arousing, the MAM is potentially more appropriate for predicting remote pilot performance under varying workloads.

3.4 Summary

This chapter has provided an overview of mental workload, behavioural models relating to arousal, adaptability, and task performance. The Yerkes-Dodson Law, Arousal Theory and Maximal Adaptability Model (MAM) were each described. Three models suggest human task performance can be determined by stress-induced arousal, which has an optimum level in facilitating performance before reducing. Alternatively, the MAM explains how task performance can be controlled by adapting to stress-induced arousal through physiological and/or psychological means, broadening the optimal region of peak task performance. For both the AT and MAM, a defined level of workload could also be considered an additional external stressor. While the AT can be used to predict performance via arousal, and for the MAM, adaptation to external stressors, it is less certain how this translates to dynamic real-world environments involving multistressor variables, like a sensory deprived applied RPAS context. With this background, three applied experiments comprising the research for this thesis/study will be presented in the next chapters. A broader discussion of the key findings and implications of the study to the field of RPAS, and by extension to other automated and teleoperated systems, will be provided in the final chapter.

CHAPTER 4 – EXPERIMENT 1: EFFECTS OF AUDITORY AND VISUAL FEEDBACK ON REMOTE PILOT FLYING PERFORMANCE

***Note. The contents of this chapter, with the addition of a brief introduction are published in the journal titled Ergonomics. Below is a reference for this publication.**

Dunn, M. J. M., Molesworth, B. R. C., Koo, T.; & Lodewijks, G., (2020). Effects of auditory and visual feedback on remote pilot manual flying performance. *Ergonomics*, 63. https://doi.org/10.1080/00140139.2020.1792561

4.1 Introduction

It was established in the first chapter that most civil and commercial RPAS operations are flown under visual-line-of-sight (VLOS). To recap, in such conditions, the remote pilot must be able to see his/her aircraft visually and unaided. However, with advancing technologies, Beyond VLOS (BVLOS) operations, which have traditionally been used in the military, are increasing within the civil and commercial contexts. There are known human factors implications evidenced by previous RPAS accidents, where reduced sensory cues and workload were identified as contributing factors. When two or more sensory cues are experienced simultaneously, specifically the presence of auditory and visual feedback, task performance can be enhanced or degraded depending on the presence (or not) of other stressors, such as workload. Furthermore, behavioural theories including the Maximal Adaptability Model (MAM; Hancock & Warm 1989; Szalma & Hancock 2011) can help explain human performance in response to external stressors.

Because auditory feedback can be perceived as both useful information and noise, behavioural adaptations in response to this stressor may be more nuanced when operating in a sensory deprived environment. Hence, the first experiment was carried out with the aim of determining the degree to which auditory feedback, visual display type, and workload experienced by the participant, influenced task performance. Specifically, the effectiveness of unimodal (visual) and multimodal (auditory & visual) cueing on remote pilot manual flying performance was measured. In pursuit of this aim, three different visual display types (unimodal) were utilised and presented with and without real-time auditory feedback (i.e., transmitted directly from the RPA), in the main experiment flying task. A workload component, namely the presence or absence of wind shear, was introduced in a sequential flying task to further understand the advantages and limitations of available sensory cueing. In all experimental conditions, task performance was defined by measuring the horizontal flying accuracy and vertical flying accuracy via photos taken from the RPA at defined points in space within the experimental environment. The time taken to complete each flying task was also recorded. As such, one research question and two hypotheses were considered.

4.1.1 Research question

How does unimodal (visual) and multimodal (auditory & visual) feedback change participant flying performance (accuracy and timeliness) under wind and no wind conditions?

4.1.2 Hypotheses

1. The real-time auditory feedback produced by the propellers will aid participant flying performance in terms of accuracy (horizontal and vertical deviation) and timeliness.
2. Participant flying performance (accuracy and timeliness) will be adversely affected by the presence of wind shear (i.e., higher workload) in all flying conditions.

These performance measures were considered central to the assessment of remote pilots' manual flying ability (CASA, 2017). Navigation and spotting tasks under different levels of audio-visual cueing and workload were used to measure participant performance while manually piloting a small multi-rotor RPA.

4.2 Method

4.2.1 Participants

Eighteen pilots (three female), all cleared to fly solo in a traditional fixed-wing aircraft were recruited from the University of New South Wales Sydney, Flying Operations Unit at Bankstown Airport, Sydney Australia. Participants held the following licenses: Recreational Pilot Licence (n = 1), Private Pilot Licence (n = 2), Commercial Pilot Licence (n = 14) and Airline Transport Pilot Licence (n = 1). Demographic information regarding participants can be found in Table 1.

Table 1

Demographic Information of the Sample Group

Descriptive Variable	M	SD	Range
Age (Years)	26.11	9.74	19 – 56
Total Flying Hours	1454.18	2697.78	70 – 10200
Total Flying Hours in Past 90 Days	51.28	55.98	0 – 200

Note. *M* refers to ‘Mean’. *SD* refers to ‘Standard Deviation’. N = 18.

All pilots held a Class 1 Medical certificate issued by the Civil Aviation Safety Authority (CASA), indicating they were in good physical and mental health. Except for one pilot who reported a total of ten hours, all other pilots recorded no previous RPAS experience. The research, including stimuli and procedure was approved in advance by the UNSW Sydney Ethics Panel (HC Number: HC180410). All participants signed a consent form prior to undertaking the experiment (Appendix 1). In addition, the research was conducted in accordance with CASR Part 101 (Unmanned aircraft and rockets) CASA Regulations. All participants were reimbursed for their time in the form of a $40 Coles/Myer gift voucher.

4.2.2 Design

Two experimental designs featured in the research. The first experimental design (Task A - Navigational task) comprised a 3 x 2 repeated measures groups design while the second (Task B - Spotting task) comprised a 3 x 2 x 2 mixed repeated measures between groups design. In both designs, the first repeated measures factor labelled Visual Cue contained three levels: VLOS (Control), BVLOS Monitor (BVLOS-M) and BVLOS Goggles (BVLOS-G). The second repeated measures factor labelled Auditory Cue contained two levels: with real-time auditory feedback from the RPA (Auditory) versus no auditory feedback from the RPA (Quiet, i.e., ambient sound). The between groups factor for the spotting task (Task B) labelled Wind, also contained two workload levels: no wind condition versus wind shear. Two distinct tasks, one navigation (Task A) and one spotting task (Task B), were created as it was not possible to create wind shear sufficiently large enough to cover the whole experimental area with the available resources. As a result, and to reduce a possible learning effect, the stimuli were presented in a Latin square counterbalanced design across all conditions. Task A was always followed by Task B; there were six counterbalanced combinations of Visual and Auditory Cues for Task A, and twelve counterbalanced combinations with the addition of the Wind component in Task B. Given the two different operational conditions (Task A vs. Task B), two separate 3 x 2 repeated measures analyses were performed. Dependent variables for each task were timeliness, horizontal deviation, and vertical deviation. A summary of the experimental design and corresponding result tables are presented in Table 2.

Table 2

Overview of the Experimental Design

<table>
<tr><th rowspan="2">Task</th><th rowspan="2">Dependent Variables</th><th colspan="3">Experimental Condition</th></tr>
<tr><td>Factor 1</td><td>Factor 2</td><td>Factor 3</td></tr>
<tr><td rowspan="6">Navigation (Task A)</td><td rowspan="6">Horizontal Deviation
Vertical Deviation
Timeliness
(Table 3)</td><td rowspan="2">VLOS (Control)</td><td>Auditory</td><td rowspan="2">-</td></tr>
<tr><td>Quiet</td></tr>
<tr><td rowspan="2">BVLOS-M</td><td>Auditory</td><td rowspan="2">-</td></tr>
<tr><td>Quiet</td></tr>
<tr><td rowspan="2">BVLOS-G</td><td>Auditory</td><td rowspan="2">-</td></tr>
<tr><td>Quiet</td></tr>
<tr><td rowspan="12">Spotting (Task B)</td><td rowspan="12">Horizontal Deviation
Vertical Deviation
Timeliness
(Table 4)</td><td rowspan="4">VLOS (Control)</td><td rowspan="2">Auditory</td><td>No Wind</td></tr>
<tr><td>Wind</td></tr>
<tr><td rowspan="2">Quiet</td><td>No Wind</td></tr>
<tr><td>Wind</td></tr>
<tr><td rowspan="4">BVLOS-M</td><td rowspan="2">Auditory</td><td>No Wind</td></tr>
<tr><td>Wind</td></tr>
<tr><td rowspan="2">Quiet</td><td>No Wind</td></tr>
<tr><td>Wind</td></tr>
<tr><td rowspan="4">BVLOS-G</td><td rowspan="2">Auditory</td><td>No Wind</td></tr>
<tr><td>Wind</td></tr>
<tr><td rowspan="2">Quiet</td><td>No Wind</td></tr>
<tr><td>Wind</td></tr>
</table>

4.2.3 Stimulus and Materials

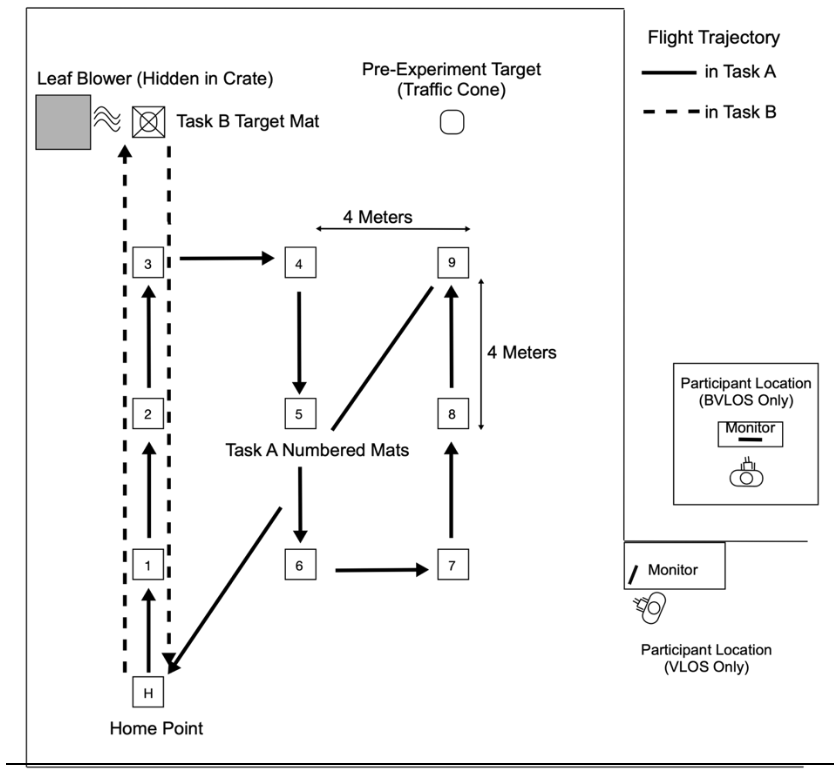
The RPAS operated by participants was a DJI (Dà-Jiāng Innovations) Phantom 4 Professional Obsidian model quadcopter, with propeller guards, and five spare batteries. The quadcopter, fitted with propeller guards and microphone transmitter weighed approximately 1.6kg, had a maximum speed of 27 knots, and a battery life of approximately 15-20 minutes in no wind conditions. A Vision Positioning System (VPS) was utilised in lieu of the Global Positioning System (GPS) for stability augmentation, employing onboard ground-facing cameras to maintain the RPA hovering position if/when participant manual control inputs temporarily ceased mid-flight. It had an on-board 20 mega-pixel camera that provided real-time visual imagery, with an 84° field-of-view (FOV). The imagery presented via the DJI GO 4 app to a compatible external display source (e.g., monitor, goggles) was identical in terms of FOV and the positioning of the displayed telemetry. The auditory recording device was secured to the quadcopter using double sided tape and Velcro straps, while the microphone (used to transmit real-time auditory cues of the rotating propellers) was secured 20cm above the central point of the quadcopter using a purpose-built holder.

To display visual imagery two Hewlett Packard (HP) Elite 23 inch (1920 x 1080 resolution) high-definition monitors (model E232) were used for the VLOS (Control) and BVLOS-M conditions, and one DJI FPV Goggles was used in the BVLOS-G condition. To record and transmit propeller auditory feedback to the participant, the RØDELink Filmmaker Kit and Bose QuietComfort 35 wireless headphones II were used. One pair of Stihl ComfortMAX hearing protector (Class 5, 29 dBA) was used in the Quiet condition only. To create wind shear in Task B, one Ryobi leaf blower was used. Wind speed was measured using one Kestrel handheld digital anemometer.

For the flying tasks, eleven carpet squares (456mm x 456mm); nine numbered from 1 through 9, one coloured bright green (Task B target mat), and one coloured red with a circular ‘H’ located in the centre (home point; coloured white, over a black circular background) were used. A bright orange cone was also used for a pre-experiment proficiency check. All numbered mats, the home point and the target mat were positioned on the ground 4.0m from each other in the horizontal X and Y planes, i.e., equidistant. The research was conducted in an enclosed aircraft hangar located at Bankstown Airport during summer daylight hours. The doors of the hangar were closed, and access was restricted to all non-participants in the hangar for the duration of the research. An additional briefing room (approximate size 2.5m x 2.7m) located in the same hangar building was used for the BVLOS-M and BVLOS-G visual conditions. The hangar area where the RPA was operated under the two BVLOS conditions was not visible from the adjacent briefing room. A representation of the experimental flying area is shown in Figure 3.

Figure 3

A Diagram of the Flying Area for Tasks A and B



Note. All Task A numbered mats were positioned on the ground at 4m from each other in the X and Y planes. For the VLOS (Control) condition, all participants had access to a monitor visual display.

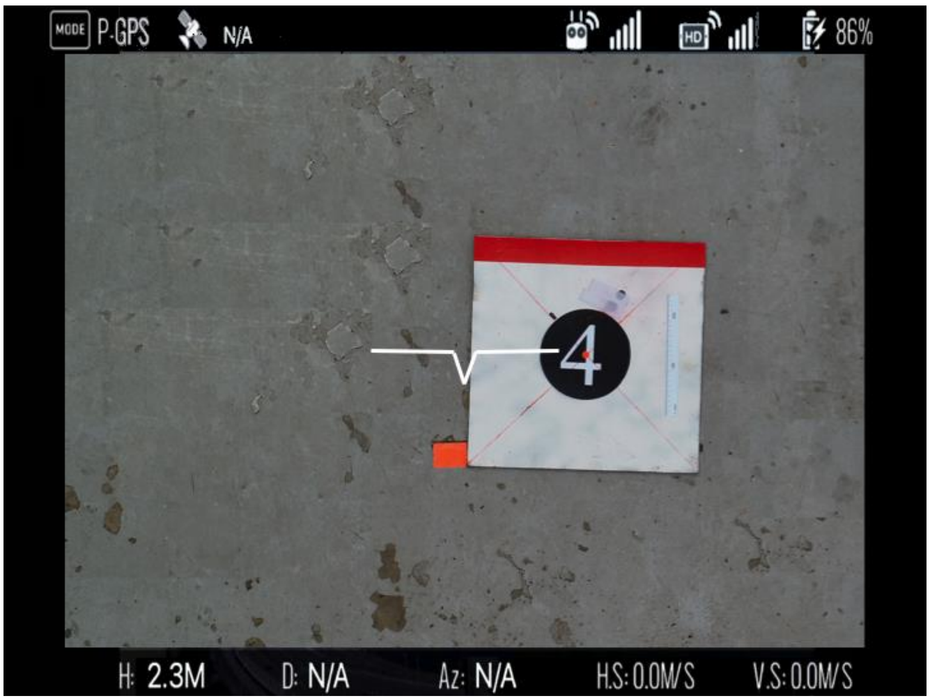
4.2.4 Procedure

Pilots were informed about the research through a series of flyers placed on notice boards at the Flying Operations Unit at UNSW Sydney. Interested pilots contacted the researcher via email or in person and made a mutually suitable time to complete the research. Prior to undertaking the main experiment, all participants (i.e., pilots) underwent training regarding RPAS operations. Instructions were provided to the pilots about the control inputs on the remote control, the capability of the RPA to hover in place if they needed to remove their hands from the remote control during the flight, the visual display interfaces, and how to obtain target photographs (depress right button on the remote control). The briefing was followed by a series of practice flights permitting them an opportunity to become comfortable with the flight controls. Once the participant reported being comfortable with the flight controls (on average 5 minutes), each participant had to demonstrate their proficiency by completing three consecutive flying tasks, under each of the three different visual conditions: VLOS (Control), BVLOS-M, and BVLOS-G. The pre-experiment proficiency test involved flying the RPA from the home point (specified cruise height at 2.0m, as shown by telemetry available on the monitor and goggles display), navigating to the target (bright orange cone), utilising the visual display interface to obtain a target photograph (orange cone) and returning to the home point (see Figure 3). On average, this proficiency test took approximately 10 minutes to complete and represented the main flying task requirements of Tasks A and B. Following the pre-experiment proficiency test, participants were provided a briefing and instructions for the main experiment including to complete each task as accurately and timely as possible. No emphasis was placed on the importance of one dependent variable versus the other.

In the case of the VLOS (Control) condition, two visual cues were presented simultaneously to the participant (RPA and monitor). The participant visually sighted the physical RPA with their own eyes, as well as the video transmitted in real-time from the RPA onboard camera to the monitor display. The participant used the monitor display to assist with aligning the target mat and referencing the altitude information to maintain the task requirements. In the BVLOS conditions (Monitor and Goggles) only one visual cue was presented (see Figure 4).

Figure 4

A Representation of the Numbered Mat Shown to Participants



Note. The imagery presented was identical in all VLOS (Control), BVLOS-M and BVLOS-G conditions with real-time telemetry information. Abbreviations located in the bottom of the screen represent H – flight altitude (in metres), H.S - horizontal speed and V.S – vertical speed (in metres per second).

Task A (navigation task) was performed once under each of the six experimental conditions. Participants were tasked to take-off from the home point and ascend to a height of 2.0m, as indicated by real-time telemetry presented on the visual display (see Figure 4).

From there, they were to proceed to the target mat labelled ‘1’, take a photograph of the mat with the mat as centred as possible, and continue this routine in numbered order to the last mat labelled ‘9’. In this experiment, nine waypoint mats presented in a 3 x 3 format provided participants an opportunity to complete two right hand turns and two left hand turns. Upon completion of photographing the last numbered mat ‘9’, participants were tasked to return to the home point and land.

In Task B (spotting task), participants were required to photograph only one mat (i.e., the target mat; see Figure 3), with and without the additional wind shear component. Specifically, participants were asked to take-off from the home point and ascend to a height of 2.0m, proceed directly to the Task B target mat, descend to a height of 0.5m over the mat, take a photograph of the target mat with the image as centred as possible, ascend back to 2.0m and then return and land at the home point. In this task, two different wind component conditions were presented to modify workload (i.e., no wind and wind), and like in Task A, each were completed with auditory feedback and without auditory feedback. In the no wind condition, the flight was performed with zero wind, i.e., the same operating conditions presented during Task A. In the wind condition, wind shear was simulated using the electric wind blower, hidden adjacent to the Task B target mat inside a wooden box and out of view of the participant. Wind, at a speed of 16 knots (as determined by a Kestrel handheld digital anemometer) was introduced overhead the Task B target mat between a height of 1.2m and 1.8m. The manufacturer recommends the RPA not to be flown in wind conditions that exceed 19.4 knots, hence this task was completed safely within the operating limits of the RPA.

In all Auditory conditions, participants adjusted the headphone volume (Bose headphones) to suit. In all Quiet conditions, the Stihl hearing protector attenuated the sound. Ambient sound levels in the hangar ranged from 40-43 dBA. The sound level of the RPA, at five metres (closest point to operator) was measured at 72 dBA. The noise reduction rating of the hearing protector was 29 dBA. The at ear sound level with the RPA operational, hearing protector on, and participant standing 5 metres from the RPA was recorded at 42 dBA (test apparatus included sound level meter with plastic ear cup mould with O-ring seal; see Burgess & Molesworth, 2016). The ambient sound level in the adjacent briefing room under the same conditions was between 39-42 dBA.

In all experimental conditions, participants were provided a seat. However, the seat was only voluntarily used by participants in the BVLOS-M and the BVLOS-G experimental conditions. In the VLOS (Control), participants elected to stand. This method is ecologically congruent with real world operations. The total time to complete the research, per participant was on average two hours.

4.2.5 Measurements

The photographs obtained had dimensions of 4864 x 3648 pixels, thus, the central point of the photo was determined to be at the 2432 x 1824 pixel coordinate. Horizontal deviation was calculated by determining the distance between the centre of the target mat, and the central point of the photograph by measuring pixel distance to one decimal point and converting to a physical length in centimetres. Vertical deviation was measured by comparing the mat width in pixels against the applicable vertical height reference photo.

The length of time to complete each flight (i.e., taking off from the home point, flying the required task, returning to the home point and landing) was recorded through the use of a stopwatch. The timing measurement was performed by the same researcher for all participants.

4.2.6 Data Analysis

The sample size was sufficient to detect between-group effects using Cohen's (1988) criteria. With alpha at .05, and power at .95, the projected sample size needed (using G*Power3.1) for this effect size is 14 (N). Three dependent variables featured in each operational condition for both Task A and Task B, namely:

- horizontal deviation (absolute, in centimetres) away from centre of target,
- vertical deviation (absolute, in centimetres) from specified target flying height, and
- time (seconds) taken to complete the task.

These variables were analysed using a repeated measures analysis of variance (ANOVA). Prior to each analysis, all data was screened for violations of the assumptions of homogeneity of variance using the interquartile range technique suggested by Moore et al. (2009), and amended using the next highest or lowest score plus one technique (stated when occurred; Tabachnick & Fidell, 2013).

Alpha was set at .05 for all analyses, including the repeated measures ANOVA and planned post-hoc pairwise comparisons in accordance with Rothman (1990) and Armstrong (2014). Since the interpretation of a single test is dependent on the performance of other tests, which is contained within that specific data set, protection against type II error was maintained. If data was repeatedly used across tests, alpha was corrected using Bonferroni correction to protect against type I error (alpha level stated when this occurred; Howell, 2009). Pre-data screening revealed 17 missing data points. Missing data resulted when the participant thought a photograph had been taken, but was not, and was managed by substituting the missing data with the mean individual value for each participant, as recommended by Tabachnick and Fidell (2013). The complete dataset was then screened for outliers. A total of 28 outliers out of 1,188 data points were identified and transformed, leaving the dataset homogeneous.

To determine the influence of pilot flight experience on the obtained results, a series of Pearson product-moment correlations (with alpha set at .05) were performed between total flight hours and performance (horizontal deviation, vertical deviation, and time) during the control condition (VLOS (Control) and no wind) for each task (Task A and Task B). The correlational analysis failed to reveal any significant relationship between total flight hours and the three dependent variables, (largest r, $r(18) = .20$, $p = .43$) for vertical deviation in Task A. Since no relationship was found between previous flight hours and task performance as measured in this experiment, flight experience was not included as a covariate in any analysis.

4.3 Results

4.3.1 Task A: Navigation Task

4.3.1.1 Horizontal and Vertical Deviation

Table 3 displays the inferential statistics from the repeated measures analysis for Task A. Only one significant main effect for visual display type was found for horizontal deviation (p = .034). No main effect or interaction were evident for vertical deviation.

Table 3

Summary of Main and Interaction Effects in the Navigation Task (Task A)

Variables	*MS**	*F*	η_p^2	*p*
Horizontal Deviation				
Visual	14.21	3.76	.18	.03
Auditory	.16	.02	.00	.90
Visual x Auditory	3.51	1.01	.06	.38
Vertical Deviation				
Visual	413.48	1.93	.10	.16
Auditory	22.49	.19	.01	.67
Visual x Auditory	52.70	.27	.02	.77
Timeliness				
Visual	3253.90	6.93	.29	.003
Auditory	1108.48	3.58	.17	.08
Visual x Auditory	142.68	.39	.02	.68

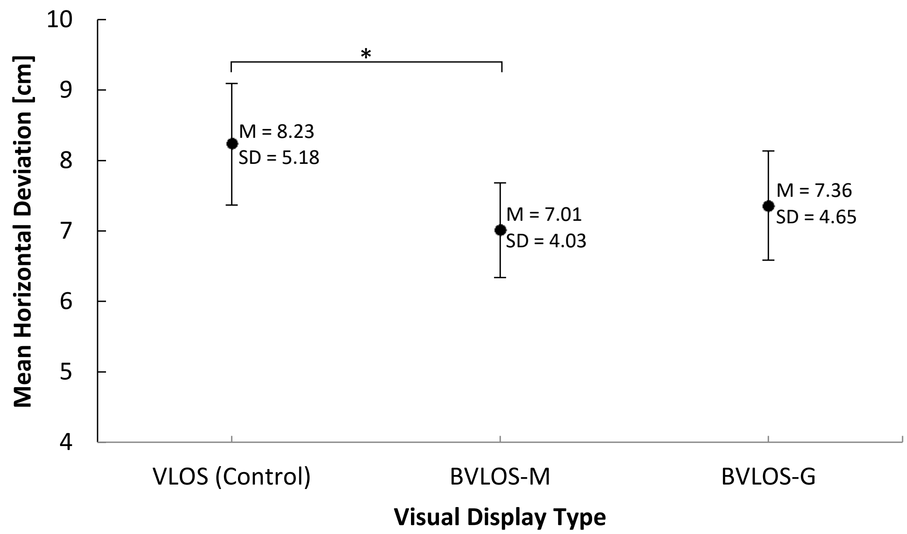
Note. *MS* refers to ‘Mean Square’.

Figure 5 displays the results of the simple effect post hoc analysis, namely a series of pairwise comparisons for the horizontal deviation for the visual condition. As can be seen from this figure, remote pilots in the BVLOS-M condition were horizontally more accurate than compared to the VLOS (Control) condition (p = .030). There was no statistically significant effect between the BVLOS-M and BVLOS-G (p = .289) conditions or the VLOS (Control) and BVLOS-G (p = .107) task conditions.

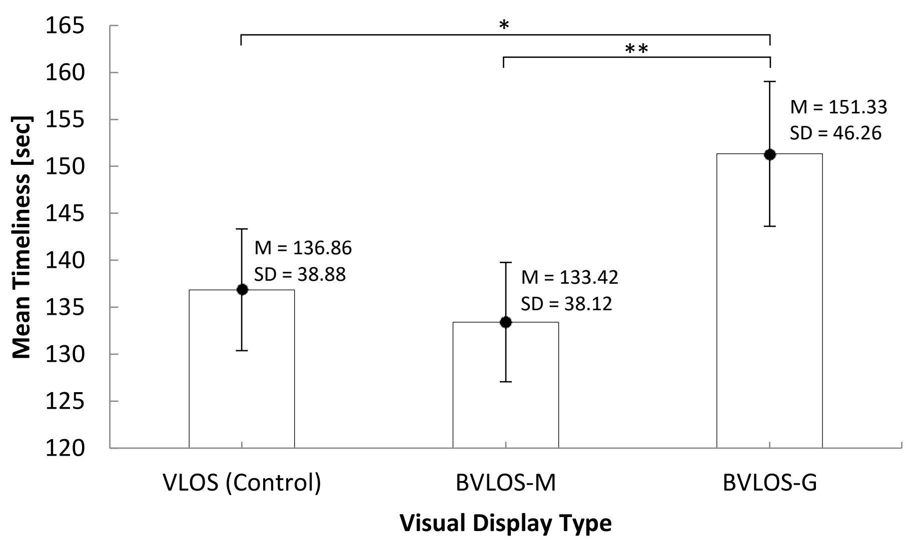
Figure 5

The Mean Horizontal Deviation (a) and Mean Timeliness (b) in Task A for Three Visual Display Types

a)



b)



Note. The error bars denote the standard error of the mean. Note: * $p < .05$, ** $p < .01$.

4.3.1.2 Timeliness

The results of the repeated measures analysis for timeliness are shown in Table 3. Only one main effect for the Visual display type was present. The results of the simple effects post-hoc analyses for visual display are presented in Figure 5. As evident in this figure, no significant effect was found between the VLOS (Control) and BVLOS-M visual conditions (p = .523). However, significant effects were found between the VLOS (Control) and BVLOS-G condition (p = .006) and BVLOS-M and BVLOS-G condition (p = .004). Participants took longer to complete the task when they were using the goggles compared to the two other conditions.

4.3.2 Task B: Spotting Task

4.3.2.1 Horizontal Deviation

Task B involved a spotting exercise under two distinct wind component conditions (No Wind vs Wind). Table 4 displays the results of the first repeated measures ANOVA examining the differences between groups based on horizontal deviation. Two significant main effects were present, one for Auditory and one for Wind.

Table 4

Summary of Main and Interaction Effects in the Spotting Task (Task B)

Variables	*MS**	*F*	η_p^2	*p*
Horizontal Deviation				
Visual	18.85	1.4	.08	.25
Auditory	36.39	7.65	.31	.01
Wind	93.91	6.26	.27	.02
Visual x Auditory	.65	.05	.00	.95
Visual x Wind	4.82	.44	.03	.65
Auditory x Wind	16.41	2.03	.11	.17
Visual x Auditory x Wind	1.10	.10	.01	.91
Vertical Deviation				
Visual	128.21	1.17	.07	.32
Auditory	22.69	.24	.01	.63
Wind	49.69	.51	.03	.48
Visual x Auditory	462.46	4.66	.21	.02
Visual x Wind	23.92	.32	.02	.73
Auditory x Wind	97.88	.55	.03	.47
Visual x Auditory x Wind	701.05	5.31	.24	.02
Timeliness				
Visual	548.46	1.65	.09	.21
Auditory	.30	.00	.00	.98
Wind	4446.30	15.92	.48	.001
Visual x Auditory	147.57	.74	.04	.48
Visual x Wind	285.57	1.77	.09	.19
Auditory x Wind	271.13	1.49	.08	.24
Visual x Auditory x Wind	91.13	1.03	.06	.37

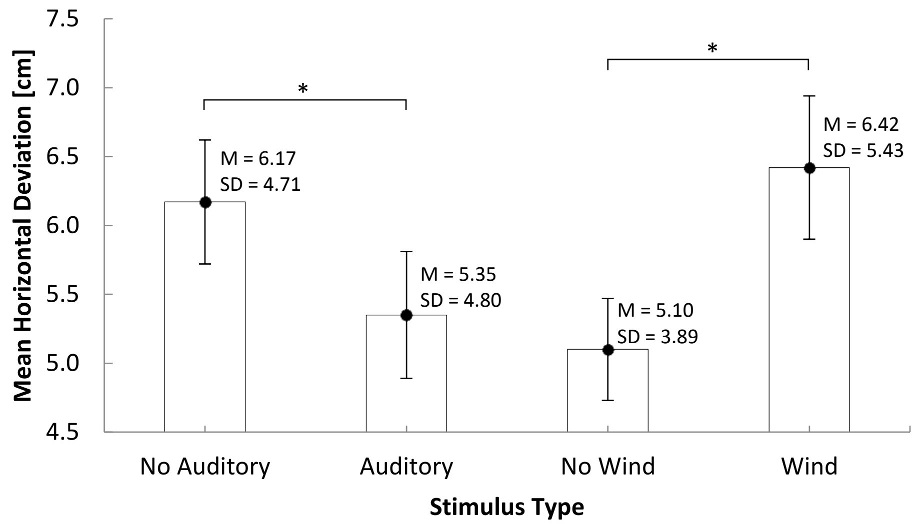
Note. *MS* refers to 'Mean Square'.

As shown in Figure 6, the presence of auditory feedback was found to improve horizontal accuracy (i.e., a lower value is more accurate). Recall in Task A, which had no wind conditions, no main effect of auditory feedback was found in terms of horizontal accuracy. There was also no Auditory by Wind interaction. Compared to Task A, the additional wind component present in Task B appeared to be enough to produce a significant effect.

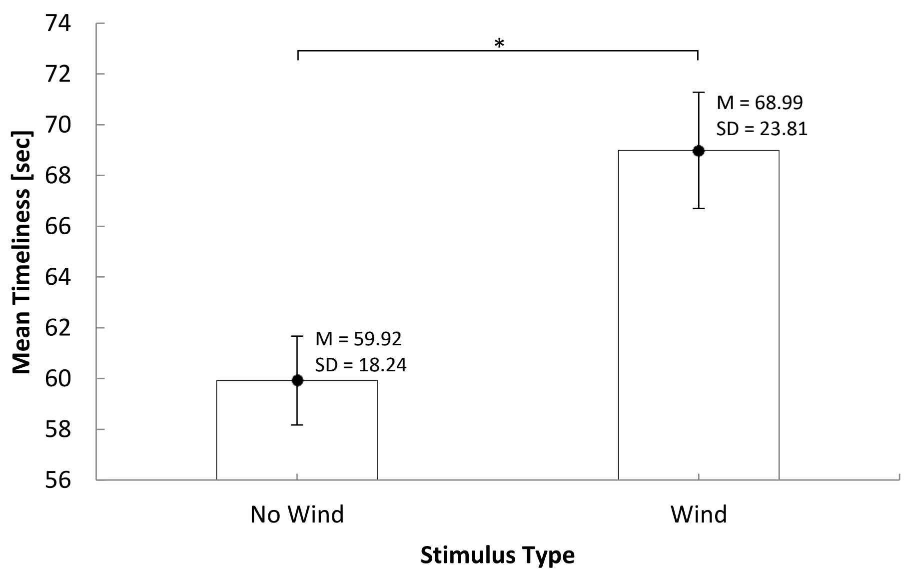
Figure 6

The Mean Horizontal Deviation (a) and Timeliness (b) for Auditory and Wind Conditions in Task B

a)



b)



Note. The error bars represent the standard error of the mean. Note: * = $p < .05$, ** = $p < .01$.

A main effect for wind is also illustrated in Figure 6. As can be seen in this figure, the addition of the wind component adversely affected participant task performance.

4.3.2.2 Vertical Deviation

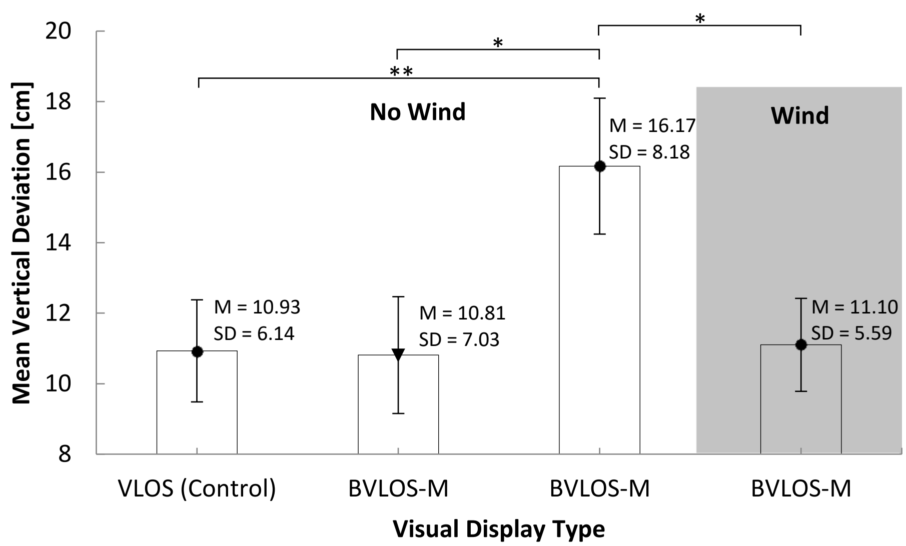
As can be seen in Table 4, a two-way interaction between the Visual and Auditory conditions, and a three-way significant interaction between the Visual, Auditory and Wind conditions was evident for vertical deviation accuracy. Since the three-way interaction captures the two-way interaction, only this interaction was investigated further. Hence, the three-way interaction for ‘Visual’ by ‘Auditory’ by ‘Wind’ was analysed by using a series of simple effect analyses. The first involved two separate two-way interactions for Visual by Auditory, one for the ‘Wind’ condition and the other for the ‘No Wind’ condition: a significant interaction was found for Visual by Auditory with Wind (p = .039) and with No Wind (p = .004) task conditions. The second simple effect analysed involved a series of two-way interactions for Auditory by Wind, for the three different visual conditions: no difference was found for the VLOS (Control) or BVLOS-G conditions, however a significant two-way interaction was found for the BVLOS-M (p = .016) visual display condition. The third involved two separate two-way interactions for Visual by Wind, for the two auditory conditions (i.e., Auditory and Quiet); no significant results were evident in either of the conditions. Each simple effect analysis that was significant was followed by a series of paired samples *t* tests with statistically significant results presented in Figure 7.

4.3.2.3 Visual by Auditory Interaction

In terms of the series of paired samples *t* tests for the Visual by Auditory interaction with Wind, no statistically significant differences were noted between the three different visual conditions based on auditory feedback, largest *t*, $t(17) = 2.11$, $p = .050$ (BVLOS-G auditory vs BVLOS-G no auditory). In addition, there were no statistically significant differences between the visual display types for the two different auditory conditions, largest *t*, $t(17) = 2.07$, $p = .054$ (VLOS (Control) no auditory vs BVLOS-G no auditory).

Figure 7

The Mean Vertical Deviation for the VLOS (Control) and BVLOS-M Visual Display Types for Auditory Cues in Task B



Note. The results are for feedback (●) and non-feedback (▼) of auditory cues in Task B. The error bars represent the standard error of the mean. Note: * $p < .05$, ** $p < .01$.

In terms of the series of paired samples *t* tests for the Visual by Auditory interaction with No Wind, the results revealed one significant difference between the two BVLOS-M conditions for Auditory and Quiet task conditions, $t(17) = 2.14$, $p = .047$. When participants were using the monitor display under BVLOS-M conditions, vertical deviation from the target mat was superior when no auditory feedback was present (M = 10.81, SD = 7.03), compared to when auditory feedback was present (M = 16.17, SD = 8.18). A medium to large effect size ($r^2 = .21$) was indicated by the mean difference of 5.36 between the two conditions, 95% CI [.08, 10.64]. One significant effect was also evident between the two visual display types, VLOS (Control) and BVLOS-M, and with auditory feedback in the same No Wind condition, $t(17) = 3.46$, $p = .003$. No other results were significant, largest *t*, $t(17) = 2.05$, $p = .056$ (VLOS (Control) Quiet vs BVLOS-M Quiet). Participants' performance for vertical deviation was also superior under the VLOS (Control) (M = 10.93, SD = 6.14) condition with auditory feedback compared to the BVLOS-M (M = 16.17, SD = 8.18) condition with auditory feedback. A large effect size ($r^2 = .41$) was indicated by the mean difference of 5.23 between the two conditions, 95% CI [.08, 10.64].

4.3.2.4 Auditory by Wind Interaction

A series of paired samples *t* tests were also performed for the two-way auditory by wind interaction for the BVLOS-M display type. In addition to the result highlighted prior between the two auditory conditions (Auditory vs Quiet) for the BVLOS-M display type, only one statistically significant result was revealed between the two wind conditions (Wind vs No Wind), $t(17) = 2.16$, $p = .046$. When wind shear (M = 11.10, SD = 5.59) was present, the participants' performance was superior compared to the No Wind task conditions (M = 16.17, SD = 8.18). A medium to large effect size ($r^2 = .21$) was indicated by the mean difference of 5.07 between the two conditions, 95% CI [.11, 10.02].

4.3.2.5 Timeliness

The final repeated measures ANOVA for timeliness revealed only one main effect, for wind, as can be seen in Table 4. Participants performed the flying task faster in the No Wind condition compared to the Wind condition, as shown in Figure 6. This result shows a detrimental effect of wind shear on manual flying performance.

4.4 Discussion

The aim of this experiment was to assess the effectiveness of unimodal (visual) and multimodal (auditory & visual) cueing on participant manual flying performance. In pursuit of this aim, two hypotheses and one research question were asked. In terms of the first hypothesis, which predicted the presence of real-time auditory feedback would aid participant accuracy and timeliness, partial support was found (Task B, however, not Task A). Moreover, in Task B, the Spotting Task, auditory feedback was found to improve horizontal accuracy. It also improved vertical accuracy, although the results for this variable are more nuanced. The BVLOS-M condition in the presence of wind, with the inclusion of real-time auditory feedback, was found to improve vertical accuracy (i.e., 31% increase) when compared to the same visual condition with no wind. Since wind shear is not reliably predictable in a real-world setting, this result highlights the benefits of the inclusion of transmitting real-time auditory feedback to RPAS operators, or at least having it available for specific operating conditions. When there was no wind, auditory feedback hindered the participants' vertical accuracy in the BVLOS-M condition, and in this instance could be seen as a distraction rather than as confirmatory of an additional stressor (i.e., an increase in workload). In the VLOS (Control) condition, auditory feedback improved their accuracy when there was no wind only.

For participant timeliness, the presence (or not) of auditory feedback did not produce a statistically significant result in either Task A or Task B. For Task A, auditory feedback increased task completion time by almost six seconds. In Task B, when wind shear was present, auditory feedback also increased task completion time by two seconds. Interestingly, when calm conditions were present (i.e., there was no wind shear) this trend reversed by almost the same margin. In this instance, auditory feedback improved task completion time by almost two seconds.

The second hypothesis predicted a main effect for wind shear (a condition only evident in Task B); however, a main effect was only evident for horizontal accuracy and timeliness. In both conditions (horizontal: 26% increase; timeliness: 15% increase), manual flying performance in the horizontal plane was less accurate and participants took longer to complete the exercise in the presence of wind shear, compared to the calm air conditions. While no main effect was evident for vertical accuracy, recall the nuanced results for this task and in relation to the first hypothesis, where real-time auditory feedback from the propellers was expected to aid participant flying performance in terms of accuracy (horizontal and vertical deviation) and timeliness. Only the BVLOS-M with auditory feedback, and in the presence of wind shear, produced a statistically significant result for superior vertical accuracy. In contrast to Task A, this task required participants to vertically descend the RPA through a layer of turbulent air. It is possible no main effect was evident for vertical accuracy due in part to the relatively simpler nature of the task (i.e., vertical descent from a single fixed point above the target mat, versus navigating between multiple waypoint mats). A further explanation could also involve the utility of one of the three visual display types: visual-line-of-sight + monitor display. While potentially disadvantageous from the perspective of two intra-modal (i.e., visual) cues and lateral movements, being able to visually sight the RPA with one's own eyes and seeing vertical movement could be assistive irrespective of the presence (or not) of a wind shear component.

The research question asked how unimodal (visual) and multimodal (auditory & visual) feedback changes participant flying performance (accuracy and timeliness) under wind and no wind conditions. The main effects evidenced in Task A and B, as well as the interaction in Task B suggest the BVLOS-M visual display type to be advantageous when compared with the VLOS (Control) and BVLOS-G visual conditions. When an auditory component was included to form a multimodal feedback combination, its influence appears to be most vulnerable to the additional workload stressor, formed by wind and evidenced in the Task B interaction. These results appear to show a trend that is representative of the expected results when considering the Maximal Adaptability Model (Hancock & Warm, 1989), and the prediction of task performance. Hence, these findings will be considered and further explored within the context of the MAM in the next two experiments.

4.4.1 Effects of Auditory and Visual Feedback on Task Performance

The results from this experiment highlight the task specificity of visual and auditory feedback on manual flying performance. The results support the findings of Frassinetti et al. (2002) that visual perception can be enhanced by auditory information, but only when there is temporal and spatial alignment between the two cues. From an applied perspective, Wickens (2002) suggests two intra-modal cue types presented together could be problematic, but two different cues could be advantageous. As shown in Task B in the BVLOS-M visual condition with auditory feedback, vertical accuracy was superior in the presence of wind shear, a result not anticipated. When auditory stimulus is perceived as being presented before visual stimulus, or when a delayed offset is perceived between the two, the Colavita Effect can be minimised or eliminated (Koppen & Spence, 2007b). However, a disadvantageous example was also seen in Task A for horizontal deviation where the VLOS (Control) condition was least accurate. It is likely two visual cues presented together (i.e., via the combined monitor display and the visual sighting of the physical RPA by the participant) caused a perceptual dissonance in the movement of the RPA in response to control inputs from the participant. A visual change displayed on the monitor display due to a left or right manual control input, could appear differently in person when perceiving the same movement change of the physical RPA. Depending on its orientation relative to the remote pilot controlling the RPA, this difference could be amplified. While remote pilots are faced with reduced sensory cues compared to their conventional pilot counterparts, redundant cues such as two visual forms of feedback presented together, may become burdensome if oppositely oriented.

From a theoretical perspective, the results from this experiment contribute to the audio-visual literature. Frassinetti et al. (2002) noted the unidirectional nature of auditory information; it will enhance visual information, but not the other way around. The results support this, however, they indicate that it may be contingent on the operating situation. Recall in Task B and the BVLOS-M condition with wind, auditory information enhanced performance. When there was no wind, performance notably reduced. In contrast, auditory information enhanced performance in the VLOS (Control) condition only when there was no wind. This finding suggests an interaction between audio-visual feedback and the task being undertaken. Auditory information is highly valued when the visual stimuli is lacking in detail. However, when the visual stimulus is sufficient, the auditory information could be considered as noise (i.e., sound that is unwanted), and cognitively taxing. As described previously in Chapter 2, auditory noise has repeatedly been shown to adversely affect task performance as a distractor, and by unnecessarily consuming individuals' limited information processing resources (Becker et al., 1995; Ljung et al., 2009; Molesworth et al., 2013; Molesworth et al., 2014; Molesworth et al., 2015).

4.4.2 Limitations and Future Research

The results of this experiment (Experiment 1) provide a valuable insight into the dynamic effects of audio-visual feedback on remote pilot task performance under two workload levels. However, these findings do need to be interpreted within the confines of the research environment. The experiment was conducted between the daylight hours of 1100 and 1600 during summer. How the time of day and visual setting afforded by daylight compared to night conditions affected participant performance remains unknown. The research was also conducted under controlled environmental conditions inside an aircraft hangar and allowed for a wind and non-wind environmental component. From an experimental perspective, this aids in conducting high quality research, however from an applied perspective this may limit the extent to which the results can be generalised. Hence, both areas for future research.

This experiment also involved manually-controlled BVLOS operations, a unique configuration normally requiring extensive pilot training and licencing to perform in a real world environment. As such, the experiment could not be conducted outdoors due to safety and regulatory requirements imposed by the national aviation regulator (CASA), and university policy. Therefore, all participants were conventional fixed-wing aircraft pilots, with the requirement that they had flown a single-engine aircraft solo. Non-pilot participants (i.e., with no previous piloting experience or training), and remote pilot participants (i.e., those who hold a remote pilot licence or equivalent and/or who had received substantial RPAS training/experience), were explicitly absent from the sample and were not targeted for recruitment. While differences in the performance of the fixed-wing pilots were measured across the different auditory, visual and workload levels, it is evident that collectively the group performed the task successfully. No crashes of the RPA were recorded, and all tasks were completed successfully. Given the task involved manual flying only, and no automation was available, how non-pilots or professional remote pilots might have navigated the course in a BVLOS setting is unknown. Furthermore, how their performance would have compared to conventional pilots between each task and the associated variables is also unknown. Hence, further areas of research.

The lack of a subjective workload assessment could also be considered a limitation of this first experiment. The additional wind component in Task B, and the requirement to manual fly the RPA through the turbulent wind layer produced obvious challenges compared to the calm conditions for the duration of all Task A flying. This was evidenced anecdotally by participants following the experiment, who suggested the additional wind component made the task more difficult to complete. While the two tasks were different, no formal secondary task (i.e., an additional task completed simultaneously) was used to induce a controlled workload change. This limitation provided an opportunity to assess subjective workload in future research, and was addressed in the second and third experiments.

Furthermore, all flying tasks required hands-on control inputs to fly the RPA throughout the research area. In a real-world setting under BVLOS operating conditions, RPA flights are normally conducted in an automated configuration involving pre-programmed flight paths. It is more common for these flights to make use of monitor displays only (i.e., no FPV goggles, and no VLOS). While it was useful in this experiment to assess participants' ability to manually control an RPA under different workload levels and auditory feedback, it remains an unknown how task performance might have differed under automated flying conditions. Hence, another area for future research.

This experiment measured the effects of auditory feedback, visual display type and two levels of workload on remote pilot manual flying task performance. While nuanced results were found, they broadly reflect the expected values associated with predictive behavioural models of task performance, namely the Maximal Adaptability Model and Arousal Theory. In a civil or commercial context, operating an RPAS can draw attention from onlookers, whether they be curious professionals or members of the public. This can be a source of distraction, and hence divert the RPA operators’ attention away from the task of monitoring and controlling the RPA, particularly if is being operated in an automated configuration. How these distractions, or additional layers of workload, impact on the task performance of the remote pilot via changing concentration or increased arousal, remains unknown. Hence, in consideration of the MAM and AT behavioural models, further investigation into remote pilot task performance will be assessed in the next experiment, specifically with further increments of workload added to the experimental conditions.

The current experiment also required participants to manually operate the RPAS under both VLOS and BVLOS conditions. However, future civil RPAS flights will increasingly be operated under automated BVLOS configurations. How performance varies under such conditions remain unknown. Hence, Experiment 2 will also measure the task performance of participants under automated conditions, while continuing the manipulation of auditory feedback and workload variables. Furthermore, the current experiment only tested participants with previous fixed-wing flying experience. While this allowed for consistency between all task conditions, it is important to also test non-pilots who increasingly have access to and are operating civil RPAS presently. Experiment 2 will expand upon the findings of this experiment by continuing the investigation into the effect of auditory feedback on remote pilot task performance, including its potential to be perceived as helpful information or noise when combined with different workload levels. Experiment 2 will also test for differences in task performance between two groups under automated RPAS flying conditions: pilot participants with previous fixed-wing flying experience, and non-pilots.

4.5 Summary

This chapter described the first of three experiments, which tested participants' manual operation of an RPAS under various conditions. All participants in this experiment had previous fixed-wing flying experience in a conventionally piloted aircraft. Task performance was measured under a series of experimental conditions comprising three different visual display types and the presence or not of both real-time auditory feedback and wind shear. Participant flying performance was found to be superior when using a monitor visual display type. The presence of additional stressors in the form of auditory feedback and workload (i.e., wind shear) also influenced task performance. Specifically, how the availability of sensory cueing in the form of real-time auditory feedback, and the degree to which workload as an added stressor influenced task performance was investigated.

The findings suggest performance including timeliness and horizontal accuracy can be enhanced under BVLOS conditions where a monitor display as opposed to FPV goggles are used. The first hypothesis predicted auditory feedback, generated from the RPA propellers, would aid the performance of the participant in terms of accuracy and timeliness. This was evident in the spotting task for horizontal deviation where a main effect was found. Vertical deviation also improved in the spotting task when two visual sources (i.e., VLOS (Control)) were available to the participant, and presented simultaneously with auditory feedback. Conversely, auditory feedback was found to have a detrimental effect on vertical deviation when utilising the BVLOS-M display type under no wind conditions. Furthermore, wind shear adversely affected participant performance in terms of horizontal accuracy and timeliness and is a partial confirmation of the second hypothesis. Hence, it is the specificity of the task in combination with different auditory cueing intensities which appears to be most influential on remote pilot performance. Although statistically significant results were not attained for all task conditions and variables, the findings broadly represent the expected numerical trends predicted by the Maximal Adaptability Model. It is apparent that the BVLOS-Monitor visual operating condition, irrespective of level of workload, reflects an optimal configuration to facilitate superior task performance. How this flying configuration is affected by further increments of workload and the presence or not of auditory feedback in an automated setting, as well as for participants with no previous flying experience, will be explored in the next experiment.

CHAPTER 5 – EXPERIMENT 2: MEASURED EFFECTS OF WORKLOAD AND AUDITORY FEEDBACK ON REMOTE PILOT TASK PERFORMANCE

***Note. The contents of this chapter, with the addition of a brief introduction are published in the journal titled Ergonomics. Below is a reference for this publication.**

Dunn, M. J., Molesworth, B. R., Koo, T., & Lodewijks, G. (2022). Measured effects of workload and auditory feedback on remote pilot task performance. *Ergonomics*, *65*(6), 886-898. https://doi.org/10.1080/00140139.2021.2003870

5.1 Introduction

Experiment 2 directly extends the findings of Experiment 1 and examines the effect of other known stressors, such as workload and real-time auditory feedback on participant task performance (i.e., spotting accuracy and spatial orientation accuracy). However, for this experiment, an automated BVLOS flying task was used. In the previous experiment, nuanced examples were found of auditory feedback effects on remote pilot manual flying performance under VLOS and BVLOS operating conditions. Specifically, in operating conditions perceived to be reflective of high workload (i.e., wind shear), auditory information presented as real-time auditory feedback from the remotely piloted aircraft (RPA) was found to assist the remote pilot in a spotting task. Conversely, in operating conditions perceived to be reflective of low workload (i.e., nil wind), real-time auditory feedback from the RPA hindered performance, possibly because it was considered as noise (i.e., unwanted, or distracting sounds). These findings were derived from traditional fixed-wing pilots manually flying an RPAS and hence, the research leaves open at least three issues.

First, changes in the level of workload were inferred based upon the manipulated meteorological conditions (i.e., an increase in wind shear). Whether changes in workload manipulated through more traditional and accepted methods (e.g., performing a secondary task during the main experiment) yields similar results remains unknown. The introduction of a subjective measure, such as the NASA-TLX, to verify this manipulation is hence warranted. Second, all participants had previous fixed-wing flying experience (i.e., involving conventional aircraft cockpits and with the associated availability of audio-visual sensory cueing). How this prior experience, and the presence or absence of auditory feedback under BVLOS conditions impact task performance, also remains unknown. Third, the task that was employed required the participant to manually control and fly the RPA in an enclosed indoor environment. RPAS operated under BVLOS conditions are commonly programmed for automated flight (Fang et al., 2018; Zmarz et al., 2018), and thus, require little or no operator input during normal operations (Politi et al., 2021). However, for non-routine operations (e.g., inflight system failures), how audio-visual feedback affects operator performance under BVLOS conditions also remains unknown. Hence, Experiment 2 was designed to address these three issues.

Continuing the prediction of task performance measured in Experiment 1, and the extent to which the AT or MAM could be used to explain those findings, three research questions and three hypotheses were considered.

5.1.1 Research Questions

1. What is the effect of real-time auditory feedback on participant perception and decision-making performance?
2. How does workload moderate this effect?
3. How does piloting experience moderate this effect?

5.1.2 Hypotheses

1. Auditory feedback will have more of an effect on the decision-making task performance than the perception task.
2. Task performance levels will deteriorate as workload increases (i.e., participant performance will correspond with the right-hand side of the inverted U-curve; see Figure 1 and 2).
3. Pilots will perform the tasks more accurately than the non-pilot group.

5.2 Method

5.2.1 Participants

Thirty-six (eight female) participants comprising eighteen undergraduate student fixed-wing pilots (2 female) and eighteen undergraduate non-pilot students (6 female) were recruited from the UNSW Flying Operations Unit, and the UNSW Sydney main campus respectively. Demographic information regarding participants can be found in Table 5.

Table 5

Demographic Information of the Sample Group

Descriptive Variable	*M*	*SD*	*Range*
Pilots (n=18)			
Age (Years)	21.10	1.70	19.00 – 25.00
Total Flying Hours	114.00	73.40	45.80 – 245.00
Total RPAS Experience (Hours)	1.10	2.70	0 – 10.00
Non-Pilots (n=18)			
Age (Years)	23.80	2.60	20.00 – 29.00
Total Flying Hours	0	0	0
Total RPAS Experience (Hours)	6.10	25.90	0 – 110.00

Note. *M* refers to ‘Mean’. *SD* refers to ‘Standard Deviation’. N = 36.

Fixed-wing pilots reported a mean of 114.0 flying hours (SD = 73.4) and non-pilots reported no previous flying experience. Participants reported no previous RPAS flying experience except for three fixed-wing pilots who reported a mean of 1.25 hours (SD = 3.15) and one non-pilot who reported a total of 110 hours. The research, including the stimuli and procedure was approved in advance by the UNSW Sydney Ethics Panel (HC Number: HC190808). The research took approximately one hour to complete per participant, for which participants were reimbursed in the form of a $20 large department store gift voucher.

5.2.2 Design

The experimental design comprised a 3 x 2 x 2 mixed repeated measures design. The first repeated measures factor labelled Workload contained three levels: Low, Moderate and High. The second repeated measures factor labelled Auditory Feedback contained two levels: with real-time auditory feedback from the RPA (Auditory) versus no auditory feedback from the RPA (Quiet, i.e., ambient). The sole between groups factor labelled Aircraft Flight Experience contained two levels and was included to answer the third research question: participants with traditional fixed-wing flying experience (Pilots) versus no flying experience (Non-Pilots). To reduce a possible learning effect, the stimuli comprising six conditions (Workload and Auditory feedback) were presented in a balanced Latin-square design (Table 6) ensuring each task and video pairing were unique for each participant.

Table 6

Balanced Latin-square Design for Task Order and Video Pairings

	Flight 1	Flight 2	Flight 3	Flight 4	Flight 5	Flight 6
Participant 1	A6	B3	F4	C5	E2	D1
Participant 2	B6	C1	A4	D3	F2	E5
Participant 3	C3	D4	B2	E6	A5	F1
Participant 4	D6	E1	C2	F5	B4	A3
Participant 5	E4	F3	D2	A1	C6	B5
Participant 6	F6	A2	E3	B1	D5	C4

Note. Letters represent the task condition (A through F), and numbers represent the paired video (1 through 6).

Dependent variables for each task were waypoint spotting accuracy (perception task) and spatial orientation accuracy (decision-making task). A summary of the experimental design is presented in Table 7.

Table 7

Overview of Experimental Design

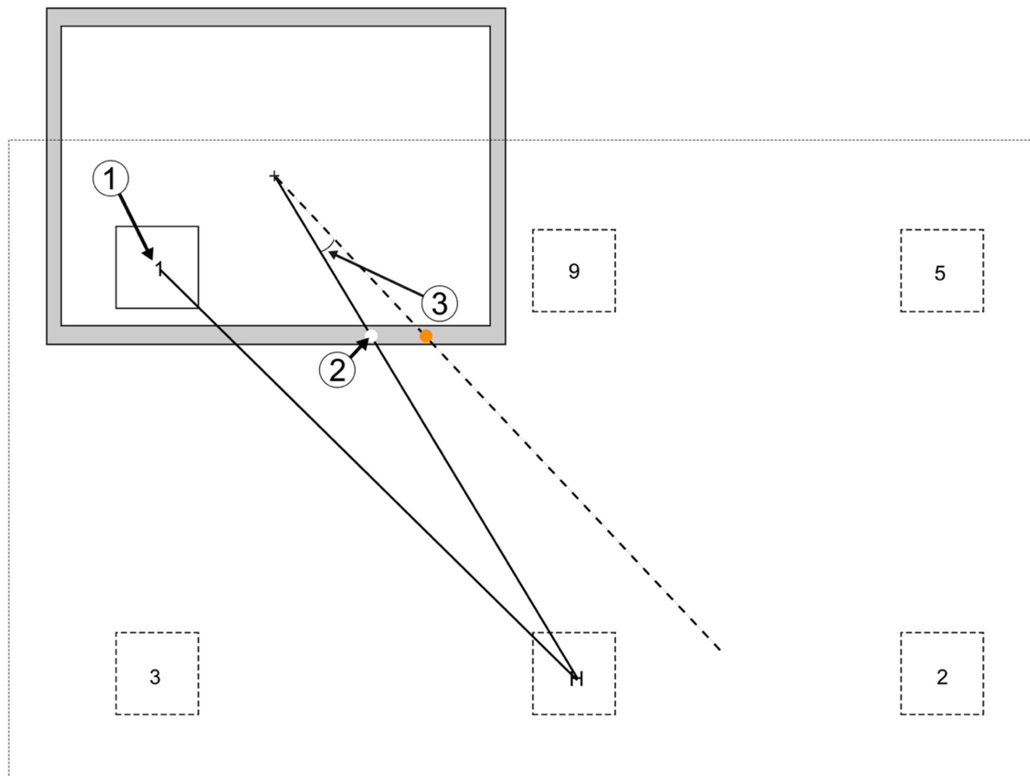
Dependent Variables	**Experimental Condition**	
	Factor 1 (Workload)	Factor 2 (Auditory Feedback)
Waypoint Spotting Accuracy (Correct Recordings)	Workload (Low)	Auditory Quiet
	Workload (Moderate)	Auditory Quiet
Spatial Orientation Accuracy (Degrees)	Workload (High)	Auditory Quiet

Waypoint spotting accuracy involved correctly identifying and recording the numeral on each waypoint mat, via an iPad display as it appeared on the monitor display. The RPA onboard camera would always face directly to the ground, hence the imagery transmitted from the RPA to the monitor display would be always consistent during the flight. Depending on the RPA's flight path, the number would appear correctly oriented, upside-down or on a side. To avoid confusion between the number 6 and number 9, participants were made aware that only the number 9 was presented as shown on the scale map provided to them and available for reference throughout the experiment. That is, only the numbers 1, 2, 3, 4, 5, 7, 8 and 9 (underlined) were presented to the participant for every flight. An error was recorded into one of two groups: when either a number was missed (labelled ‘Miss’) or was incorrectly identified (labelled ‘Wrong ID’).

Spatial orientation accuracy involved the participant correctly placing a Velcro dot on the perimeter of the monitor display, in the 360-degree horizontal plane direction of the Home Point relative to the final position of the RPA at the end of each flight. At the conclusion of the experiment, the researcher calculated the spatial orientation error for each flight (measured in degrees). Using a 360-degree protractor, the difference in angle (in degrees) between the perceived Home Point direction (location of the Velcro dot) and the true Home Point direction (determined by string intercept) was determined (see Figure 8).

Figure 8

A Representation of the Method Used to Calculate Spatial Orientation Error



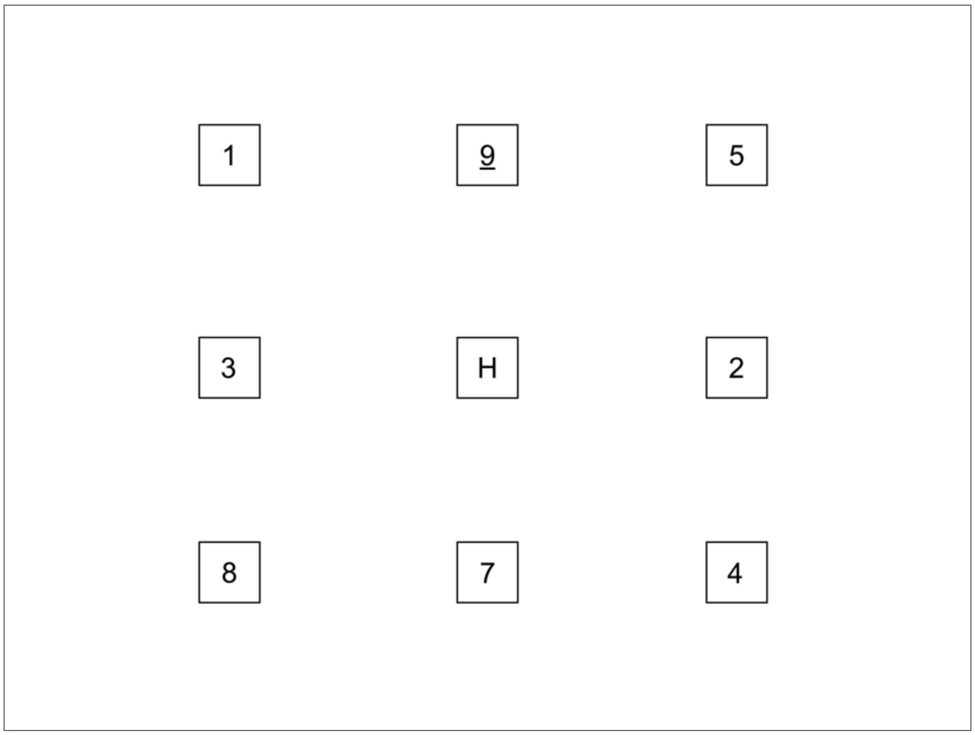
Note. The flying area is overlayed for reference. One end of a piece of string ① was placed over the centre of the final waypoint mat as it appeared paused on the computer monitor visual display. The end of a second piece of string was then placed on the centre of the computer monitor visual display. The two pieces of string were then held in the researcher’s hand and tensioned to the centre of the Home Point location (not shown on screen, but overlayed in this figure for reference, middle bottom square) until the two strings intercepted. The correct direction of the Home Point was determined by the Researcher ② who then placed a white reference Velcro dot on the edge of the computer screen. The angle difference in degrees (absolute) between the white reference Velcro dot and the participant’s Velcro dot (shown on the computer monitor visual display edge in this example in orange colour) was then measured ③ from the centre of the screen.

5.2.3 Stimuli and Materials

Six pre-recorded videos representing an automated RPAS flight were presented to participants. The apparatus comprised: a DJI (Dà-Jiāng Innovations) Phantom 4 Professional Obsidian model quadcopter (for recording of videos, utilising on-board 20 mega-pixel, 84° FOV camera), a Samsung 10 Galaxy phone (to record auditory feedback in four phases of flight: take-off, stationary/hovering, turning (yaw), and forward movement), the Movavi Video Editor Plus 2020 software (for pairing auditory feedback and video imagery), one computer, one computer mouse, two keyboards and two HP 23 inch (1920 x 1080 resolution) high-definition monitors (model LA2306x; for displaying pre-recorded videos; the first monitor was used by the participant and the second by the researcher in mirrored mode enabling the loading of videos out of sight of each participant), an Apple iPad (used by the participant in the spotting task to record the waypoint mat numbers), DT 770 Pro headphones (worn by the participant only in the auditory feedback conditions), and six pre-recorded videos depicting an RPAS flight over eight waypoints (order randomly assigned) and one home point (labelled ‘H’), which were equally spaced 8 meters apart (main experimental task; see Figure 9).

Figure 9

A Representation of the Flying Area

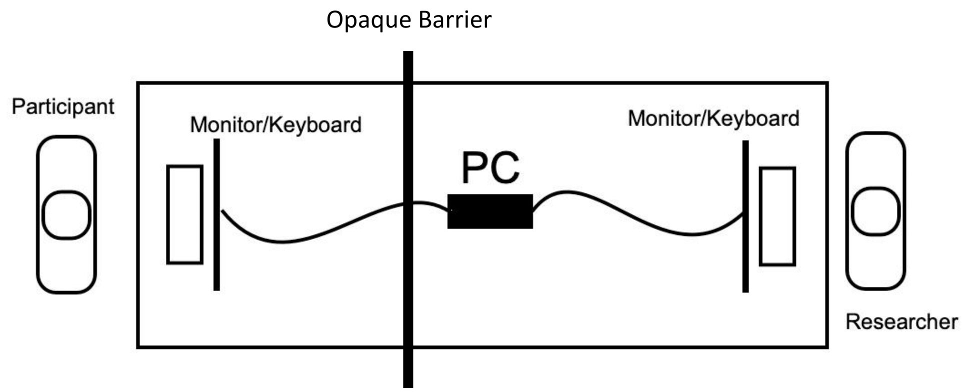


Note. All squares are equidistant in the X and Y planes.

The research was conducted in similarly configured rooms (approximate size 4m x 6m) at Bankstown Airport for pilots, and at the UNSW Sydney main campus, with a partition separating the researcher and participant (see Figure 10).

Figure 10

A Diagram of the Experimental Area Setup (Top-down View)



The materials comprised: a paper copy of the experiment instructions, a scale map of the flying area, a demographic questionnaire (i.e., age, flying experience, gender), the Serial 3s secondary task (for Moderate workload; Castro et al., 2019), the Serial 7s secondary task (for High workload; Hayman, 1942; Williams et al., 1996; Scholey et al., 2001), and the NASA-TLX questionnaire (unweighted and widely used to capture subjective workload assessment, see Appendix 2; Hart & Staveland, 1988).

5.2.4 Procedure

Participants were informed about the research through online internal UNSW advertisements and broadcast emails. Interested participants contacted the researcher, and a mutually suitable time was arranged to complete the research. Prior to undertaking the main experiment, all participants received a written and verbal briefing about the research. Following the briefing, and signing of the consent form (Appendix 3), participants were provided instructions (in paper form) about the main experiment and were asked to complete the supplied demographic questionnaire. A shortened trial task representing the main experiment (duration = 1 minute 30 seconds) was provided to each participant to demonstrate their proficiency in using the apparatus and to acknowledge the instructions given for the main experiment. Two of the six experimental conditions involved a low level of workload (Low). Two involved a moderate level of workload (Moderate) and two involved a high level (High). Utilising the Serial 3s (i.e., for Moderate workload) and Serial 7s (i.e., for High workload) procedures respectively, these secondary tasks were performed concurrently with the primary spotting task. Prior to the main experiment, baseline readings were obtained for the Moderate and High workload conditions and involved verbally counting backwards (i.e., subtracting from the previous number) in 3s for two minutes, followed by 7s for two minutes. The Researcher recorded on paper the verbal numbers stated by the participant. Following this, the main experiment commenced, and the six tasks were presented in the counterbalanced experimental order.

The primary task of spotting and recording waypoints was performed in each of the six experimental conditions. Participants were tasked to start the video by pressing the space bar of the keyboard which would commence the flight. The RPA would then take-off from the Home Point vertically to a height of approximately 4 meters before flying automatically towards the first of several waypoints (represented by the numbered mats on the ground) programmed for that flight. The waypoints were not flown in numerical order, but in a predefined and random pattern unknown to the participant and varied for each flight. Participants were then asked to record each waypoint by noting them in the order they appeared on the monitor screen using the supplied iPad (i.e, spotting - perception task). Each waypoint mat appeared on screen for an average of 6.37 seconds (SD = 1.48 seconds). The RPA flew horizontally between the waypoints at a speed of approximately 4km/hr (1m/s). At all times, a scale map (in paper form) of the flying area was available to the participant to assist with navigation. At a pre-designated waypoint unknown to the participant, the RPA would cease its mission (i.e., simulate a failure) and hover over the final waypoint mat. The participant was tasked to press the spacebar key to stop the video as soon as they realised the mission had failed. The final task in each of the six conditions required the participant to place a Velcro dot on the side of the monitor to indicate their perceived direction of the Home Point from the final RPA position (i.e., spatial orientation - decision-making task).

In four of the six conditions, a secondary task was introduced to alter the perceived level of workload. For the Moderate workload (with and without auditory feedback) conditions, participants were required to verbally count aloud backwards, starting from a predefined number between 790 and 799 in 3s (e.g., with a starting number of 798 the participant would say aloud ‘798, 795, 792’, and so forth). Likewise, in the High workload conditions, participants were required to verbally count aloud backwards from the same predefined numbers, however this time in 7s (e.g., with a starting number of 790 the participant would say aloud ‘783, 776, 769’, and so forth). The predefined starting number varied to ensure no two conditions were the same, thus minimising any learning effect. Participants were instructed to continue counting if they forgot their previous number or counted a wrong number. In all three conditions with auditory feedback (i.e., real-time, from the RPA propellers), participants adjusted the headphone (DT 770 Pro) volume to suit. Finally, at the end of each task condition, participants completed the subjective workload measure, the NASA-TLX questionnaire.

Ambient noise levels in the room ranged from 40 – 43 dBA in the Bankstown Airport briefing room for pilots and 39 – 42 dBA in the Kensington briefing room for non-pilots (test apparatus included sound level meter with plastic ear cup mould with O-ring seal; see Burgess & Molesworth, 2016). In all experimental conditions, participants were provided a seat which is ecologically congruent with real world operations.

5.2.5 Data Analysis

An a priori analysis using G*Power 3.1 determined a minimum of 16 participants per group was sufficient to detect a medium to large between groups effect (i.e., Cohen's, (1988) criteria). Power was calculated using the following assumptions: effect size 0.20, alpha error probability 0.05, power 0.8, number of groups 2, number of measurements 6. Since there were six experimental conditions, sample size featured in multiples of six, and thus 18 participants per group. Three dependent variables featured in each operational condition:

- Subjective NASA-TLX Score,
- Waypoint Spotting Accuracy, and
- Spatial Orientation Accuracy.

These variables were analysed using a repeated measures analysis of variance (ANOVA). The NASA-TLX was used as a proxy to confirm any differences between each of the manipulated workload conditions. Prior to all analyses, the data was screened for violations of the assumptions of homogeneity of variance using the interquartile range technique suggested by Moore et al. (2009). If outliers were present, the data was transformed using the next highest or lowest score plus one technique as suggested by Tabachnick and Fidell (2013; stated when occurred).

Alpha was set at .05 for all analyses, including the repeated measures ANOVA and planned post-hoc pairwise comparisons in accordance with Rothman (1990) and Armstrong (2014). Since the interpretation of a single test is dependent on the performance of other tests, which is contained within that specific data set, protection against type II error was maintained. If data was repeatedly used across tests, alpha was corrected using Bonferroni correction to protect against type I error (alpha level stated when this occurred; Howell, 2009).

Pre-data screening revealed one missing data point which resulted when the participant erased the recorded waypoints accidentally. The missing data point was managed by substituting the missing data with the mean individual value for that participant, as recommended by Tabachnick and Fidell (2013). The complete dataset was then screened for outliers and transformed where identified, leaving the dataset homogeneous.

5.3 Results

5.3.1 Subjective Workload

The first analysis sought to determine the effectiveness of the secondary task, namely the workload manipulation of the serial 3s (Moderate workload) and 7s (High workload) procedure. Thus, subjective total scores on the NASA-TLX under the three non-auditory conditions were examined by means of three planned comparisons (*t* tests). With alpha adjusted to control familywise error (Bonferroni adjustment .05/2), the results revealed significant differences between all three pairings (i.e., Low vs. Moderate, Low vs. High, and Moderate vs. High; smallest *t*, $t(35) = 3.378$, $p = .002$, $r^2 = .25$, for Moderate vs. High workload). As can be seen in Table 8, the subjective level of workload increased in line with expectations.

Table 8

Average NASA-TLX Values (Unweighted; in %) for Perceived Workload

Descriptive Variable	*M*	*SD*	*Range*
Workload – Low	16.44	12.88	1.67 – 47.50
Workload – Moderate	43.58	16.94	4.17 – 79.17
Workload – High	50.60	16.20	15.00 – 80.00

Note. *M* refers to 'Mean'. *SD* refers to 'Standard Deviation'. N = 36.

5.3.2 Waypoint Spotting Accuracy (Perception Task)

The primary task was completed by most participants with a high level of accuracy. That is, 182 out of the 216 possible tasks (i.e., 6 tasks per participant x 36 participants) were completed with no waypoint recording errors. As such, an analysis of this dependent variable could not be completed using the intended repeated measures procedure with an acceptable level of confidence due to the high number of 'zero' (i.e., perfect) recording errors.

Since errors in this task could result from a miss or wrong identification, an analysis was conducted to examine if one type of error was more common. Due to the type of data (frequency counts), a series of one-way chi-square analyses were employed. The first analysis examined differences between groups (Miss vs. Wrong Identification – Wrong ID). With alpha set at .05, the results revealed a significant difference between the two groups, χ^2 (1, N = 54) = 6.00, p = .014. As can be seen in Table 9, if an error was to occur, a Miss was twice as frequent as a Wrong ID.

Table 9

The Frequency Count for Miss and Wrong ID Error Types Across Each Task Condition

Task Condition	Error Type (Miss)	Error Type (Wrong ID)
Low Workload – No Auditory	1.00	3.00
Low Workload – Auditory	0.00	0.00
Moderate Workload – No Auditory	5.00	4.00
Moderate Workload – Auditory	6.00	4.00
High Workload – No Auditory	10.60	4.00
High Workload – Auditory	13.00	3.00
Total	35.60	18.00

Note. Count values include one substituted missing data point using the mean individual value, as described in Section 5.2.5.

The next analysis examined the effect of auditory feedback and workload on the two error types; hence a series of one-way chi-square analyses were performed. With alpha set at .017 (.05/3) to control for familywise error, the results revealed two statistically significant differences for the Miss error type; the first between the Low and Moderate workload groups (χ^2 (1, N = 12) = 8.33, p = .004), and the second between the Low and High workload groups (χ^2 (1, N = 25) = 21.16, p < .000). In contrast, there was no significant difference between the Moderate and High workload groups for the Miss error type (χ^2 (1, N = 35) = 4.83, p < .028) nor between any of the three workload groups with the Wrong ID error type, largest χ^2 (1, N = 11) = 2.27, p = .132 (between the Low and Moderate workload groups). Comparing errors based on workload between the two error types, a significant difference was only found in the High workload condition, χ^2 (1, N = 31) = 9.32, p = .002. This result indicates that under a high level of workload, when an error was to occur, it was more likely to be a Miss than a Wrong ID.

The presence (or not) of auditory feedback did not have the same effect as workload on error type, as the frequency counts were similar between the two error groups, largest χ^2, χ^2 (1, N = 18) = 0.89, p = .346 (for Wrong ID error type). However, as can be seen in Table 10, when auditory feedback was present, the Miss error type was approximately three times greater than a Wrong ID (χ^2 (1, N = 26) = 5.54, p = .019; alpha set at .025 to control for familywise error). In the absence of auditory information, no differences between the two groups were evident, χ^2 (1, N = 28) = 1.29, p = .257.

Table 10

The Frequency Count for Miss and Wrong ID Error Types Across Workload and Auditory Feedback Experimental Conditions, and Pilot Experience

Group	Error Type (Miss)	Error Type (Wrong ID)
Experimental Condition		
Low Workload	1.00	3.00
Moderate Workload	11.00	8.00
High Workload	23.60	7.00
No Auditory	16.60	11.00
Auditory	19.00	7.00
Pilot Experience		
Pilots	19.00	6.00
Non-Pilots	16.60	12.00

Note. Count values include one substituted missing data point using the mean individual value, as described in Section 5.2.5.

In terms of previous flying experience, no statistically significant effects were found between the Pilot and Non-Pilot groups in terms of Miss and Wrong ID error type, largest χ^2, χ^2 (1, N = 18) = 2.00, p = .157. However, for pilots the error type was approximately three times more likely to be a Miss than a Wrong ID, χ^2 (1, N = 25) = 6.76, p = .009 (alpha set at .025 to control for familywise error). In comparison for non-pilots, no statistical difference was evident between the two error types, χ^2 (1, N = 29) = 0.86, p = .353.

5.3.3 Spatial Orientation Accuracy (Decision-making Task)

The results of the 3 x 2 x 2 mixed repeated measures analysis for spatial orientation accuracy did not show a significant main effect for Workload, Auditory Feedback, or Flying Experience. As can be seen in Table 11, however, a significant two-way interaction between Auditory Feedback and Workload was evidenced.

Table 11

Summary of the Main and Interaction Effects for Spatial Orientation Accuracy

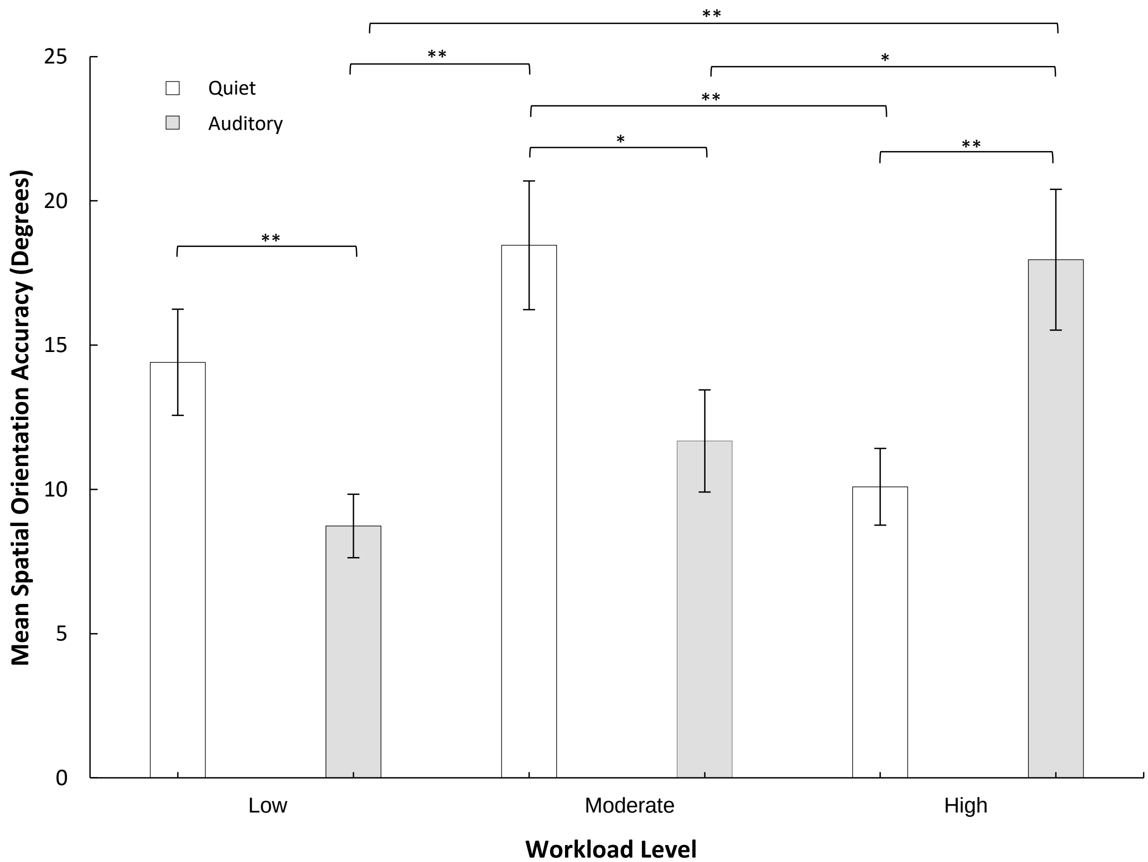
Variables	*MS**	*F*	η_p^2	*p*
Auditory	126.19	1.57	.04	.22
Auditory x Participant	23.40	.29	.01	.59
Workload	232.65	2.09	.06	.13
Workload x Participant	88.96	.80	.02	.13
Auditory x Workload	1198.62	10.82	.24	.000
Auditory x Workload x Flying Experience	67.361	.61	.02	.55
Flying Experience (Pilot and Non-Pilot)	1.65	.01	.00	.93

Note. *MS* refers to 'Mean Square'.

A series of simple effects post-hoc tests, namely paired samples *t* tests revealed three significant effects, one in each of the three workload conditions (i.e., Low workload with and without auditory feedback, Moderate workload with and without auditory feedback and High workload, with and without auditory feedback; smallest *t*, $t(35) = 2.62$, $p = .013$, $r^2 = .16$, for the Moderate workload condition). As can be seen in Figure 11, with the addition of auditory feedback, participants' performance was superior in the Low workload condition (M = 8.73, SD = 6.61) compared to the same condition without auditory feedback (M = 14.41, SD = 10.86). The difference in means of 5.68, 95% CI [1.66, 9.69] represented a medium-large effect (r^2 = .19). A similar effect was also evident under Moderate workload conditions. Auditory feedback enhanced accuracy (M = 11.68, SD = 10.44) when compared with the Quiet condition (M = 18.46, SD = 13.32). The difference in means of 6.78, 95% CI [1.52, 12.04] also represented a medium-large effect (r^2 = .16). However, in the High workload condition, the reverse was found: the presence of auditory feedback adversely affected performance. With a difference in means of 7.87, 95% CI [2.93, 12.82] representing a medium-large effect (r^2 = .16), spatial orientation accuracy was found to be superior in the Quiet condition (M = 10.09, SD = 7.85) when compared to the Auditory condition (M = 17.96, SD = 14.58).

Figure 11

The Mean Spatial Orientation Accuracy in Degrees (Absolute)



Note. The error bars denote standard error of the mean. High values on the vertical axis represent lower accuracy, and vice versa. Note: * = $p < .05$, ** = $p < .01$.

A significant effect was also found between the Moderate and High workload conditions in the absence of auditory feedback (i.e., Quiet condition), $t(35) = 3.40$, $p = .002$ (see Figure 11). Under Moderate workload conditions, spatial orientation accuracy was less accurate (M = 18.46, SD = 13.32) when compared with the High workload (M = 10.10, SD

= 7.85). A mean difference of 8.37, 95% CI [3.37, 13.38] represented a large effect (r^2 = .25). When auditory feedback was introduced, a significant effect was found between the Low and High workload conditions (i.e., each with auditory feedback), $t(35) = 3.28$, p = .002, and between the Moderate and High workload conditions (i.e., each with auditory feedback), $t(35) = 2.57$, p = .015. Conversely, spatial orientation accuracy was superior in the Low workload condition with auditory feedback (M = 8.73, SD = 6.61) when compared to the High workload condition (M = 17.96, SD = 14.58; see Figure 11). A mean difference of 9.23, 95% CI [3.51, 14.95] represented a medium-large effect (r^2 = .23). Similarly, the Low workload condition with auditory feedback was superior compared to the Moderate workload and Quiet condition, evidenced by a significant effect, $t(35) = 4.40$, $p < .001$. Spatial orientation accuracy was also found to be superior in the Moderate workload condition (M = 11.68, SD = 10.44) when compared to the High workload condition in the presence of auditory feedback (see Figure 11). A medium-large effect (r^2 = .16) was represented by a mean difference between the two conditions of 6.28, 95% CI [1.32, 11.24].

5.4 Discussion

Experiment 2 was designed to examine the effect of real-time auditory feedback on participant perception and decision-making task performance. Furthermore, how different workloads and previous piloting experience moderated this effect were also investigated. Auditory feedback was set by each participant to a volume that was comfortable to them. Workload was manipulated via a secondary task, subjectively measured by utilising the NASA-TLX questionnaire, and found to confirm the expected incremental perceived differences of participants between the three workload levels. Participants were separated into two groups, those with previous fixed-wing flying experience in a conventionally piloted aircraft, and those without.

The first research question sought to determine what effect real-time auditory feedback had on participant task performance. In terms of the spotting task (i.e., perception), the results revealed a high level of successful performance, irrespective of the level of workload or participant group. That is, all participants correctly identified and recorded the waypoint mat numbers as the automated flight of the RPA passed overhead each mat, with a high level of accuracy. For the decision-making task (i.e., spatial orientation accuracy), real-time auditory feedback was found to influence performance, evidenced by an interaction with workload. Under Low and Moderate workload conditions without auditory feedback, spatial orientation accuracy was lower when compared to the High workload condition. When auditory feedback was introduced, spatial orientation performance improved in the Low and Moderate workload conditions only, while it deteriorated in the High workload condition. As with the spotting task, no statistically significant differences in performance were noted between the Pilot and Non-Pilot participant groups.

Two of the three hypotheses were confirmed. The presence of real-time auditory feedback did indeed influence task performance and was more effective on the decision-making task than the perception task. Furthermore, task performance levels did deteriorate as workload increased. However, in terms of previous piloting experience, no statistically significant difference was found in task performance between two groups. These findings, from both a theoretical and applied perspective, will be further discussed.

5.4.1 Low Error Rate of Perception Task

The two possible error types for this task were separated into either a Miss or Wrong ID. A Miss error occurred when the participant simply missed recording the waypoint mat. A Wrong ID occurred when the participant incorrectly recorded the waypoint number (e.g., the number ‘7’ was incorrectly recorded as the number ‘2’). Of the total errors that did occur, Miss errors were twice as frequent as Wrong ID errors. Furthermore, both pilots and non-pilots committed more Miss errors than Wrong ID errors. However, a significant difference between error types was found for the Pilot group only. While there are clear differences between error type frequency in these results, the reasons for these can be explained by acknowledging a Miss error requires the participant to simply not see the stimuli which could be due to many reasons (e.g., distraction, boredom, or looking elsewhere at the required moment). However, for a Wrong ID, it needed to be misinterpreted and incorrectly identified. Given the numbers were all 1-digit (i.e., 1 through to 9, and no number 6) and displayed onscreen for an average of six seconds each, it is reasonable to consider this to be a simple task, hence the likelihood of misidentification was less than for a complete miss. While these results could also be reflective of task difficulty, the limited effect of auditory feedback on attention/perception task performance, compared with more cognitively challenging decision-making tasks, is predicted by behavioural models such as the MAM. This is evident in the present findings.

5.4.2 Research Findings in the Context of Behavioural Models

The results from Experiment 2 can be interpreted in relation to the Arousal and MAM theories. Moreover, with respect to both theories, under low and moderate workload, auditory feedback could be regarded as stimulatory (i.e., arousing), thus aiding in performance (i.e., task accuracy in terms of spatial orientation was highest). Low workload with auditory feedback produced the most accurate spatial orientation and possibly reflects an optimum level of arousal (though not statistically different to the Moderate workload and auditory condition). Conversely, moderate workload without auditory feedback produced the least accurate performance out of all six conditions, including the three non-auditory conditions (though not statistically different to Low workload and no auditory condition), thus a possible reflection of participant boredom associated with a low level of arousal and/or a cognitive underload (i.e., hypostress).

As highlighted in Hancock and Warm's (1989) adaptability model of performance and stress, where an underload or overload of work exists (i.e., external stressors), the behavioural adaptability potential of the individual to these external stressors is at its lowest. Prior to reaching the plateau of maximal adaptability, which in the present experiment is represented by the measured task performance, a zone of instability exists whereby slight increments in the level of stress can rapidly raise the adaptability transition level from minimal to maximal. The reverse occurs where slight increments to the stress level reaches the zone of hyperstress, and behavioural adaptability rapidly diminishes from the maximal to minimal levels. The findings from Experiment 2 indeed match the predicted curve. At low and moderate workload, the addition of auditory stimuli has a transforming effect on participant task performance by improving their spatial orientation accuracy during the task. It would appear a behavioural response is induced, requiring adaptation to the additional auditory stressor, which by its location in turn assists in task performance to a point. At a high level of workload, it is the addition of auditory stimuli that appears to push the level of stressor experienced by participants into the region of hyperstress, thereby reducing the level of adaptability to the stimuli and resulting in task performance deterioration.

As was shown by the interaction, there is an inflexion or tipping point where auditory stimuli adversely affects task performance. Because this point occurred under high workload, it is plausible that auditory feedback no longer had an arousing or stimulating effect on participants but was in fact perceived and experienced as noise, and thus became an added stressor. What is unclear, however, is where this tipping point precisely occurred, and how it fits onto the predicted adaptability, and by extension, performance curve. While the three workload levels are each incrementally different, it is unclear if these increments are objectively equidistant. That is, while the subjective assessments obtained via the NASA-TLX can provide an indication of perceived workload, there are limitations to relying solely on this measure (as discussed in Chapter 3). As such, while the results suggest a transition occurs whereby the auditory stimuli changes from useful feedback into perceived noise, the resolution of this change point requires refinement. The results from this experiment in terms of spatial orientation accuracy, therefore, reflect the expected performance based on the MAM, however it is unclear exactly where on the curve they can be superimposed, as well as the extent of this overlay.

5.4.3 The Performance of Pilots and Non-Pilots

The significant results found for spatial orientation task performance occurred uniformly across the two participant groups (i.e., Pilot and Non-Pilot). All participants in this experiment were required to perform monitoring tasks that are reflective of an automated RPA BVLOS visual flying condition. The use of automation is congruent with similar tasks performed by traditional pilots in conventional aircraft, except for the reduced sensory cueing available. As hypothesised, it would be reasonable to expect a group (i.e., Pilot) with previous flying experience, and a demonstrated aptitude in performing similar tasks would be advantaged versus a group without this experience (i.e., Non-Pilot). However, the lack of statistically significant results between the two groups in this experiment show this is not the case. While it is not clear whether the measured responses are due to the stressors of both workload and auditory stimuli exceeding individuals' cognitive resources or due to the overarousal effect, it is possible both explanations may be suitable given the two different participant groups. That is, the mechanism responsible for the decline in spatial orientation accuracy task performance in each participant group may differ. However, given the methods utilised to measure task performance in Experiment 2, it is not possible to ascertain this with certainty.

5.4.4 Applied Research Findings

From an applied perspective, the findings from this experiment are important to the field of RPAS in terms of the design of future systems. These findings support the inclusion of additional sensory systems, and have implications for the recruitment, selection and training of the individuals operating the RPA systems, and regulatory bodies assessing the risk involved with BVLOS operations in a civil setting. The moderating effect of previous piloting experience was not evident in this experiment, which is an important finding. The BVLOS visual operating conditions which this experiment attempted to replicate are considered highly challenging by aviation regulatory authorities (i.e., CASA, FAA, CAA), and therefore can only be performed by skilled and appropriately trained remote pilots (e.g., commercial, or military pilots with an Instrument Flight Rule (IFR) rating). Assuming the mechanism involved in the non-significant difference between the two participant groups are the same, this requirement, at least for operations requiring minimal control inputs and/or interaction between the remote pilot and RPA, is not supported by the findings.

There are also important applications in professions beyond the traditional RPAS field, including other teleoperations, such as Urban Air Mobility (Fu et al., 2019) which seeks to utilise vertical take-off and landing (VTOL) aircraft to transport people, cargo and medical equipment/inventory, over relatively short distances in and around city and suburban environments (Koo et al., 2022). Similarly, workplaces transitioning to automated and/or remotely operated practices could also benefit from these findings, particularly when considering the training and recruitment of personnel required to teleoperate the equipment.

Furthermore, adding auditory stimuli to tasks which are potentially boring or monotonous may increase individuals' performance, particularly at critical times if the stimuli is targeted, or expected changes to perceived workload occur as the task progresses. Recognising the adverse effect of auditory stimuli when combined with high workload is also important, and thus in such situations where a change in workload level is unanticipated, individuals should be educated about these effects and provided an opportunity to eliminate the stimuli. Similarly, where changes to workload are anticipated, the introduction of dynamic and adaptive auditory stimuli could also be warranted. In terms of RPAS, the developers of future systems should consider the inclusion of real-time auditory feedback (i.e., stressor) from teleoperated equipment that can be linked to the stage of flight or workload experienced by the human operator. For operations that are varying in complexity, or during periods of changing workload, consideration could also be given to the addition of an auditory component for automated tasks that is also adaptable.

5.4.5 Limitations and Future Research

Like the first experiment, the results from the second experiment need to be interpreted with a level of caution. This experiment focused on a series of simulated, automated RPA flights, under BVLOS operating conditions and required no manual takeover of flying control. Therefore, while two aspects of BVLOS operations were tested (spotting accuracy and spatial orientation accuracy), the manual flying skills of each participant were not. While the experimental design using simulation offered several advantages including the ability to test non-pilot participants without the need for regulatory approvals, as well as controlling the experimental environment consistently across each participant, how this compares to the real world is unknown. Future research under real-world conditions would therefore be beneficial. A real-world experiment involving non-pilots operating RPA (and under BVLOS) may be difficult to achieve from a regulatory perspective, however in an enclosed large outdoor environment, where access to such facilities is feasible, there may be some benefit in determining their task performance abilities.

The Pilot and Non-Pilot groups showed no statistically significant differences in task performance. However, it is still unknown how previous traditional flying experience and the pressures associated with external, commercial operating conditions contribute to task performance. Although a lack of observed differences between the two user groups may not extend directly to some real-world settings (e.g., in non-segregated airspace; a future commercial environment, and currently beyond the scope of most current civil RPAS users), it is possible differences between pilots and non-pilots may exist. For example, variations between the two groups could lie in aeronautical judgement and airmanship as opposed to monitoring and physical manipulation tasks. Hence, this is an avenue for future research. Nevertheless, using the results from this experiment as a baseline and determining whether differences in performance occur in a commercial real-world setting, these findings have the potential to contribute to the design of future RPAS training. Furthermore, regulatory changes when higher usage of automation are implemented, may also be warranted. For non-pilots, this could be a substantial change, given at present the many barriers to entry including financial, the requirement for extensive previous aviation experience, and special licencing.

This experiment was also relatively short in duration in comparison to traditional flying, and long-endurance RPAS BVLOS operations. As RPAS technology progresses in terms of flight endurance and performance, it would be valuable to understand the effects of dynamic sensory cueing (e.g., adaptive auditory feedback) linked to the temporal workloads of long-duration BVLOS flights on human performance. How an extended task duration, and the inclusion of other environmental variables such as those relating to visual degradation through meteorological or night conditions may affect performance is also an area for future research. However, as with other fields of research, limiting factors include access to resources. In the context of RPAS and human performance, these limitations are even more acute when regulatory limitations are included.

Experiment 1 and Experiment 2 each measured the effect of auditory feedback and workload on participant task performance. The results obtained from these experiments broadly reflect the predicted values of task performance in the context of behavioural models including the MAM. However, the transition or inflection points whereby performance declines in conjunction with additional stressors (i.e., auditory feedback, and/or workload) is not at sufficient resolution to confidently superimpose on the model. There are at least three possible reasons for this: 1) the changes to the variables manipulated are not at sufficient increments; 2) the context of remote pilots and RPAS involves performing tasks in a sensory deprived environment, which could influence how stressors are perceived and responded to by participants; and 3) a combination of reason 1 and reason 2. Hence, Experiment 3 will seek to further expand upon the findings previously determined, and specifically retain the methodology of Experiment 2.

The next experiment will also introduce an element of adaptive auditory cueing, by maintaining three workload levels, but increasing auditory feedback by two additional levels, pre-determined individually by each participant. Task performance relative to the predicted values of the MAM will be further explored by using incremental changes above and below a comfortable volume setting. Furthermore, an understanding of the individual subjective strategies used to deal with the presence of auditory feedback, and specifically at sound intensities which are considered as either beneficial information or noise, will be investigated.

5.5 Summary

In this chapter, the design and results of Experiment 2 were described. Specifically, this experiment sought to determine the effect of real-time auditory feedback on remote pilot spotting (i.e., perception) and spatial orientation (i.e., decision-making) task performance, between two participant groups: Pilots and Non-Pilots. While the mechanism responsible for the measured task performances is unclear (i.e., auditory feedback perceived as noise, and exceeding individuals' cognitive resources or creating an overarousal effect), the results build upon previous findings related to workload and auditory feedback. Specifically, the findings are relevant to remote pilots operating in sensory deprived environments, and more broadly others utilising teleoperated systems. Three workload levels were introduced to further measure the effect on task performance. Two additional participant groups were also measured to determine if task performance is moderated by previous piloting experience, however no statistically significant effects were found between the Pilot and Non-Pilot groups.

This experiment further expanded upon the validity of the Maximal Adaptability Model and Arousal Theory in predicting the associated level of performance when external stressors are introduced, with reference to remote pilots operating RPAS under automated BVLOS visual flight conditions. Workload and auditory feedback were found to have influenced performance of the perception task in terms of error types committed, with Miss errors more frequent than Wrong ID errors in occurrence. However, decision-making task performance was benefited by the stimulatory effect of real-time auditory feedback from the RPA under low and moderate workload only. Moreover, under high workload conditions, a tipping point in participant task performance occurred. Auditory feedback affected task performance evidenced by degraded spatial orientation accuracy, a result consistent with the detrimental effects of noise (i.e., an additional stressor). The results from this experiment suggest performance is broadly aligned with behavioural models of adaptability and will be discussed further in the discussion in the final chapter (Chapter 7).

CHAPTER 6 – EXPERIMENT 3: COMBINED WORKLOAD AND INCREMENTED AUDITORY FEEDBACK EFFECTS ON TASK PERFORMANCE

6.1 Introduction

The results of Experiment 1 revealed superior task performance when a monitor display was used under BVLOS conditions. Furthermore, under higher workload and with auditory feedback (i.e., Task B), horizontal flying accuracy also improved. The results of Experiment 2 revealed superior task performance (as measured through spatial orientation error accuracy (i.e., decision-making task)) under low and moderate workloads with a comfortable level of auditory feedback, when compared to the Quiet (i.e., ambient) auditory conditions. The results from both experiments broadly reflected the predicted values of the Maximal Adaptability Model (Hancock & Warm, 1989). However, the two auditory feedback conditions were limited in range (i.e., Quiet versus 'comfortable'). Therefore, to better understand the effects of stressors on performance, and to identify where auditory feedback transitions more precisely to noise, further increments of auditory volume need to be measured.

In both Experiment 1 and 2, participants controlled the volume at which the auditory feedback was presented. They did this based on the instructions that asked them to set it at a volume they considered comfortable. From an applied perspective, this may not always be possible. Therefore, in Experiment 3, 'dynamic' auditory feedback on remote pilot task performance will be examined. As discussed in Chapter 2, auditory feedback can provide useful information, aiding the operator. It can also be unwanted, interpreted as noise, thereby hindering performance. Understanding the effect of ‘dynamic' auditory feedback on RPAS operator performance is important, as well as what, if any strategies individuals employ to deal with variations in dynamic auditory feedback. Hence, three research questions and one hypothesis were proposed for this experiment.

6.1.1 Research Questions

1. What is the effect of +/-10 dBA (i.e., incremental) from a comfortable level of auditory feedback on participant perception and decision-making task performance (i.e., waypoint spotting accuracy and spatial orientation)?
2. What are the effects of combined workload and incremental auditory feedback on participant task performance?
3. Does the auditory management style employed by participants mediate the effect of auditory feedback and workload on decision-making task performance (i.e., spatial orientation accuracy)?

6.1.2 Hypothesis

1. Task performance (i.e., waypoint spotting accuracy and spatial orientation) in the Baseline condition (i.e., consisting of an average of the two best auditory conditions from Experiment 2 - Low and Moderate workload under Comfortable auditory feedback conditions) will be superior to the Moderate Workload (Soft), Moderate

Workload (Loud), High Workload (Comfortable), High Workload (Soft) and High Workload (Loud) workload-auditory paired task conditions.

6.2 Method

6.2.1 Participants

The results from Experiment 2 strongly influenced the design of Experiment 3. Since there were no statistically significant differences in task performance between the two participant groups (i.e., Pilots and Non-Pilots) in Experiment 2, all participants recruited for Experiment 3 were non-pilots. Twenty-four (eleven female) participants were recruited from the UNSW Sydney main campus. Demographic information regarding participants can be found in Table 12. As can be seen in this table, the mean age of the participants were 24.75 (SD = 4.07) years, and the mean hours of RPAS experience was 2.72 (SD = 8.46). The research, including stimuli and procedure was approved in advance by the UNSW Sydney Ethics Panel (HC Number: 210309). All participants signed a consent form prior to commencing the experiment (Appendix 4). The research took approximately one and a half hours to complete per participant, for which participants were provided a $30 large department store gift voucher.

Table 12

Demographic Information of the Sample Group

Descriptive Variable	*M*	*SD*	*Range*
Age (Years)	24.75	4.07	19.00 – 32.00
Total Flying Hours	0	0	0
Total RPAS Experience (Hours)	2.75	8.46	0 – 30.00

Note. *M* refers to 'Mean'. *SD* refers to 'Standard Deviation'. N = 24.

6.2.2 Design

Experiment 3 comprised a 3 x 4 repeated measures design. The first repeated measures factor labelled Workload contained three levels: Low, Moderate and High. The second repeated measures factor labelled Auditory Feedback contained four levels: the first level was determined by the participant at a volume setting that they deemed to be comfortable with (i.e., labelled Auditory-Comfortable), the second level was 10 dBA above the comfortable level (i.e., labelled Auditory-High), the third level was 10 dBA below the comfortable level (i.e., labelled Auditory-Low) and the fourth level was without auditory feedback (i.e., ambient; labelled Quiet). To reduce a possible learning effect, the stimuli comprising twelve conditions (Workload x Auditory feedback conditions) were presented in a balanced Latin-square design ensuring each task and video pairing were unique for each participant. Dependent variables for each task were consistent with Experiment 2 (i.e., waypoint spotting accuracy and spatial orientation accuracy), however with one small difference. Due to limited resource availability, way point mat numbers were recorded via pen and paper instead of iPad, as was previously used. A summary of the experimental design is presented in Table 13.

Table 13

Overview of Experimental Design

<table>
<tr><th rowspan="2">Dependent Variables</th><th colspan="2">Experimental Condition</th></tr>
<tr><td>Factor 1 (Workload)</td><td>Factor 2 (Auditory Feedback)</td></tr>
<tr><td rowspan="3">Waypoint Spotting Accuracy (Correct Recordings)

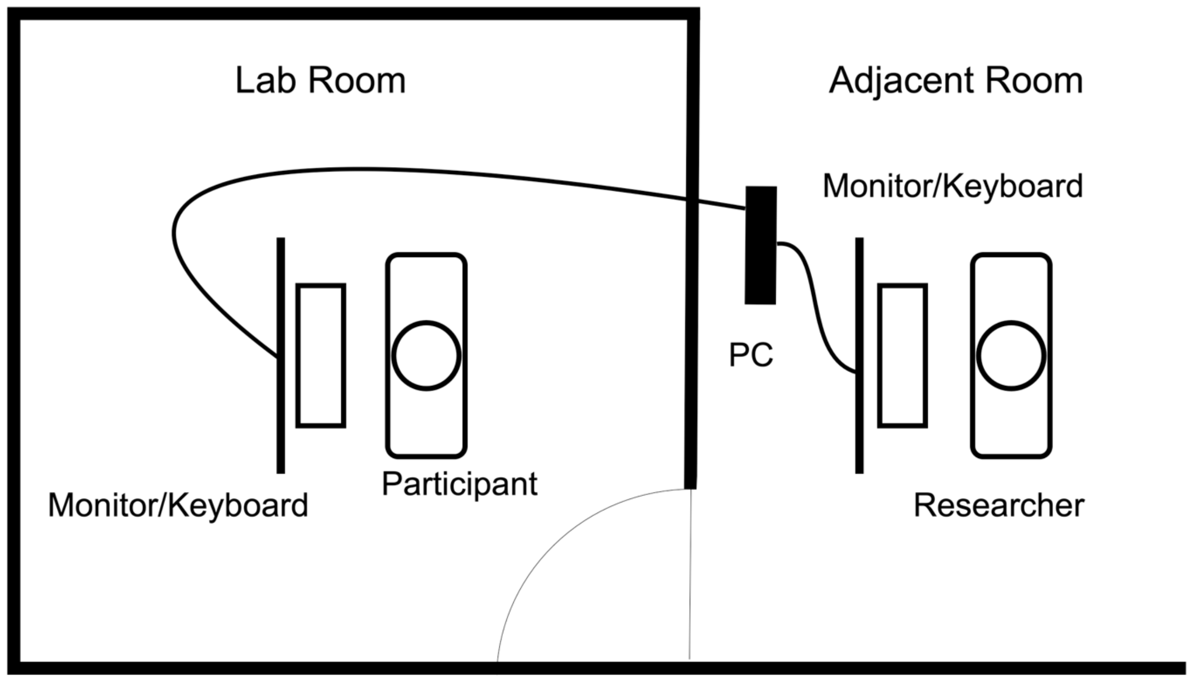
Spatial Orientation Accuracy (Degrees)</td><td>Workload (Low)</td><td>Auditory – Comfortable
Auditory – High
Auditory – Low
Quiet</td></tr>
<tr><td>Workload (Moderate)</td><td>Auditory – Comfortable
Auditory – High
Auditory – Low
Quiet</td></tr>
<tr><td>Workload (High)</td><td>Auditory – Comfortable
Auditory – High
Auditory – Low
Quiet</td></tr>
</table>

6.2.3 Stimuli and Materials

Twelve pre-recorded videos representing a series of automated RPAS flights were presented to the participants. The apparatus comprised: a DJI (Dà-Jiāng Innovations) Phantom 4 Professional Obsidian model quadcopter (for recording of videos, utilising on-board 20 mega-pixel, 84° FOV camera), a Samsung 10 Galaxy telephone (to record auditory feedback in four phases of flight: take-off, stationary/hovering, turning (yaw), and forward movement), the Movavi Video Editor Plus 2020 software (for pairing auditory feedback and video imagery), one IBM compatible computer, one computer mouse, two keyboards and two HP 23-inch (1920 x 1080 resolution) high-definition monitors (model LA2306x; i.e., for displaying pre-recorded videos; the first monitor was used by the participant and the second by the researcher in a mirrored mode, located in an adjacent room enabling the loading of videos out of sight of each participant), DT 770 Pro headphones (worn by the participant only in the auditory feedback conditions), and twelve pre-recorded videos depicting an automated RPAS flight over eight waypoints (i.e., the main experimental task; see Experiment 2 - Chapter 5). The research was conducted at the UNSW Sydney main campus, with a physical wall separating the researcher and participant (Figure 12).

Figure 12

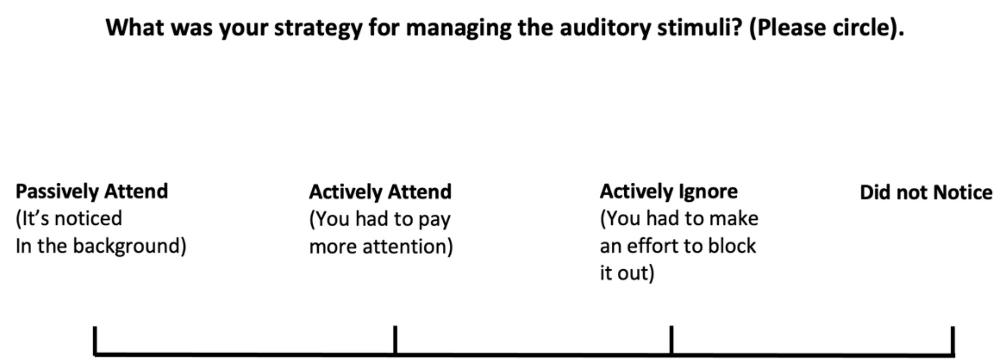
A Diagram of the Experimental Area Setup (Top-down View)



The materials comprised: a paper copy of the experiment instructions, a scale map of the flying area, a demographic questionnaire (i.e., age, flying experience, gender), the Serial 3s and Serial 7s (for Moderate and High workload; see Chapter 5), the NASA-TLX questionnaire (unweighted; see Appendix 2), a reference table to determine correct volume level for the auditory feedback conditions (see Appendix 5), and an Auditory Management Strategy questionnaire. The Auditory Management Strategy questionnaire contained a single question, where participants were asked to state their strategy employed to manage the auditory stimuli. Four potential strategies were presented, as can be seen in Figure 13. Participants had to circle the strategy they employed.

Figure 13

The Auditory Management Strategy Questionnaire



6.2.4 Procedure

Participants were informed about the research through online internal UNSW advertisements, broadcast emails and noticeboard paper flyers located throughout the university. Interested participants contacted the researcher, and a mutually suitable time was arranged to complete the research. Prior to undertaking the main experiment, all participants received a written and verbal briefing about the experiment. Following the briefing, and signing of the consent form, participants were provided instructions (in paper form) about the main experiment and were asked to complete the supplied demographic questionnaire. A shortened trial task representing the main experiment (duration = 1 minute 30 seconds) was provided to each participant to demonstrate proficiency using the apparatus and to acknowledge the instructions for the main experiment. Four of the twelve experimental conditions involved a low level of workload (Low). Four involved a moderate level of workload (Moderate) and Four involved a high level (High), utilising the Serial 3s and Serial 7s procedures respectively, and performed concurrently with the primary spotting task. Prior to the main experiment, baseline readings were obtained for the Moderate and High workload conditions and involved verbally counting backwards in 3s for two minutes, followed by 7s for two minutes from pre-defined starting numbers. The Researcher recorded on paper the verbal numbers stated by the participant. Following this, the twelve videos were presented in the counterbalanced experimental order. The primary task of spotting and recording waypoints was performed once under each of the twelve experimental conditions. The experimental procedure was the same as performed in Experiment 2 (see Chapter 5).

In the Auditory-Comfortable condition (i.e., with real-time sound from the RPA propellers), the headphone volume was adjusted by each participant to suit, determined prior to the main experiment commencing. In the Auditory – Low, and Auditory – High conditions, the volume was adjusted by the researcher, utilising a pre-defined table (see Appendix 5), and based upon the individual pre-determined comfortable sound level. Under the Low auditory setting, the volume was decreased by 10 dBA below the Auditory – Comfortable level. Similarly, under the High auditory setting, the volume was increased by 10 dBA above the Auditory – Comfortable level. The 10 dBA value was elected based on the results of a pilot study with five participants, which revealed noticeable audible differences, instead of a +/-5 dBA level (Fastl & Zwicker, 2007). After each of the twelve tasks, participants completed the subjective workload measure (i.e., NASA-TLX questionnaire), consistent with Experiment 2. The Auditory Management Strategy questionnaire was then completed.

Ambient noise levels in the room ranged from 39-42 dBA (i.e., test apparatus included sound level meter with plastic ear cup mould with O-ring seal; see Burgess & Molesworth, 2016). In all experimental conditions, participants were provided a seat which is ecologically congruent with real world operations.

6.2.5 Data Analysis

An a priori analysis using G*Power 3.1 determined a minimum of 19 participants was sufficient to detect a medium to large effect (i.e., Cohen's (1988) criteria). Power was calculated using the following assumptions: repeated measures, effect size .20, alpha error probability = 0.05, power = 0.8, number of measurements = 12. Since there were twelve experimental conditions, sample size featured in multiples of twelve, and therefore, 24 participants in total. Four dependent variables featured in each task condition:

- Subjective NASA-TLX Score,
- Waypoint Spotting Accuracy,
- Spatial Orientation Accuracy, and,
- Subjective Auditory Management Strategy (and relationship to Spatial Orientation Accuracy).

The NASA-TLX featured as a manipulation check, consistent with Experiment 2, to determine if task workload and auditory feedback influenced the subjective level of perceived workload by participants using a 3 (workload) x 4 (auditory feedback) repeated-measures analysis. A Chi-square analysis was used to analyse the Waypoint Spotting Accuracy in terms of error type, while a series of Fisher's Exact tests (due to some task conditions with very low error counts, i.e., ≤ 5) were used to analyse error type (i.e., Miss Error or Wrong ID; refer to Chapter 5) based on the level of Auditory Feedback and Workload, and between the Baseline and other task condition pairings.

Because the results from Experiment 2 for spatial orientation accuracy also informed the design of the present experiment, a series of a-priori planned comparisons were made. The first planned comparison was a manipulation check to ensure that spatial orientation accuracy performance in the Low Workload (Comfortable) and Moderate Workload (Comfortable) groups were similar (i.e., no statistically significant difference), and thus consistent with Experiment 2. Following this confirmation, data from these two groups were collapsed to form a single Baseline group. Significant differences were expected to be found between the Baseline group and the five other auditory conditions with Moderate and High workloads. The remaining five task conditions (i.e., Low workload, and other Quiet task conditions) were included in the analysis, however no statistically significant effects were expected to be found. The expected significant and non-significant differences between the Baseline condition and all other conditions for spatial orientation accuracy (i.e., decision-making task) are shown in Table 14.

Table 14

Ten Baseline-Group Pairings for Expected and Not Expected Statistically Significant Differences in the Planned Comparisons Analysis

Significant Differences Expected	**Significant Differences Not Expected**
Baseline and: Moderate Workload (Soft) Moderate Workload (Loud) High Workload (Comfortable) High Workload (Soft) High Workload (Loud)	Baseline and: Low Workload (Quiet) Low Workload (Soft) Low Workload (Loud) Moderate Workload (Quiet) High Workload (Quiet)

Finally, a Kruskal-Wallis and series of Mann-Whitney tests (post-hoc test) were used for measuring the auditory management strategy (and in addition, the relationship to spatial orientation accuracy) variables. Prior to all analyses, the data was screened for violations of the assumptions of homogeneity of variance using the interquartile range technique suggested by Moore et al. (2009). If outliers were present, the data was transformed using the next highest or lowest score plus one technique as suggested by Tabachnick and Fidell (2013). This was stated when occurred. Alpha was set at 0.05 for all analyses. Pre-data screening revealed one missing data point which was managed by substituting the missing data with the mean individual value for that participant, as recommended by Tabachnick and Fidell (2013). The complete dataset was then screened for outliers and transformed where identified, leaving the dataset homogeneous.

6.3 Results

6.3.1 Subjective Workload

The first analysis (i.e., a manipulation check) sought to confirm the effect of Workload and Auditory feedback on participants' subjective level of perceived workload. Mean total scores on the NASA-TLX were compared between all conditions using a 3 x 4 repeated-measures ANOVA. The results revealed two main effects, one for Workload ($F(2, 46) = 108.85$, $p < .001$, $\eta^2 = .83$), and one for Auditory Feedback ($F(3, 69) = 14.30$, $p < .001$, $\eta^2 = .38$). There were no interaction effects. Table 15 shows a summary of the NASA-TLX data distributed across groups.

Table 15

A Summary of the Mean Participant NASA-TLX Values (Unweighted; in %) for Perceived Workload

Independent Variable	*M*	*SD*	*Range*
Workload			
Low	20.73	16.19	0.00 – 64.17
Moderate	40.59	18.10	5.00 – 85.00
High	48.20	17.51	9.17 – 90.83
Auditory Feedback			
Quiet	32.14	20.14	0.00 – 82.50
Soft	37.77	21.74	2.50 – 90.83
Comfortable	36.18	19.56	2.50 – 80.83
Loud	39.94	21.10	2.50 – 86.25

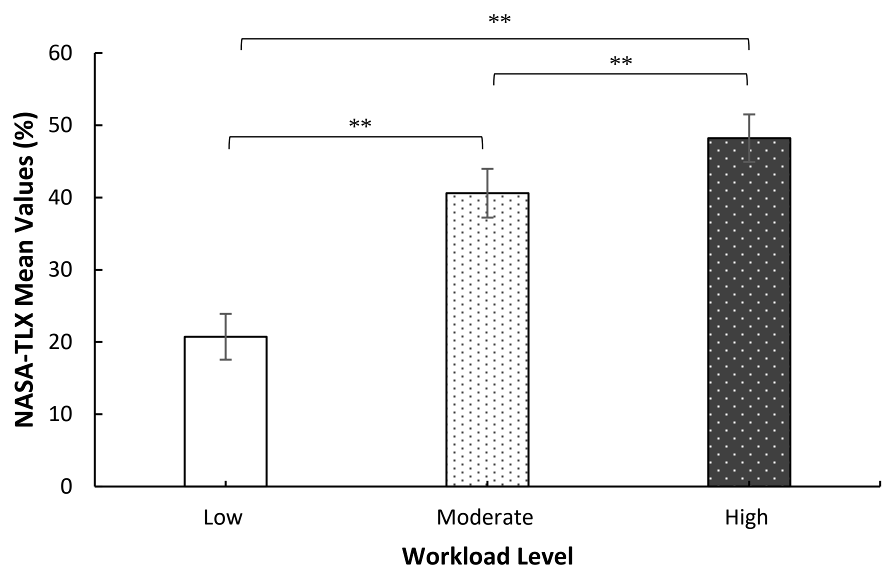
Note. *M* refers to ‘Mean’. *SD* refers to ‘Standard Deviation’. N = 24.

Post hoc tests revealed statistically significant differences between all three levels for workload ($p < .001$; largest difference between the Low and High workload conditions), with results shown in Figure 14a. For the auditory condition, the post hoc tests revealed significant differences between four of the six auditory level combinations. Statistically significant differences were found between the Quiet group and all three other auditory groups, and between the Comfortable and Loud groups (+10 dBA). Significant effects were not reached between the Soft (-10 dBA) and Comfortable, and Soft (-10 dBA) and Loud groups (+10 dBA).

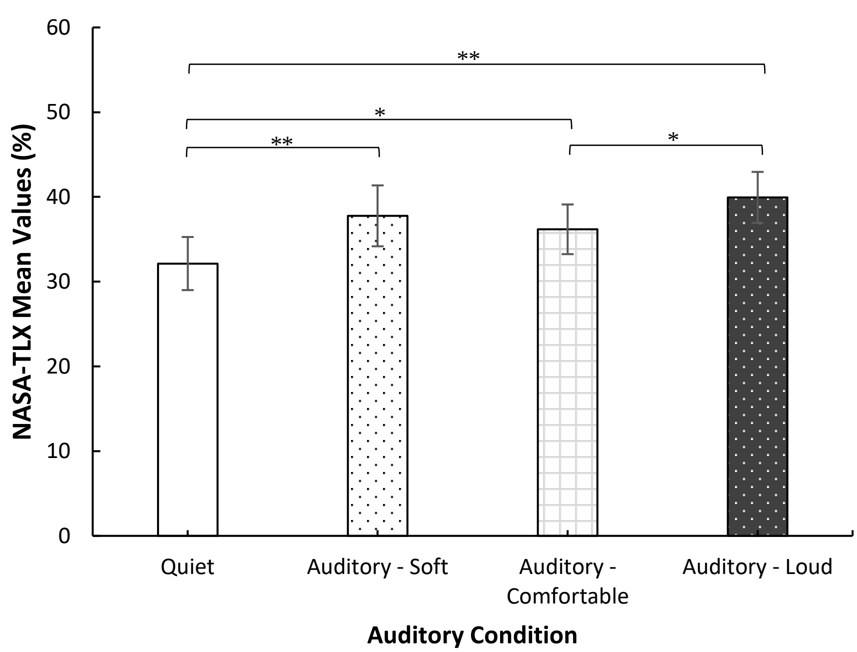
Figure 14

The Mean NASA-TLX Values (Unweighted) for a) Workload and b) Auditory Condition

a)



b)



Note. The error bars denote standard error of the mean. Note: * = $p < .05$, ** = $p < .01$.

6.3.2 Waypoint Spotting Accuracy (Perception Task)

The waypoint spotting accuracy task was completed by most participants with a high level of accuracy. That is, 232 out of 288 possible tasks (i.e., 12 tasks per participant x 24 participants) were completed with no waypoint recording errors. Consistent with Experiment 2, and since errors in this task could result from a 'miss' or 'wrong identification', an analysis was first conducted to examine if one type of error was more common. With alpha set at .05, the one-way Chi-square analysis failed to reveal a significant difference between groups (i.e., Miss vs. Wrong Identification (Wrong ID)), χ^2 $(1, n = 55) = 3.07$, $p = .08$.

The next analysis examined the effect of Auditory Feedback and Workload on the two error types. Because the expected frequency count was ≤ 5 in several task conditions, a series of Fisher's exact tests were employed. The results failed to reveal any significant differences for Auditory Feedback type (lowest *p*, $p = .05$ between Auditory - Quiet and Auditory – Comfortable), or Workload (lowest *p*, $p = .36$ (between Low Workload and Moderate Workload). A summary of the frequency counts for waypoint recording error types can be found in Table 16.

Table 16

The Frequency Count for Miss and Wrong Identification Error Types Across Workload and Auditory Feedback Experimental Conditions

Group	Waypoint Error (Miss)	Waypoint Error (Wrong ID)
Auditory Condition		
Quiet	10.10	3.00
Soft	10.00	5.00
Comfortable	5.00	8.00
Loud	9.00	5.00
Workload		
Low	0.00	1.00
Moderate	13.10	7.00
High	21.00	13.00

Note. Count values include one substituted missing data point using the mean individual value, as described in Section 6.2.5.

The final analyses for waypoint spotting accuracy sought to determine differences between the Baseline and the 10 other task conditions in terms of error type. A series of Fisher's exact tests were again utilised due to the expected frequency count being ≤ 5 for several task conditions (i.e., they were successfully completed with no errors). All tests (i.e. 55 possible pairs) failed to reach statistical significance (lowest *p*, $p = .23$ between the Baseline and Moderate Workload (Quiet) conditions; see Table 17). These results will be discussed in the Discussion.

Table 17

The Frequency Count for Waypoint Error (Miss) and Waypoint Error (Wrong ID) Error Types Across Each Task Condition

Task Condition	Waypoint Error (Miss)	Waypoint Error (Wrong ID)
Baseline	1.00	2.00
Low Workload (Quiet)	0.00	0.00
Low Workload (Soft)	0.00	0.00
Low Workload (Loud)	0.00	0.00
Moderate Workload (Quiet)	4.10	1.00
Moderate Workload (Soft)	3.00	2.00
Moderate Workload (Loud)	4.00	1.00
High Workload (Quiet)	6.00	2.00
High Workload (Comfortable)	3.00	4.00
High Workload (Soft)	7.00	3.00
High Workload (Loud)	5.00	4.00
Total	33.1	19.00

Note. Count values include one substituted missing data point using the mean individual value, as described in Section 6.2.5.

6.3.3 Spatial Orientation Accuracy (Decision-making Task)

The manipulation check examining differences in spatial orientation accuracy between the Low Workload (Comfortable) and Moderate Workload (Comfortable) conditions failed to reveal any statistical differences, as predicted ($t(23) = 1.55$, $p = .14$). As stated prior, the data from these two groups were combined and featured as the ‘Baseline’ group. Having effectively established the Baseline condition, reflecting the results of Experiment 2 (see Chapter 5 and Dunn et al., 2022), the next analyses involved the series of a priori planned comparisons. The results revealed a statistically significant result for unequal variances between the Baseline and all other workload-auditory groups (i.e., Levene’s test, $F(10, 253) = 4.82$, $p < .001$), hence the results assuming unequal variance were used. Significant effects were found between the Baseline and Moderate Workload (Soft) groups $t(28.11) = 2.26$, $p = .032$, the Baseline and the Moderate Workload (Loud) groups, $t(28.80) = 2.29$, $p = .030$, and the Baseline and High Workload (Loud) groups, $t(29.74) = 2.20$, $p = .036$. A summary of the mean values, and *t* statistics between the Baseline and all other task conditions are shown in Table 18.

Table 18

Summary of Spatial Orientation Accuracy Means (degrees) for the Baseline and Other Task Conditions

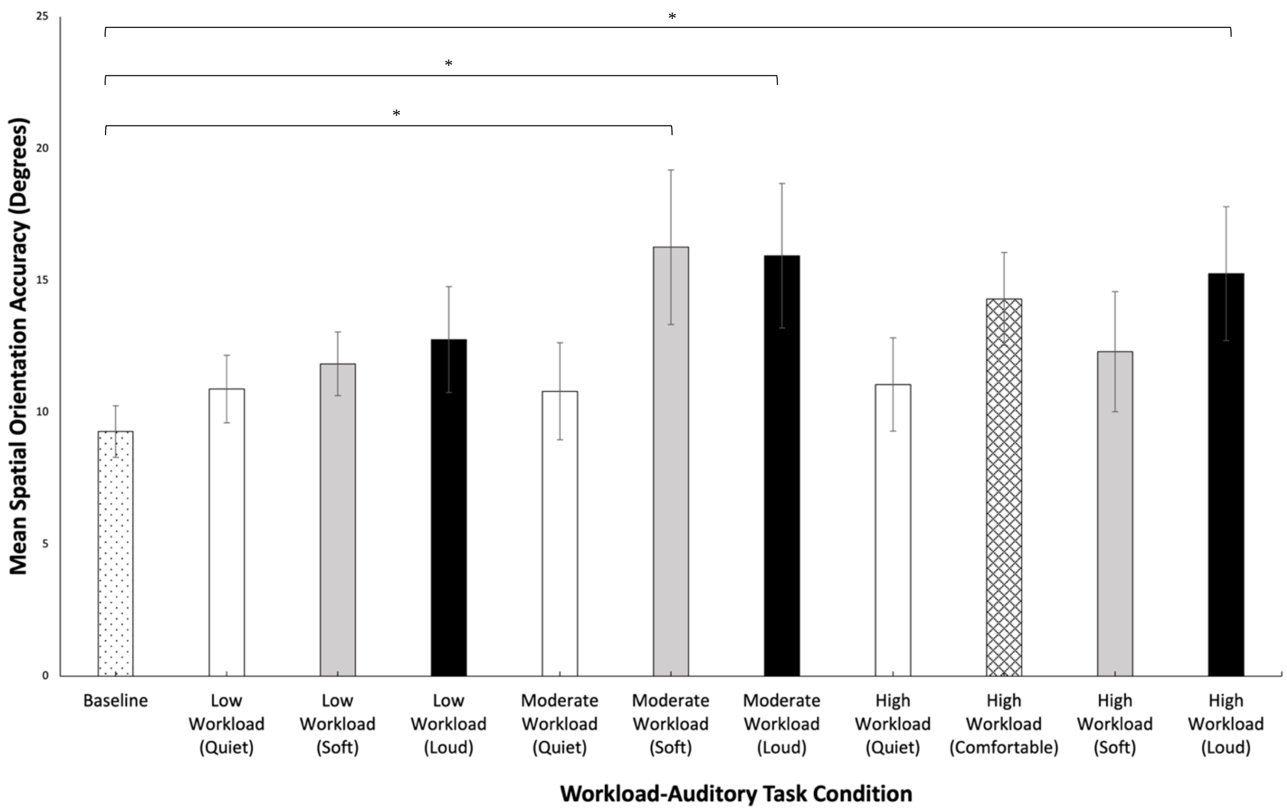
Task Condition	*M*	*SD*	*t*
Baseline	9.28	4.81	-
Low Workload (Quiet)	10.89	6.28	1.00
Low Workload (Soft)	11.84	5.93	1.65
Low Workload (Loud)	12.76	9.86	1.56
Moderate Workload (Quiet)	10.80	9.00	0.73
Moderate Workload (Soft)	16.27	14.34	2.26*
Moderate Workload (Loud)	15.94	13.45	2.29*
High Workload (Quiet)	11.06	8.66	0.88
High Workload (Comfortable)	14.30	11.19	1.50
High Workload (Soft)	12.30	8.65	2.02
High Workload (Loud)	15.26	12.43	2.20*

Note. Note: * = $p < .05$, ** = $p < .01$. *M* refers to 'Mean'. *SD* refers to 'Standard Deviation'. N = 24. The *t* values are between the Baseline and specific Task Condition.

While the two further a priori comparisons (Baseline and High Workload (Soft); Baseline and High Workload (Comfortable)) failed to reach statistically significant levels, highest *t*, $t(31) = 2.02$, $p = .052$ (Baseline and High Workload (Comfortable)), their means were each higher than the Baseline condition indicating a numerically reduced performance. These results are further highlighted in Figure 15.

Figure 15

The Mean Spatial Orientation Accuracy in Degrees (Absolute) for Each Workload-auditory Task Condition Compared with the Baseline



Note. The error bars denote standard error of the mean. Note: * = $p < .05$, ** = $p < .01$.

6.3.4 Auditory Management Strategy

The final analysis investigated the auditory management strategy used by participants and how this contributed to spatial orientation accuracy performance. Only tasks with auditory feedback were included for analysis (i.e., the three Quiet task conditions did not require a strategy to deal with the nil auditory feedback). Because the four potential management strategies that could be utilised by participants in each task condition were unequal, a Kruskal-Wallis nonparametric test was used for the Baseline and seven other auditory groups. No significant results were revealed for each of the four auditory management strategies (Passively Attend, χ^2 (7, N = 94) = 5.82, p = .56, η^2 = .06; Actively Attend, χ^2 (7, N = 74) = 2.49, p = .93, η^2 = .03; Actively Ignore, χ^2 (7, N = 32) = 5.54, p = .60, η^2 = .18; and Did Not Notice, χ^2 (7, N = 16) = 8.27, p = .22, η^2 = .55). A summary of these results is shown in Table 19.

Table 19

Summary of Spatial Orientation Accuracy Means (degrees) and Counts for the Baseline and Other Task Conditions

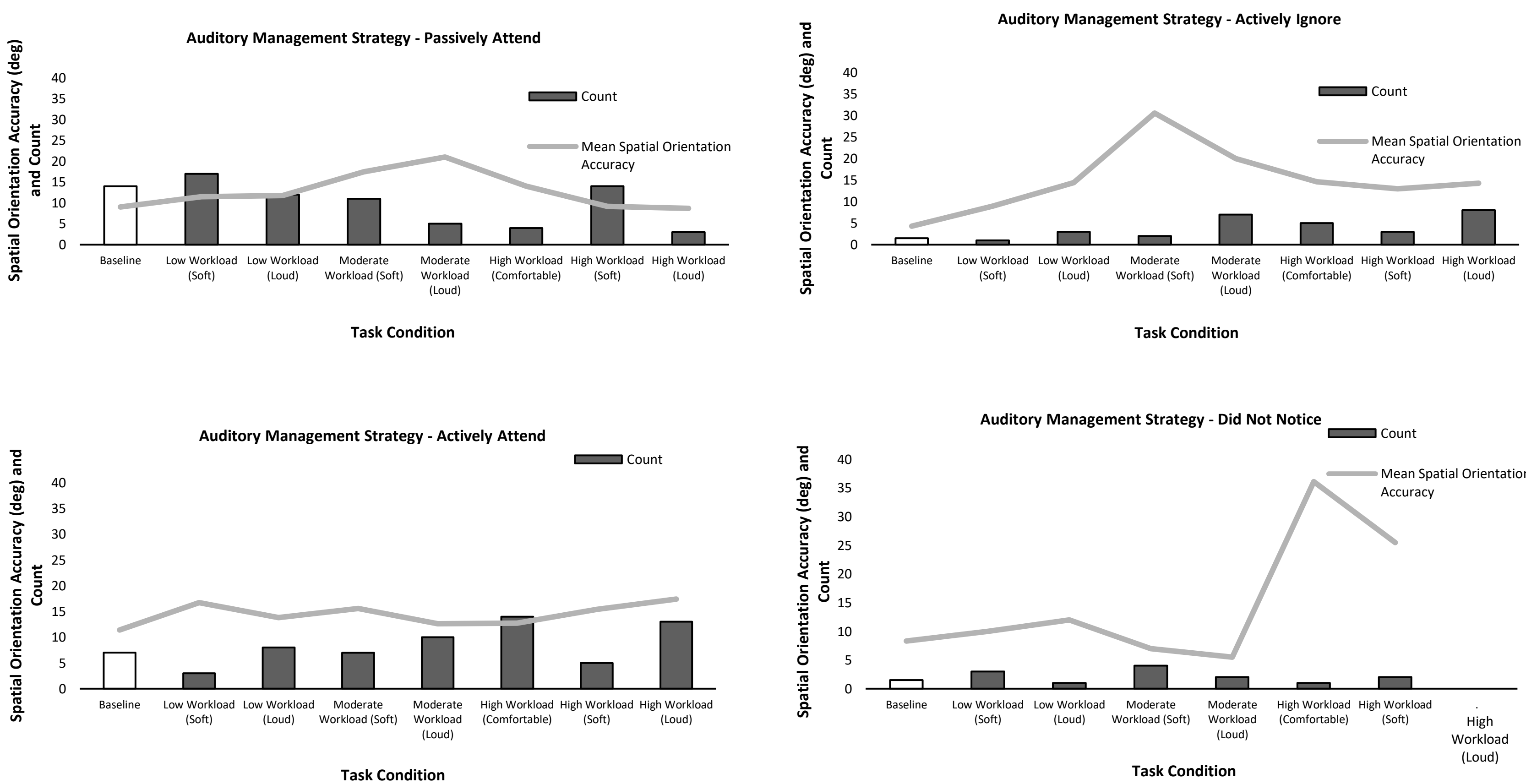
	Passively Attend			*Actively Attend*			*Actively Ignore*			*Did Not Notice*		
Task Condition	*M*	*SD*	*Count*	*M*	*SD*	*Count*	*M*	*SD*	*Count*	*M*	*SD*	*Count*
Baseline	9.04	6.97	**14.00**	11.37	8.88	**7.00**	4.33	1.15	**1.50**	8.33	3.51	**1.50**
LW (Soft)	11.47	5.05	**17.00**	16.70	7.53	**3.00**	9.00	-	**1.00**	6.50	9.90	**3.00**
LW (Loud)	11.75	9.76	**12.00**	11.75	9.75	**8.00**	14.37	15.75	**3.00**	12.00	-	**1.00**
MW (Soft)	17.47	16.57	**11.00**	15.59	12.47	**7.00**	30.55	14.92	**2.00**	7.00	4.55	**4.00**
MW (Loud)	21.02	14.59	**5.00**	12.61	11.68	**10.00**	20.04	15.70	**7.00**	5.50	4.95	**2.00**
HW (Comfortable)	14.03	14.93	**4.00**	12.71	9.82	**14.00**	14.60	10.53	**5.00**	36.10	-	**1.00**
HW (Soft)	9.15	7.27	**14.00**	15.40	6.77	**5.00**	13.03	12.22	**3.00**	25.50	2.12	**2.00**
HW (Loud)	8.67	2.31	**3.00**	17.39	13.33	**13.00**	14.26	13.15	**8.00**	-	-	**0.00**

Note. *M* refers to 'Mean'. *SD* refers to 'Standard Deviation'. N = 24. LW = Low Workload. MW = Moderate Workload. HW = High Workload.

Despite no statistically significant effects being found for the Auditory Management Strategy used, as can be seen in Figure 16, clear numerical trends are evident in the data. There are differences in the spatial orientation accuracy values between the Soft and Loud conditions. Moreover, a transition appears to occur from 'Passive' to 'Actively Attend/Ignore' responses as the associated level of workload and auditory feedback also changes. While these results will be elaborated on in the Discussion, it is of note that under the soft auditory conditions a more Passive response is elicited by participants, and with higher spatial orientation accuracy. Conversely, it is apparent the 'Loud auditory' conditions induce more Active responses and with lower accuracies, most notably at Low and High workloads.

Figure 16

Mean Spatial Orientation Accuracies (Degrees) and Frequency Counts for the Four Auditory Management Strategies Used



Note. See Appendix 6 for large-sized figures.

6.4 Discussion

The role of auditory feedback on remote pilot task performance was explored in Experiments 1 and 2. However in both experiments, auditory feedback levels were fixed (i.e., comfortable as set by the participant). In Experiment 3, auditory feedback was manipulated, providing two additional volume increments (+/-10 dBA). The subjective management strategies adopted by participants to deal with the incremental auditory feedback were also examined. Workload conditions remained consistent with Experiment 2 from Low to Moderate to High. However, combined with four auditory levels, a doubling of total conditions per participant occurred, from six to twelve.

6.4.1 Subjective Workload and Waypoint Spotting Accuracy Task Results

The results from the NASA-TLX assessment of subjective workload were consistent with the findings from Experiment 2. In terms of auditory feedback, workload was perceived to be at its lowest with no auditory feedback (i.e., Quiet/ambient), followed by the Comfortable volume level. Significant effects were found for all three auditory conditions when compared with the Quiet level, and only one significant effect was found between the Auditory-Comfortable and Auditory-High conditions.

In terms of waypoint spotting accuracy, the task was completed with a high level of accuracy, as determined by correct scores. This result is consistent with Experiment 2. However, and unlike Experiment 2, no analyses reached statistical significance. In terms of the first and second research questions, for waypoint spotting accuracy, the three workload and four incremental auditory feedback levels had no significant effect on task performance. Numerically, the error frequency counts reflected the probable values based upon the condition difficulty. That is, error occurrence increased in value from Low to High workload. One possible explanation for the non-significant result could be related to the experimental design. In comparison to the previous, this experiment doubled the total number of task conditions per participant from six to twelve. Furthermore, the total spotting errors committed by participants remained low, and when further divided into two categories (i.e., 'miss' or 'wrong identification'), the level of precision seen in these categories amplified.

6.4.2 Spatial Orientation Accuracy

This experiment introduced two additional real-time auditory feedback conditions: 10 dBA below the participant-defined comfortable level, and 10 dBA above the comfortable level. These additions were used instead of only comparing a single comfortable auditory feedback condition with a quiet condition under different workloads, as was the case for Experiment 1 and Experiment 2. In Experiment 2, Low and Moderate workloads, combined with comfortable auditory feedback produced the most accurate performance in terms of spatial orientation error. Hence, an average of these two auditory conditions was used for the Baseline level in Experiment 3.

In terms of the first two research questions, task performance differences were observed with incremental changes to auditory volume, when combined with the two greater workload levels. For task conditions with the single stressor of workload only (i.e., the Moderate and High workload with Quiet auditory conditions), performance was numerically unaffected, albeit with a lower accuracy compared to the Baseline group. However, when two stressors were present, and the stressors were not able to be controlled by the participant (i.e., at Soft and Loud auditory levels), performance deteriorated. With two stressors, and the ability to control one to reduce its effect (i.e., the participant-defined Comfortable level), task accuracy improved (see Table 18).

The findings revealed for High workload, additional auditory feedback perceived as noise influenced performance, but not in a uniform way. Although the hypothesis was validated numerically, a statistically significant difference was not observed between the Baseline and High Workload (Comfortable or Soft) conditions (refer to Figure 15). In comparison to Experiments 1 and 2, the number of task conditions doubled from six to twelve. However in terms of exposure to auditory feedback (i.e., Experiment 1 and 2 each had three quiet conditions and three auditory conditions), there was a tripling of task conditions that included this stressor, from three to nine. The observed differences between the Baseline and the High Workload (Comfortable or Soft) conditions, indicate spatial orientation accuracy declined as the level of workload increased to the highest level. These results need to be interpreted with caution, as too few errors were evident to make a meaningful conclusion. However it is likely several factors contributed to task performance observed below the threshold of statistical significance. Firstly, High workload combined with a lower intensity of auditory feedback was arousing, albeit transitioning to an overload, evidenced numerically by an accuracy decline. Secondly, the number of additional tasks involving auditory feedback compared with the two previous experiments, possibly gave participants the opportunity to develop an adequate response (i.e., adapt).

The results furthermore indicate with a Moderate level of workload, the addition of Loud or Soft auditory feedback is cognitively taxing, evidenced by the two lowest mean accuracies recorded. When Loud auditory feedback was combined with High workload, a similar effect was also noticeable.

6.4.3 Auditory Management Strategies

In terms of the third research question, it is possible the management strategy employed by participants mediated the effect of workload and auditory feedback on task performance. In dealing with auditory feedback as a stressor, individuals reported to employ different strategies. If the sound was at a Soft level, it was noticed in the background. While this perception of the Soft sound continued to consume cognitive resources (Edwards, 2016), it is unlikely that under Low workload, this will adversely affect task performance. Conversely, under Moderate or High workload, Soft auditory feedback, irrespective of it being consciously perceived as noise (or not), consumed cognitive resources to a threshold, deteriorating performance.

When the auditory feedback was Loud and perceived as noise, different strategies could be used to deal with it, including actively ignoring it, or actively attending to it. Each of these strategies consume cognitive resources to a level where performance deteriorates. In this experiment, when the auditory feedback was Soft, the addition of this sound consumed individuals limited cognitive resources (i.e., evidenced by lower accuracies), an effect that did not appear to be obvious to the participant. This was evident by their self-reported management strategy (i.e., noticed in background as opposed to making an active response). When the sound was perceived as subjectively Comfortable, its effect may have been arousing. While still consuming cognitive resources, it could have stimulated the individual to adapt, thereby performing the task more optimally.

6.4.4 Limitations and Future Research

In this experiment, several limitations could not be overcome. The auditory conditions were extended to four increments, and tested non-pilots only. While the specific volume level defined as 'Comfortable' varied in intensity between participants, the higher and lower relative volume increments of +/- 10 dBA remained equal for all. While valuable in an experimental setting, it may not be possible to adjust volume as precisely or consistently, in a real-world setting. Hence, the limited utility of these findings should be acknowledged. Furthermore, the Auditory Management Strategy questionnaire was introduced to better understand the noise effect on participants. While the questionnaire underwent a series of validity tests, as well as a pilot test with five participants, it remains a subjective measure, and as with all subjective measures relies on individuals' ability to accurately appraise the situation. This is no easy feat, and remains a limitation. Future research might investigate other objective methods/measures (e.g., physiological) to determine the noise effect on cognition within an RPAS setting.

As with Experiment 2, participants were also instructed to respond to a simulated failure, and therefore had prior knowledge that a failure was going to occur ahead of time. While the participants did not know when the failure would occur, there was again an anticipation that could not be removed due to the experimental design. This associated anticipation could, therefore, be considered an additional stressor that was not explicitly tested. Although workload was manipulated via secondary tasks in both Experiment 2 and 3, these data sets were not analysed. Future research could examine the accuracy of participants completing the secondary tasks as a measure of workload, and in response to the perceived intensities of concurrent auditory information within an RPAS setting. Furthermore, the secondary task involved mental arithmetic requiring some level of environmental rejection (Lacey, 1959; see Chapter 3 section 3.2). This was in opposition to the spotting task whereby environmental intake was necessary. In future research, the primary and secondary task directions could be reversed (i.e., environmental intake versus rejection) or paired, supporting unidirectional and multi-directional measurement.

Finally, the failure to reach statistical significance in the perception (i.e. spotting) tasks, and between the Baseline and High-Quiet, High-Soft and High-Comfortable pairs for decision-making (i.e., spatial orientation) accuracy, is perhaps a function of the experimental design. The spotting task was determined to be relatively simple in Experiment 2, compared to the spatial orientation task, however it was maintained in Experiment 3 for consistency. Furthermore, in contrast to the first two experiments, each with six task conditions, this experiment had twelve. While unavoidable, the number of tasks increased in order to accommodate the additional auditory volume increments. Although the task conditions were counterbalanced, the doubling of total tasks performed may have allowed the participants the opportunity to establish or adopt a coping response in order to deal with the stressors, and complete the tasks more effectively. Hence, future experiments might benefit from using more extreme task conditions, and/or with relatively higher contrasts in terms of task difficulty (i.e., via workload), volume level, and under different operating conditions (e.g., manual control instead of automated).

6.5 Summary

This chapter described the design and results of Experiment 3. In contrast to the previous two experiments, this experiment utilised non-pilot participants only. Expanding upon the previous findings, the same methodology used in Experiment 2 was also adopted, but with a notable difference. Specifically, the effect of incremental real-time auditory feedback, relative to a participant-defined comfortable level, was measured. Furthermore, the subjective auditory management strategies used by participants in dealing with the incremental changes to auditory feedback, paired with different workloads, was also compared with their decision-making performance. The subjective workload results were consistent with expected values.

For the spotting accuracy task, a similar numerical trend was in line with the findings from Experiment 2. While no results reached statistically significant values in this experiment, it was likely due to a combination of high participant accuracy, and a dilution of the possible task pairs through a doubling of total tasks completed per participant (i.e., six in Experiment 2 versus twelve in Experiment 3). The results for the decision-making task showed a decline in performance when auditory feedback was considered ‘loud’ (i.e., +10 dBA from the comfortable level) or ‘soft’ (i.e., -10 dBA from the comfortable level), and when there was a sufficient level of increased workload (i.e., Moderate or High workload conditions). In terms of dealing with the auditory feedback, task performance was mediated by auditory management strategies employed. Irrespective of how the feedback was consciously noticed when the volume level was soft (i.e., not noticed or passively attended to in the background), it seemingly remained cognitively taxing when combined with moderate or high workloads. Conversely at loud levels, an active strategy of attending to, or ignoring, was required.

For comfortable levels (i.e., the Baseline condition), whilst still consuming cognitive resources, many participants attended to the stressor either passively or actively. With a high level of spatial orientation accuracy under this condition, it is likely the auditory feedback had a stimulatory or arousing effect on participants. Hence, it provides further consensus with the expected results of the MAM relating to predicted performance, and exposure to stressors.

This chapter concludes the research component of this thesis. In the next and final chapter, a general discussion of the results from all three experiments will be described. In addition to the limitations of the three experiments collectively, implications and avenues for future research will be presented.

CHAPTER 7 – DISCUSSION AND CONCLUSION

7.1 General Discussion

Remote pilots and their conventional pilot counterparts are differentiated by one key distinction, physical separation from the aircraft they control. Sensory cues benefit conventional pilots located within a cockpit, having access to the full suite of stimuli for their perception. However, for remote pilots, the scope of sensory cues available are limited, and/or have reduced quality. As a result of this, their ability to perform a task sufficiently may become challenged. Primarily this study measured how the presence or absence of a sensory cue not normally available to remote pilots, namely real-time auditory feedback, influenced task performance. Furthermore, this was measured in combination with other external stressors (i.e., workload, control configuration, and visual display type). How behavioural models of adaptability and arousal could be used to predict task performance in applied RPAS settings were also considered. To facilitate this research, three experiments were conducted to measure performance under different intensities of real-time auditory feedback from the RPA propellers, and workloads.

Experiment 1 examined how auditory feedback and workload pairings influenced the task performance of conventional fixed-wing pilots. They manually operated a small multi-rotor RPA within an enclosed environment, under different visual conditions. The counterbalanced repeated measures design comprised two tasks (i.e., Task A and Task B; completed in consecutive order), performed under six experimental conditions. For Task A and B, two real-time auditory feedback conditions (with and without), and three visual display configurations were utilised: VLOS ((Control); i.e., VLOS + Monitor Display), BVLOS-Monitor (i.e., BVLOS with

Monitor Display only) and BVLOS-Goggles (i.e., BVLOS with Goggles display only). In Task B, workload was increased, via the inclusion of a wind shear component (i.e., wind versus no wind). Task performance was measured in terms of vertical and horizontal flying accuracy, and timeliness to complete each task. The results revealed several findings. Firstly, the BVLOS-Monitor display type enabled superior task performance in comparison to the Control and Goggles conditions. Secondly, auditory feedback influenced task performance but not in the hypothesised direction. It was thought that real-time auditory feedback would assist participant task performance because it is not a sensory cue that is normally available to remote pilots under BVLOS conditions. However, statistically significant results were found in Task B for the spotting task accuracy only.

Superior task performance was also only found in the presence of the additional wind shear component. When the workload level was low (i.e., no wind shear), task accuracy declined. Hence, auditory feedback was demonstrated to have both a stimulatory and distracting effect on participant performance. The third finding related to the effect of auditory feedback on timeliness for both Task A and B. Although not statistically significant, the results revealed a numerical trend consistent with the predicted values in reference to the MAM. In Task A, auditory feedback hindered performance by increasing timeliness by six seconds, and by two seconds in Task B when wind shear was present. However, when wind shear was removed in Task B, auditory feedback was found to improve task performance by almost two seconds. Hence, the results demonstrated the variable influence of auditory feedback on task performance. It was perceived as both useful information and noise (i.e., auditory information that is unwanted), evidenced by changes in performance results.

Experiment 2 sought to understand the differences between conventional fixed-wing pilots and non-pilot participant groups involving an automated RPA flight task. Under these settings, participants were not required to perform manual control inputs, and thus had limited interaction with the RPA. Performance was again measured under real-time auditory feedback directly from the RPA propellers, however workload was presented differently by utilising a secondary task, and was expanded to include three levels. In contrast to Experiment 1, all tasks were performed under automated BVLOS visual flying conditions only, utilising pre-recorded videos of actual RPAS flights. Hence, while a spotting task was maintained, a different procedure was implemented. In each task, participants were required to note and record via an iPad the waypoint numbered mat as it appeared on the monitor display (i.e., not presented in numerical order, and in different order for each flight). At a random mat, and unknown to the participant, a failure would occur, and the RPA would hover continuously over a single mat without proceeding onwards, thereby signalling the end of the flight. Following this failure, participants were instructed to perform a spatial orientation task by indicating the correct home point direction (i.e., horizontal, 360-degrees) relative to the final waypoint mat location. With the RPA onboard camera oriented towards the ground for the entirety of the flight, the waypoint numbers would sometimes appear correctly oriented, upside down or on their sides. For support, a reference map was provided to participants to assist them maintain orientation for the duration of the flight.

No statistically significant differences in spotting accuracy or spatial orientation performance between the pilot and non-pilot groups were found. This was an important finding. Given the task being measured (i.e., an automated BVLOS flight) is normally restricted to professional pilots, the non-pilots performed the task to an equivalent level of accuracy. In terms of the spotting task, it was performed with high accuracy. For the errors that did occur, ‘Misses’ occurred almost twice as frequently as ‘Wrong Identifications’. Workload was also found to moderate errors, a finding consistent with expectations (i.e., more errors occurred under high workload, and fewer under low workload). However, the presence or not of auditory feedback did not have the same linear effect. In terms of spatial orientation, superior differences in accuracy occurred under Low and Moderate workloads, when auditory feedback was included. When workload increased to a High level, and auditory feedback was also included, accuracy declined. Again in this experiment, both the positive and negative effects of auditory feedback on task performance were evident. Depending on the workload level, and like the findings from Experiment 1, auditory feedback could be perceived as both positive sound containing important information, and noise. The results from Experiment 2 also appeared to align with the predicted performance values more closely in terms of the MAM. While these findings remained broadly consistent with those obtained in Experiment 1, the explicit location of task performance relative to the MAM curve was unclear. Hence, further refinement of the independent variables was required.

In pursuit of this objective, and using the same methodology as Experiment 2, the third experiment introduced four auditory feedback levels (i.e., two additional levels of auditory feedback in comparison to Experiment 1 and 2). The three workload levels were also maintained. The aim of this modification was to further investigate the role of auditory feedback on task performance by using a finer grade of incremental conditions. Furthermore, because Experiment 2 established no difference in performance between the pilot and non-pilot groups, only participants without previous flying experience were recruited. The procedure was identical to the second experiment, except for an additional questionnaire included at the end of each task. This questionnaire recorded the auditory management strategy utilised by participants to deal with the auditory feedback. In terms of the data analysis, the third experiment differed from the second. The results from the second experiment were used to define the research question, hence a planned comparisons a priori analysis was utilised to measure task performance via spatial orientation accuracy. The analysis consisted of a Baseline group, made up of the two best performing auditory feedback groups in the second experiment (i.e., with Low and Moderate levels of workload), and the other task conditions involving three workload and four auditory levels.

In terms of the findings, firstly, the subjective workloads assessed via the NASA-TLX questionnaire were in line with expected values, and consistent with the results obtained in Experiment 2 (i.e., the mean scores increased as workload increased from Low to High, via the secondary task). Importantly, subjective workload increased consistently for the four auditory feedback levels in order: Quiet, Comfortable, Soft and Loud. Secondly, waypoint spotting accuracy was again completed successfully with high accuracy. No statistically significant findings were revealed, however observed differences for frequency counts (i.e., how many errors occurred in total) in each of the task conditions were in line with expectations, consistent with Experiment 2. Thirdly, in terms of spatial orientation accuracy, the planned comparisons analysis revealed statistically significant differences between the Baseline condition and other task conditions with two stressors (e.g., Moderate or High workload combined with Soft or Loud auditory feedback). Under Moderate workload, task accuracy declined when both Soft and Loud auditory conditions were included, suggesting the volume of auditory feedback was perceived as noise and/or cognitively taxing. Hence, while the two additional auditory levels were quite different in terms of intensity, their effect on performance was similarly negative.

Finally, the auditory management strategies used by participants in responding to the auditory feedback stressor, were also measured via the Auditory Management Questionnaire. When sound was perceived as Soft, although noticed in the background, it was found to induce poorer performance. When coupled with a Low level of workload, this volume may not be problematic in terms of task performance. However, when workload was High, the effect on task performance was found to be similarly detrimental. This was evident even if the strategies used to actively attend to, or ignore the auditory component, were different. Furthermore, when auditory feedback was at the Comfortable level, as defined individually by each participant, its effect under Low and Moderate workload conditions (i.e., the Baseline) improved task performance. It seems possible that while consuming cognitive resources, it also had an arousing or stimulatory effect, reflected by superior task accuracy. The results from Experiment 3 provide further evidence of how auditory feedback effects remote pilot task performance; dependent on the volume level of the sound, as well as the combined level of workload.

Consistently in all three experiments, different levels of real-time auditory feedback from the RPA propellers were maintained. Throughout the study, examples of the real-time auditory feedback being perceived as both sound and noise, via performance deterioration were demonstrated. Implications of the key findings, as well as limitations and opportunities for future research will now be discussed.

7.1.1 Key Findings and Implications

The first key finding from this research showed that in an environment with limited sensory cueing availability, the transitional nature of auditory information being perceived as both useful auditory feedback and noise, influenced participants' performance. For remote pilots who do not usually have access to a real-time auditory cue, it is reasonable to expect that any additional cue would be advantageous. However, this was shown not to be the case. Moreover, sound and workload pairings as two separate external stressors, interacted with each other and did not affect performance uniformly. The addition or removal of one, could tip performance from advantageous to disadvantageous, but was also related to the workload the participant was exposed to. Auditory information consumes cognitive resources for processing. The cocktail party effect whereby sound that is noticed and processed, even if subconsciously, still consumes cognitive resources to do so. The differences in performance could depend on individuals' preferences for sound, and particularly in the three experimental conditions, the type of sound (i.e., directly from the RPA propellers). Furthermore, filter theories including Broadbent's selective filter theory of attention, Treisman's attenuation model of selective attention, and late-selection models could also be used to contextualise and explain some of these research findings. However, regardless of the mechanisms involved, it is important to recognise the mere addition of a sensory cue to an environment where it is usually non-existent does not automatically result in increased levels of performance. Furthermore, there are implications for future human-centred RPAS design and operation. By finding examples of positive and negative effects of auditory feedback on task performance, the inclusion of an additional sensory cue type should consider the type of task being undertaken (e.g., long endurance, automated BVLOS versus short endurance VLOS), as well as the stage of flight involving different workloads. Hence, there is justification for the inclusion of a dynamic or adaptive auditory cueing (as seen in Experiment 3) component made available to remote pilots, where it can be included and removed as necessary.

The second key finding came from Experiment 2, where no statistically significant differences in task performance between the Pilot and Non-Pilot groups were evident. In the context of current civil and commercial RPAS operations this is significant, as technologies that were historically exclusive to military and government operators are now accessible to non-aviation specialists. In the present regulatory environment, restrictions are placed upon remote pilots conducting BVLOS operations. The experimental settings involved automated BVLOS operating conditions, which are currently inaccessible (at least legally) to remote pilots without the appropriate licencing and training. It is perhaps an assumption that individuals without an aviation background or experience in the operating environment pose a greater threat to the safety of other airspace users. However, for specific tasks such as automated BVLOS flight involving limited user input, it is apparent that previous piloting experience is not necessarily an advantage when the focus is on perception (i.e., spotting) or decision-making (i.e., spatial orientation) tasks. These findings are limited to specific operating conditions. However as future RPA users and uses change, the importance for remote pilots to have extensive qualifications, licensing and experience, could change. Regulatory amendments may also be warranted.

The third key finding from this research relates to the behavioural models of performance, involving adaptability and arousal, in response to external stressors. In Experiment 1, participant performance was found to broadly reflect the expected values in consideration of the MAM. This was further demonstrated in Experiment 2 and 3, whereby task performance, particularly in terms of spatial orientation accuracy, was benefited by a comfortable auditory component. It also improved from low workload to superior performance as workload increased to a moderate level. Task performance then declined as workload further increased, or when the sound was perceived as noise.

When the auditory component featured additional increments of volume, at a level below and above the comfortable, each had negative effects on task performance. Hence, to include auditory feedback in future RPAS should be considered with caution, and with an acknowledgement of the variable effect on performance that auditory information can have.

7.1.2 Limitations and Future Research

As acknowledged in the experiment chapters, unavoidable limitations exist with the research performed. Controlled environments are advantageous for maintaining homogeneous testing conditions, and to reduce the influence of external variables on participant performance. However, by their nature, controlled environments also impose limitations on the spectrum of data that can be collected and the generalisability of the results.

Experiment 1 examined manual flying RPA performance under BVLOS operating conditions, in a controlled indoor environment. The second experiment measured the performance of non-pilots against pilots, in an automated configuration. How non-pilots would have performed in Experiment 1, requiring manual control of the RPAS under BVLOS conditions is an unknown. At present, it is not possible to ascertain this information in many countries, including Australia, due to regulatory and institutional limits, and the necessity of remote pilots to hold appropriate licensing for BVLOS flights. While Experiment 2 and 3 each tested non-pilot participants under automated BVLOS flying conditions, other factors including real-world distractions and the associated pressures of performing a difficult task with monetary and/or safety consequences, could not be captured entirely. This remains a limitation of an emerging sector within the broader aviation industry.

In all three experiments, a multi-rotor RPAS was used. While the physical type and specifications of the RPAS remained consistent for all three experiments, it was not possible to use a fixed-wing RPA due to the confines of the research area, and experimental design. Like conventionally piloted aircraft and helicopters, the two different RPA types can be used for separate operating purposes, and are more suitable for certain environments. Hence, how the performance measurements of participants might have differed, if at all, by using a fixed-wing RPA remains unknown. Furthermore, how task performance could have changed under a differently sounding auditory cue (i.e., single engine/motor versus four motor propellers), when combined with the workload increments used in Experiment 2 and 3, also remains unknown, and hence is an area for future research.

By extension, all experiments used a sound type generated from the aircraft engines. In the real-world, it is feasible that additional sounds including warning alerts, radio communications, and personal chatter between other colleagues or clients, may exist and occur simultaneously. This would likely create a distraction. How these combined sounds might influence task performance, or 'tip' a comfortable sound into one that is perceived as noise, is unknown. Thus, future research may benefit in utilising a real-world setting that includes other environmental sounds and distractions occurring simultaneously as the task is being measured.

Workload was varied in all three experiments, however measured subjectively in two. While the results reflected the hypothesised predictions and subjective reports, future research should consider alternate measures of workload, with a focus on objectively quantifying the workload, thus permitting a direct comparison between different experimental conditions.

This study measured several participant groups with different piloting experience. However, there is some research, albeit limited that indicates personality can influence individual behavioural responses to noise (Belojević et al., 1997; Beheshti et al., 2019; Golmohammadi et al., 2021), which can ultimately change performance success. In the presence of loud noise, extroverts tend to perform better than introverts, while in quieter conditions the reverse effect is possible (Geen et al., 1985; Furnham & Strbac, 2002; Molesworth, 2016). Hence, future research should investigate the relationship between personality and auditory feedback/noise on RPAS performance.

Finally, future research should investigate the above-mentioned limitations in relation to a combination of automated and manual RPA flying tasks. Moreover, the point at which the flights ended in Experiment 2 and Experiment 3 could be extended in future research, where a takeover request requiring manual RPA control to a pre-defined location could be employed. As this could introduce new perceived differences in workload, secondary tasks coupled with a manual takeover may provide further insight into decision-making and overall task performance.

7.2 Conclusion

Remotely piloted aircraft systems offer new capabilities in a growing civilian field. Whilst the technologies continue to develop, unlike the established broader aviation industry which has adopted a human-centred system approach, the same is not universal for remotely piloted aircraft. This study presented three experiments testing the effects of real-time auditory feedback from the RPA, combined with incremental workload, on participant task performance. Nuanced effects of auditory feedback were found, where it was perceived as both useful information and helpful, or as noise, reflected by changes in task performance measurements. When exposed to the external stressors of auditory feedback and workload, the results broadly met the expectations of predictive models of adaptational behavioural performance, including the Maximal Adaptability Model.

The sensory deprived environment in which remote pilots interact with RPAS is unique, however this also provides a 'clean slate' opportunity for the design of future systems. The research findings provide evidence for the usefulness (or not) of different auditory feedback intensities when combined with varying workloads. Depending on the type of RPAS mission and stage of flight, it could be reasonable to include additional sensory cues like real-time auditory feedback from the RPA to enhance the remote pilot's adaptability and assist with optimal performance. Similarly, at other stages of the flight, the real-time auditory cue could be removed to reduce the cognitive load placed on the operator (i.e., utilise an adaptive system). How these cues are interchanged is dependent on the mission, the task itself, individual differences of the remote pilot, and the level of RPAS automation involved. Finally, no statistically significant differences were observed between the pilot and non-pilot participants, under automated conditions. Individuals without conventional flying experience might, therefore, be suitable for more complex RPAS operations, including highly automated BVLOS flights.

REFERENCES

Adamantidis, A., & De Lecea, L. (2008). Physiological arousal: a role for hypothalamic systems. *Cellular and Molecular Life Sciences*, *65*(10), 1475-1488. https://doi.org/10.1007/s00018-008-7521-8

Air Accidents Investigation Branch. (2015). *AAIB investigation to G-OPFA and S037996, AAIB Bulletin: 10/2015*. https://bit.ly/3Aq12xe

Air Accidents Investigation Branch. (2016a). *AAIB investigation to F-GSBM and SAS Wildthing, AAIB Bulletin: 1/2016*. https://bit.ly/3tGg7XK

Air Accidents Investigation Branch. (2016b). *AAIB investigation to HZ-IBN, AAIB Bulletin: 12/2016*. https://bit.ly/2FWatX1

Air Accidents Investigation Branch. (2017). *AAIB investigation to Quest Q-200 UAS, AAIB Bulletin: 12/2017*. https://bit.ly/3tFqSJT

Air Accidents Investigation Branch. (2020a). *AAIB investigation to DJI Phantom 4, (UAS, registration n/a), AAIB Bulletin 9/2020.* https://bit.ly/3Oh0RtS

Air Accidents Investigation Branch. (2020b). *AAIB investigation to DJI Mavic Pro 2, AAIB Bulletin 6/2021*. https://bit.ly/3UNm9BX

Akamatsu, M., MacKenzie, I. S., & Hasbroucq, T. (1995). A comparison of tactile, auditory, and visual feedback in a pointing task using a mouse-type device. *Ergonomics*, *38*(4), 816-827. https://doi.org/10.1080/00140139508925152

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257-262. https://doi.org/10.1016/j.cub.2004.01.029

Alais, D., Newell, F., & Mamassian, P. (2010). Multisensory processing in review: From physiology to behaviour. *Seeing and Perceiving*, *23*(1), 3-38. https://doi.org/10.1163/187847510X488603

Albeaino, G., Eiris, R., Gheisari, M., & Issa, R. R. (2021). DroneSim: A VR-based flight training simulator for drone-mediated building inspections. *Construction Innovation*. http://doi.org/10.1108/CI-03-2021-0049

Antuñano, M. J., & Spanyers, J. P. (2006). *Hearing and noise in aviation* (No. AM-400-98/3). United States. Federal Aviation Administration. https://rosap.ntl.bts.gov/view/dot/38652

Arent, S. M., & Landers, D. M. (2003). Arousal, anxiety, and performance: A reexamination of the inverted-U hypothesis. *Research Quarterly for Exercise and Sport*, *74*(4), 436-444. https://doi.org/10.1080/02701367.2003.10609113

Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Pozzi, S., & Babiloni, F. (2016). A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks. *Progress in Brain Research*, *228*, 295-328. https://doi.org/10.1016/bs.pbr.2016.04.021

Arlinger, S., Lunner, T., Lyxell, B. & Pichora-Fuller, M. K. (2009). The emergence of Cognitive Hearing Science. *Scandinavian Journal of Psychology*, *50*, 371–384. https://doi.org/10.1111/j.1467-9450.2009.00753.x

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34*(5), 502-508. https://doi.org/10.1111/opo.12131

Arrabito, G. R., Ho, G., Lambert, A., Rutley, M., Keillor, J., Chiu, A., Au, H., Hou, M. (2011). *Human factors issues for controlling uninhabited aerial vehicles: Preliminary findings in support of the Canadian Forces Joint Unmanned Aerial Vehicle Surveillance Target Acquisition System Project*. Ottawa, ON: Defence R&D Canada.

ATSB. (2018). *Aviation occurrence statistics 2007 to 2016* (AR-2017-2104). https://www.atsb.gov.au/media/5773880/ar-2017-104_final.pdf

ATSB. (2020). *Aviation occurrence statistics 2010 to 2019* (AR-2020-014). https://www.atsb.gov.au/media/5777724/ar-2020-014_final.pdf

Avila, C., Furnham, A., & McClelland, A. (2012). The influence of distracting familiar vocal music on cognitive performance of introverts and extraverts. *Psychology of Music*, *40*(1), 84-93. https://doi.org/10.1177/0305735611422672

Azman, A., Meng, Q., & Edirisinghe, E. (2010, August). Non intrusive physiological measurement for driver cognitive distraction detection: Eye and mouth movements. In *2010 3rd international conference on advanced computer theory and engineering (ICACTE)* (Vol. 3, pp. V3-595). IEEE. https://doi.org/10.1109/ICACTE.2010.5579547

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. (1986). *Working memory*. Oxford: Clarendon Press.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559. https://doi.org/10.1126/science.1736359

Bainbridge, L., Dorneich, M. (2009) Processing underlying human performance. In Wise, J.A., Wise, Hopkin, V. D., & Garland, D.J., (Eds.), *Handbook of aviation human factors* (pp. 7.1–7.68). CRC Press.

Baklanov, V. S. (2016). Role of structural noise in aircraft pressure cockpit from vibration action of new-generation engines. *Acoustical Physics, 62*(4), 456-461. https://doi.org/10.1134/S1063771016040047

Baldwin, C. L. (2012). *Auditory cognition and human performance: Research and applications*. Boca Raton, FL: Taylor & Francis Group.

Baldwin, C. L., Spence, C., Bliss, J. P., Brill, J. C., Wogalter, M. S., Mayhorn, C. B., & Ferris, T. K. (2012, September). Multimodal cueing: The relative benefits of the auditory, visual, and tactile channels in complex environments. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 56, No. 1, pp. 1431-1435). SAGE Publications. https://doi.org/10.1177%2F1071181312561404

Balog, C. R., Terwilliger, B. A., Vincenzi, D. A., & Ison, D. C. (2017). Examining human factors challenges of sustainable small unmanned aircraft system (sUAS) operations. In *Advances in human factors in robots and unmanned systems* (pp. 61-73). Springer, Cham. https://doi.org/10.1007/978-3-319-41959-6_6

Banbury, S., & Berry, D. C. (1997). Habituation and dishabituation to speech and office noise. *Journal of Experimental Psychology: Applied, 3*(3), 181–195. https://doi.org/10.1037/1076-898X.3.3.181

Banbury, S., & Berry, D. C. (1998). Disruption of office-related tasks by speech and office noise. *British Journal of Psychology, 89*(3), 499-517. https://doi.org/10.1111/j.2044-8295.1998.tb02699.x

Banbury, S. P., Macken, W. J., Tremblay, S., & Jones, D. M. (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors, 43*(1), 12-29. https://doi.org/10.1518/001872001775992462

Baum, A., Grunberg, N. E., & Singer, J. E. (1982). The use of psychological and neuroendocrinological measurements in the study of stress. *Health Psychology*, *1*(3), 217. https://doi.org/10.1037/0278-6133.1.3.217

BEA (2012). *Final report, F-GZCP, Air France 447*. https://bea.aero/docspa/2009/f-cp090601.en/pdf/f-cp090601.en.pdf

Becker, A. B., Warm, J. S., Dember, W. N., & Hancock, P. A. (1995). Effects of jet engine noise and performance feedback on perceived workload in a monitoring task. *The International Journal of Aviation Psychology*, *5*(1), 49-62. https://doi.org/10.1207/s15327108ijap0501_4

Beheshti, M. H., Taban, E., Samaei, S. E., Faridan, M., Khajehnasiri, F., Khaveh, L. T., Jebeli, M., Mehri, A., & Tajpoor, A. (2019). The influence of personality traits and gender on noise annoyance in laboratory studies. *Personality and Individual Differences*, *148*, 95-100. https://doi.org/10.1016/j.paid.2019.05.027

Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A. A., Masure, M. C., & Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience*, *10*(4), 536-540. https://doi.org/10.1162/089892998562834

Belojević, G., Jakovljević, B., & Aleksić, O. (1997). Subjective reactions to traffic noise with regard to some personality traits. *Environment International*, *23*(2), 221-226. https://doi.org/10.1016/S0160-4120(97)00008-1

Bergmann, K. (2017). The remarkable X-47b–The future of combat aviation?. *Asia-Pacific Defence Reporter, 43*(1), 14-16.

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience, 14*(10), 693. https://doi.org/10.1038/nrn3565

Blumer, T., Wargo, C., & Hunter, G. (2018, April). UAS situation awareness shortcomings, gaps, and future research needs. In *2018 integrated communications, navigation, surveillance conference (ICNS)* (pp. 2G1-1). IEEE. https://doi.org/10.1109/ICNSURV.2018.8384850

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, *44*, 58-75. https://doi.org/10.1016/j.neubiorev.2012.10.003

Borgman, C. L. (1986). The user's mental model of an information retrieval system: An experiment on a prototype online catalog. *International Journal of Man-Machine Studies*, *24*(1), 47-64. https://doi.org/10.1016/S0020-7373(86)80039-6

Boyd, S. P. (1983, October). Assessing the validity of SWAT as a workload measurement instrument. In *Proceedings of the human factors society annual meeting* (Vol. 27, No. 2, pp. 124-128). SAGE Publications. https://doi.org/10.1177/154193128302700202

Bravo, A., Vieira, D., & Ferrer, G. (2022). Emissions of future conventional aircrafts adopting evolutionary technologies. *Journal of Cleaner Production*, 347, 131246. https://doi.org/10.1016/j.jclepro.2022.131246

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT Press. https://doi.org/10.1121/1.408434

Brill, J. C., Lawson, B., Rupert, A. H., & Gagliano, I. (2015). Audiotactile displays for improving situation awareness and mitigating spatial disorientation. In *Proceedings of the American helicopter association* (pp. 1-4). Virginia, USA.

Britton, L. A., & Delay, E. R. (1989). Effects of noise on a simple visual attentional task. *Perceptual and Motor Skills*, 68(3), 875-878. https://doi.org/10.2466%2Fpms.1989.68.3.875

Broadbent, D. E. (1958). *Perception and communication*. Elsevier.

Broadbent, D. E. (1978). The current state of noise research: Reply to Poulton. *Psychological Bulletin*, *85*(5), 1052–1067. https://doi.org/10.1037/0033-2909.85.5.1052

Broadhurst, P. L. (1959). The interaction of task difficulty and motivation: The Yerkes Dodson law revived. *Acta Psychologica, Amsterdam*, *16*, 321-338. https://doi.org/10.1016/0001-6918(59)90105-2

Bronkhorst, A. W., Veltman, J. A., & Van Breda, L. (1996). Application of a three-dimensional auditory display in a flight task. *Human Factors, 38*(1), 23-33. https://doi.org/10.1518/001872096778940859

Brookhuis, K. A., & De Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention, 42*(3), 898-903. https://doi.org/10.1016/j.aap.2009.06.001

Bruggen, A. (2015). An empirical investigation of the relationship between workload and performance. *Management Decision, 53*(10), 2377-2389. https://doi.org/10.1108/MD-02-2015-0063

Burke, J. L., Prewett, M. S., Gray, A. A., Yang, L., Stilson, F. R., Coovert, M. D., & Redden, E. (2006). Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In *Proceedings of the 8th international conference on multimodal interfaces* (pp. 108-117). https://doi.org/10.1145/1180995.1181017

Burgess, M., & Molesworth, B. R. C. (2016). The noise reduction provided by aviation headsets. *Acoustics Australia*, *44*, 95-99. https://doi.org/10.1007/s40857-016-0050-y

Byers, J. C., Bittner Jr, A. C., Hill, S. G., Zaklad, A. L., & Christ, R. E. (1988, October). Workload assessment of a remotely piloted vehicle (RPV) system. In *Proceedings of the human factors society annual meeting* (pp. 1145-1149). SAGE Publications. https://doi.org/10.1177%2F154193128803201704

Cacioppo, J. T., Berntson, G. G., & Andersen, B. L. (1991). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome, 1991: Contributions from social psychophysiology. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3*(3), 321–336. https://doi.org/10.1037/1040-3590.3.3.321

Calvert, E. S. (1954). Visual judgments in motion. *The Journal of Navigation*, *7*(3), 233-251. https://doi.org/10.1017/S0373463300020907

CASA. (2017). *Guidance for conducting RPAS theory and operations (practical) training*. https://www.casa.gov.au/sites/default/files/guidance-conducting-rpas-theory-operation-training.pdf

Casner, S. M., Geven, R. W., & Williams, K. T. (2013). The effectiveness of airline pilot training for abnormal events. *Human Factors*, *55*(3), 477-485. https://doi.org/10.1177/0018720812466893

Cassidy, G., & MacDonald, R. A. (2007). The effect of background music and background noise on the task performance of introverts and extraverts. *Psychology of Music*, *35*(3), 517-537. https://doi.org/10.1177/0305735607076444

Castro, S. C., Strayer, D. L., Matzke, D., & Heathcote, A. (2019). Cognitive workload measurement and modeling under divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(6), 826–839. https://doi.org/10.1037/xhp0000638

Charles, R. L., & Nixon, J. (2019). Measuring mental workload using physiological measures: A systematic review. *Applied Ergonomics, 74*, 221-232. https://doi.org/10.1016/j.apergo.2018.08.028

Chen, J. Y., Haas, E. C., & Barnes, M. J. (2007). Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *37*(6), 1231-1245. https://doi.org/10.1109/TSMCC.2007.905819

Chepesiuk, R. (2005). Decibel hell: The effects of living in a noisy world. *Environmental Health Perspectives*, *113*(1), A34. https://doi.org/10.1289/ehp.113-a34

Choe, C. S., Welch, R. B., Gilford, R. M., & Juola, J. F. (1975). The “ventriloquist effect”: Visual dominance or response bias?. *Perception & Psychophysics*, *18*(1), 55-60. https://doi.org/10.3758/BF03199367

Clark, J. J., & Yuille, A. L. (1990). *Data fusion for sensory information processing systems* (Vol. 105). Springer Science & Business Media. https://doi.org/10.1007/978-1-4757-2076-1

Cockerton, T., Moore, S., & Norman, D. (1997). Cognitive test performance and background music. *Perceptual and Motor Skills*, *85*(3 Pt 2), 1435-1438. https://doi.org/10.2466/pms.1997.85.3f.1435

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409-412. https://doi.org/10.3758/BF03203962

Colle, H. A., & Reid, G. B. (1998). Context effects in subjective mental workload ratings. *Human Factors*, *40*(4), 591-600. https://doi.org/10.1518/001872098779649283

Cooper, G. E., & Harper Jr, R. P. (1969). *The use of pilot rating in the evaluation of aircraft handling qualities* (No. NASA-TN-D-5153). https://ntrs.nasa.gov/api/citations/19690013177/downloads/19690013177.pdf

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: The psychology of visual illusion*s. Lawrence Erlbaum. https://doi.org/10.4324/9781003050681

Cowan, B., Rojas, D., Kapralos, B., Moussa, F., & Dubrowski, A. (2015). Effects of sound on visual realism perception and task performance. *The Visual Computer, 31*(9), 1207-1216. https://doi.org/10.1007/s00371-014-1006-6

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

Crognale, M. A., & Krebs, W. K. (2011). Performance of helicopter pilots during inadvertent flight into instrument meteorological conditions. *The International Journal of Aviation Psychology*, 21(3), 235-253. https://doi.org/10.1080/10508414.2011.582443

Cummings, M. L., Myers, K., & Scott, S. D. (2006, November). Modified Cooper Harper evaluation tool for unmanned vehicle displays. In *Proceedings of UVS Canada: Conference on unmanned vehicle systems Canada*.

Cummings, M. L., Mastracchio, C., Thornburg, K. M., & Mkrtchyan, A. (2013). Boredom and distraction in multiple unmanned vehicle supervisory control. *Interacting with Computers, 25*(1), 34–47. https://doi.org/10.1093/iwc/iws011

Cummings, M. L., Bertucelli, L. F., Macbeth, J., & Surana, A. (2014). Task versus vehicle-based control paradigms in multiple unmanned vehicle supervision by a single operator. *IEEE Transactions on Human-Machine Systems, 44*(3), 353-361. https://doi.org/10.1109/THMS.2014.2304962

Cunio, R. J., Dommett, D., & Houpt, J. (2019, November). Spatial auditory cueing for a dynamic three-dimensional virtual reality visual search task. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 63, No. 1, pp. 1766-1770). SAGE Publications. https://doi.org/10.1177/1071181319631045

Dang, T., & Tapus, A. (2013, May). User modelling in adjustable control system. In *2013 International conference on collaboration technologies and systems (CTS)* (pp. 236-240). IEEE. https://doi.org/10.1109/CTS.2013.6567235

De la Torre, G., Ramallo, M. A., & Cervantes, E. (2016). Workload perception in drone flight training simulators. *Computers in Human Behavior*, *64*, 449-454. https://doi.org/10.1016/j.chb.2016.07.040

Deatherage, B. H. (1972). Auditory and other sensory forms of information presentation. In Van Cott, H.P., Kinkade, R.G., (Eds.), *Human engineering guide to equipment design*, (pp 123-166). US Government Printing Office.

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. *Frontiers in Neuroscience*, *14*, 268. https://doi.org/10.3389/fnins.2020.00268

Dell'Agnola, F., Momeni, N., Arza, A., & Atienza, D. (2020, July). Cognitive workload monitoring in virtual reality based rescue missions with drones. In *International conference on human-computer interaction* (pp. 397-409). Springer, Cham. https://doi.org/10.1007/978-3-030-49695-1_26

Denjean, S., Roussarie, V., Kronland-Martinet, R., & Velay, J. L. (2012, April). How does interior car noise alter driver's perception of motion? Multisensory integration in speed perception. In *Proceedings of the acoustics 2012 Nantes conference* (pp.1063-1068).

Dennett, D. C. (1995). Cognitive science as reverse engineering several meanings of “Top-down” and “Bottom-up”. In *Studies in logic and the foundations of mathematics* (Vol. 134, pp. 679-689). Elsevier. https://doi.org/10.1016/S0049-237X(06)80069-8

Deroy, O., Spence, C., & Noppeney, U. (2016). Metacognition in multisensory perception. *Trends in Cognitive Sciences*, *20*(10), 736-747. https://doi.org/10.1016/j.tics.2016.08.006

Dey, A., & Mann, D. D. (2010). Sensitivity and diagnosticity of NASA-TLX and simplified SWAT to assess the mental workload associated with operating an agricultural sprayer. *Ergonomics*, *53*(7), 848-857. https://doi.org/10.1080/00140139.2010.489960

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*(3), 271-285. https://doi.org/10.1518/155534307X255627

Diamond, D. M., Campbell, A. M., Park, C. R., Halonen, J., & Zoladz, P. R. (2007). The temporal dynamics model of emotional memory processing: a synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson law. *Neural Plasticity*, *2007*, 60803. https://doi.org/10.1155/2007/60803

Dicke, C., Jakus, G., & Sodnik, J. (2013, July). Auditory and head-up displays in vehicles. In *International conference on human-computer interaction* (pp. 551-560). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39262-7_62

Dixon, S. R., & Wickens, C. D. (2006). Automation reliability in unmanned aerial vehicle control: A reliance-compliance model of automation dependence in high workload. *Human Factors*, *48*(3), 474-486. https://doi.org/10.1518%2F001872006778606822

Dixon, S. R., Wickens, C. D., & Chang, D. (2005). Mission control of multiple unmanned aerial vehicles: A workload analysis. *Human Factors*, *47*(3), 479-487. https://doi.org/10.1518%2F001872005774860005

Doohan, M. A., Stewart, I. B., Coulter, T. J., Drovandi, C., Weier, S., & Borg, D. N. (2022). Modified Stroop task performance when wearing protective clothing in the heat: An evaluation of the maximum adaptability model. *Physiology & Behavior*, *246*, 113690. https://doi.org/10.1016/j.physbeh.2021.113690

Dorneich, M. C., Ververs, P. M., Whitlow, S. D., Mathan, S., Carciofini, J., & Reusser, T. (2006, October). Neuro-physiologically-driven adaptive automation to improve decision making under stress. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 50, No. 3, pp. 410-414). SAGE Publications. https://doi.org/10.1177%2F154193120605000342

Dosseville, F., Laborde, S., & Scelles, N. (2012). Music during lectures: Will students learn better?. *Learning and Individual Differences, 22*(2), 258-262. https://doi.org/10.1016/j.lindif.2011.10.004

Drury, J. L., & Scott, S. D. (2008). Awareness in unmanned aerial vehicle operations. *The International C2 Journal, 2*(1), 1-16.

Duffy, E. (1957). The psychological significance of the concept of" arousal" or" activation.". *Psychological Review*, *64*(5), 265. https://doi.org/10.1037/h0048837

Dunn, M. J., Molesworth, B. R., Koo, T., & Lodewijks, G. (2020). Effects of auditory and visual feedback on remote pilot manual flying performance. *Ergonomics*, *63*(11), 1380-1393. https://doi.org/10.1080/00140139.2020.1792561

Dunn, M. J., Molesworth, B. R., Koo, T., & Lodewijks, G. (2022). Measured effects of workload and auditory feedback on remote pilot task performance. *Ergonomics*, *65*(6), 886-898. https://doi.org/10.1080/00140139.2021.2003870

Dye, E., & Wells, N. (2017). Subjective and objective measurement of neonatal nurse practitioner workload. *Advances in Neonatal Care*, *17*(4), E3-E12. https://doi.org/10.1097/ANC.0000000000000353

Edwards, B. (2007). The future of hearing aid technology. *Trends in Amplification*, *11*(1), 31-45. https://doi.org/10.1177/1084713806298004

Edwards, B. (2016). A model of auditory-cognitive processing and relevance to clinical applicability. *Ear and Hearing*, *37*, 85S-91S. https://doi.org/10.1097/AUD.0000000000000308

Edwards, E. (1972). Man and machine: Systems for safety. In *Proceedings of the British airline pilots associations technical symposium.* British Airline Pilots Associations, London.

Edworthy, J., & Hellier, E. (2000). Auditory warnings in noisy environments. *Noise and Health*, *2*(6), 27.

Ekman, G. (1959). Weber's law and related functions. *The Journal of Psychology*, *47*(2), 343-352. https://doi.org/10.1080/00223980.1959.9916336

Ekman, G., Goude, G., & Waern, Y. (1961). Subjective similarity in two perceptual continua. *Journal of Experimental Psychology*, *61*(3), 222. https://doi.org/10.1037/h0041065

Endsley, M. R. (1988, October). Design and evaluation for situation awareness enhancement. In *Proceedings of the human factors society annual meeting* (pp. 97-101). SAGE Publications. https://doi.org/10.1177/154193128803200221

Engineer, N. D., Engineer, C. T., Reed, A. C., Pandya, P. K., Jakkamsetti, V., Moucha, R., & Kilgard, M. P. (2012). Inverted-U function relating cortical plasticity and task difficulty. *Neuroscience*, *205*, 81-90. https://doi.org/10.1016/j.neuroscience.2011.12.056

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162-169. https://doi.org/10.1016/j.tics.2004.02.002

Estes, S. (2015). The workload curve: Subjective mental workload. *Human Factors*, *57*(7), 1174-1187. https://doi.org/10.1177/0018720815592752

Eysenck, M. W. (1994). Perception and Communication: DE Broadbent. *The British Journal of Psychiatry, 164*(5), 711-712. https://doi.org/10.1192/S0007125000034759

Fan, J., & Smith, A. P. (2017, June). The impact of workload and fatigue on performance. In *International symposium on human mental workload: Models and applications* (pp. 90-105). Springer, Cham. https://doi.org/10.1007/978-3-319-61061-0_6

Fang, S.X., O'Young, S., & Rolland, L. (2018). Development of small UAS Beyond-Visual-Line-of-Sight (BVLOS) flight operations: System requirements and procedures. *Drones, 2*(2). https://doi.org/10.3390/drones2020013

Fastl, H., & Zwicker, E. (2007). Just-noticeable sound changes. In *Psychoacoustics* (pp. 175-202). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-68888-4_7

Ferris, T., Penfold, R., Hameed, S., & Sarter, N. (2006, October). The implications of crossmodal links in attention for the design of multimodal interfaces: A driving simulation study. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 50, No. 3, pp. 406-409). SAGE Publications. https://doi.org/10.1177/154193120605000341

Fox, J. G., & Embrey, E. D. (1972). Music—an aid to productivity. *Applied Ergonomics*, *3*(4), 202-205. https://doi.org/10.1016/0003-6870(72)90101-9

Frassinetti, F., Bolognini, N., & Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147*(3), 332-343. https://doi.org/10.1007/s00221-002-1262-y

Freeman, F. G., Mikulka, P. J., Scerbo, M. W., & Scott, L. (2004). An evaluation of an adaptive automation system using a cognitive vigilance task. *Biological Psychology*, *67*(3), 283-297. https://doi.org/10.1016/j.biopsycho.2004.01.002

Frid, E., Moll, J., Bresin, R., & Sallnäs Pysander, E. L. (2019). Haptic feedback combined with movement sonification using a friction sound improves task performance in a virtual throwing task. *Journal on Multimodal User Interfaces*, 13(4), 279-290. https://doi.org/10.1007/s12193-018-0264-4

Fu, M., Rothfeld, R., & Antoniou, C. (2019). Exploring preferences for transportation modes in an urban air mobility environment: Munich case study. *Transportation Research Record*, 2673(10), 427-442. https://doi.org/10.1177/0361198119843858

Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteguh, C., & Owen, G. (1999). Flight deck automation issues. *The International Journal of Aviation Psychology*, *9*(2), 109-123. https://doi.org/10.1207/s15327108ijap0902_2

Furnham, A., & Bradley, A. (1999). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology*, *11*(5), 445-455. https://doi.org/10.1002/(SICI)1099-0720(199710)11:5%3C445::AID-ACP472%3E3.0.CO;2-R

Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. *Ergonomics*, *45*(3), 203-217. https://doi.org/10.1080/00140130210121932

Gaillard, A. W. K. (1993). Comparing the concepts of mental load and stress. *Ergonomics, 36*(9), 991-1005. https://doi.org/10.1080/00140139308967972

Gao, S., & Wang, L. (2020, July). Effects of mental workload and risk perception on pilots' safety performance in adverse weather contexts. In *International conference on human-computer interaction* (pp. 278-291). Springer, Cham. https://doi.org/10.1007/978-3-030-49183-3_22

Gasaway, D. C. (1986). Noise levels in cockpits of aircraft during normal cruise and considerations of auditory risk. *Aviation, Space, and Environmental Medicine*, *57*(2), 103-112.

Geen, R. G., McCown, E. J., & Broyles, J. W. (1985). Effects of noise on sensitivity of introverts and extraverts to signals in a vigilance task. *Personality and Individual Differences*, *6*(2), 237-241. https://doi.org/10.1016/0191-8869(85)90114-X

Geiselman, E. E., Johnson, C. M., & Buck, D. R. (2013). Flight deck automation: Invaluable collaborator or insidious enabler?. *Ergonomics in Design*, *21*(3), 22-26. https://doi.org/10.1177/1064804613491268

Ghanbary Sartang, A., Ashnagar, M., Habibi, E., & Sadeghi, S. (2016). Evaluation of Rating Scale Mental Effort (RSME) effectiveness for mental workload assessment in nurses. *Journal of Occupational Health and Epidemiology*, *5*(4), 211-217. https://doi.org/10.18869/acadpub.johe.5.4.211

Gibb, R. W. (2007). Visual spatial disorientation: Revisiting the black hole illusion. *Aviation, Space, and Environmental Medicine*, *78*(8), 801-808.

Gibb, R., Schvaneveldt, R., & Gray, R. (2008). Visual misperception in aviation: Glide path performance in a black hole environment. *Human Factors*, *50*(4), 699–711. https://doi.org/10.1518%2F001872008X288619

Golmohammadi, R., Darvishi, E., Motlagh, M. S., & Faradmal, J. (2021). Role of individual and personality traits in occupational noise-induced psychological effects. *Applied Acoustics*, *173*, 107699. https://doi.org/10.1016/j.apacoust.2020.107699

Greenlee, E. T., DeLucia, P. R., & Newton, D. C. (2018). Driver vigilance in automated vehicles: Hazard detection failures are a matter of time. *Human Factors*, *60*(4), 465-476. https://doi.org/10.1177/0018720818761711

Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object?. *Nature Reviews Neuroscience*, *5*(11), 887. https://doi.org/10.1038/nrn1538

Groemer, G., Gruber, V., Bishop, S., Peham, D., Wolf, L., & Högl, B. (2010). Human performance data in a high workload environment during the simulated Mars expedition "AustroMars". *Acta Astronautica*, *66*(5-6), 780-787. https://doi.org/10.1016/j.actaastro.2009.08.017

Grosz, J., Rysdyk, R. T., Bootsma, R. J., Mulder, J. B. A., van der Vaart, J. H. C., & van Wieringen, P. C. (1995). Perceptual support for timing of the flare in the landing of an aircraft. In Hancock, P., Flach, J., Caird, J., & Vicente, K. (Eds.), *Local applications of the ecological approach to human-machine systems* (pp. 104-121). CRC Press. https://doi.org/10.1201/9780203748749

Guastavino, C. (2018). Everyday sound categorization. In Virtanen, T., Plumbley, M., Ellis, D., (Eds.), *Computational analysis of sound scenes and events* (pp.183-213). Springer. Cham. https://doi.org/10.1007/978-3-319-63450-0_7

Günther, S., Kratz, S., Avrahami, D., & Mühlhäuser, M. (2018, June). Exploring audio, visual, and tactile cues for synchronous remote assistance. In *Proceedings of the 11th pervasive technologies related to assistive environments conference* (pp. 339-344*)*. Corfu, Greece. https://doi.org/10.1145/3197768.3201568

Haas, E. C., & van Erp, J. B. (2014). Multimodal warnings to enhance risk communication and safety. *Safety Science*, *61*, 29-35. https://doi.org/10.1016/j.ssci.2013.07.011

Haines, M. M., Stansfeld, S. A., Job, R. S., Berglund, B., & Head, J. (2001). Chronic aircraft noise exposure, stress responses, mental health and cognitive performance in school children. *Psychological Medicine*, *31*(2), 265-277. https://doi.org/10.1017/S0033291701003282

Hancock, P.A. (1987). Arousal Theory, Stress And Performance: Problems Of Incorporating Energetic Aspects Of Behavior Into Human-Machine Systems Function. In: Mark, L.S., Warm, J.S., Huston, R.L. (Eds.), *Ergonomics and human factors. Recent research in psychology*. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-4756-2_17

Hancock, P. A. (2022). In defense of the maximal adaptability model. *Physiology & Behavior*, 113844. https://doi.org/10.1016/j.physbeh.2022.113844

Hancock, P. A., & Ganey, H. C. (2003). From the Inverted-U to the Extended-U: The Evolution of a Law of Psychology. *Journal of Human Performance in Extreme Environments*. https://doi.org/10.7771/2327-2937.1023

Hancock, P. A., Mercado, J. E., Merlo, J., & Van Erp, J. B. (2013). Improving target detection in visual search through the augmenting multi-sensory cues. *Ergonomics*, *56*(5), 729-738. https://doi.org/10.1080/00140139.2013.771219

Hancock, P. A., & Szalma, J. L. (Eds.). (2008). *Performance under stress*. Ashgate Publishing, Ltd.

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

Hancock, P. A., & Warm, J. S. (2003). A dynamic model of stress and sustained attention. *Journal of Human Performance in Extreme Environments*, *7*(1), 4. https://doi.org/10.7771/2327-2937.1024

Harris, W. C., Hancock, P. A., Arthur, E. J., & Caird, J. K. (1995). Performance, workload, and fatigue changes associated with automation. *The International Journal of Aviation Psychology*, *5*(2), 169-185. https://doi.org/10.1207/s15327108ijap0502_3

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

Haslbeck, A., & Hoermann, H. J. (2016). Flying the needles: flight deck automation erodes fine-motor flying skills among airline pilots. *Human Factors*, *58*(4), 533-545. https://doi.org/10.1177/0018720816640394

Hawkins, F. H., & Orlady, H. W. (2017). *Human factors in flight*. Routledge. https://doi.org/10.4324/9781351218580

Hayman, M. A. X. (1942). Two minute clinical test for measurement of intellectual impairment in psychiatric disorders. *Archives of Neurology & Psychiatry*, *47*(3), 454-464. https://doi.org/10.1001/archneurpsyc.1942.02290030112010

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychological Review*, *62*(4), 243. https://doi.org/10.1037/h0041823

Hecht, D., & Reiner, M. (2009). Sensory dominance in combinations of audio, visual and haptic stimuli. *Experimental Brain Research*, *193*(2), 307-314. https://doi.org/10.1007/s00221-008-1626-z

Hellier, E., Naweed, A., Walker, G., Husband, P., & Edworthy, J. (2011). The influence of auditory feedback on speed choice, violations and comfort in a driving simulation game. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(6), 591-599. https://doi.org/10.1016/j.trf.2011.07.004

Hendrix, C., & Barfield, W. (1996). The sense of presence within auditory virtual environments. *Presence: Teleoperators and Virtual Environments*, *5*(3), 290-301. https://doi.org/10.1162/pres.1996.5.3.290

Henelius, A., Hirvonen, K., Holm, A., Korpela, J., & Muller, K. (2009, September). Mental workload classification using heart rate metrics. In *2009 annual international conference of the IEEE engineering in medicine and biology society* (pp. 1836-1839). IEEE. https://doi.org/10.1109/IEMBS.2009.5332602

Hidalgo-Muñoz, A. R., Béquet, A. J., Astier-Juvenon, M., Pépin, G., Fort, A., Jallais, C., Tattegrain, H., & Gabaude, C. (2019). Respiration and heart rate modulation due to competing cognitive tasks while driving. *Frontiers in Human Neuroscience*, *12*, 525. https://doi.org/10.3389/fnhum.2018.00525

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: mandatory fusion within, but not between, senses. *Science*, *298*(5598), 1627-1630. https://doi.org/10.1126/science.1075396

Hulusic, V., Aranha, M., & Chalmers, A. (2008). The influence of cross-modal interaction on perceived rendering quality thresholds. In: *16th international conference in central Europe on computer graphics, visualization and computer vision*. University of W Bohemia, Plzen, Czech Republic.

Ho, C., & Spence, C. (2008). *The multisensory driver: Implications for ergonomic car interface design*. CRC Press. https://doi.org/10.1201/9781315555423

Hobbs, A., & Lyall, B. (2016). Human factors guidelines for unmanned aircraft systems. *Ergonomics in Design, 24*(3), 23-28. https://doi.org/10.1177%2F1064804616640632

Hobbs, A. (2017). Remotely Piloted Aircraft. In Landry, S. J. (Ed.). *Handbook of human factors in air transportation systems*. (pp. 379-395). CRC Press. https://doi.org/10.1201/9781315116549

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*(1-3), 73-93. https://doi.org/10.1016/S0301-0511(96)05223-4

Hooey, B. L., Kaber, D. B., Adams, J. A., Fong, T. W., & Gore, B. F. (2017). The underpinnings of workload in unmanned vehicle systems. *IEEE Transactions on Human-Machine Systems*, *48*(5), 452-467. https://doi.org/10.1109/THMS.2017.2759758

Howe, S. (2017). *The leading human factors deficiencies in unmanned aircraft systems.* https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20170005590.pdf

Howell, D. C. (2009). *Statistical methods for psychology*. Belmont, CA: Thomson.

Howse, W. R. (2011). Knowledge, skills, abilities, and other characteristics for remotely piloted aircraft pilots and operators. *Air force personnel center strategic research and assessment., No. DAS-2011-04.* Gurnee, IL: Damos Aviation Services Inc.

Huang, G., Steele, C., Zhang, X., & Pitts, B. J. (2019, November). Multimodal cue combinations: a possible approach to designing in-vehicle takeover requests for semi-autonomous driving. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 63, No. 1, pp. 1739-1743). SAGE Publications. https://doi.org/10.1177/1071181319631053

Huddleston, H. F., & Wilson, R. V. (1971). An evaluation of the usefulness of four secondary tasks in assessing the effect of a lag in simulated aircraft dynamics. *Ergonomics, 14*(3), 371-380. https://doi.org/10.1080/00140137108931257

Hughes, M. O. (2004). Anatomy of the anterior eye for ocularists. *Journal of Ophthalmic Prosthetics*, *8*, 25-35.

Hulusic, V., Aranha, M., & Chalmers, A. (2008). The influence of cross-modal interaction on perceived rendering quality thresholds. In *16th international conference in central Europe on computer graphics, visualization and computer vision* (pp. 41-48).

Hygge, S., Evans, G. W., & Bullinger, M. (2002). A prospective study of some effects of aircraft noise on cognitive performance in schoolchildren. *Psychological Science*, *13*(5), 469-474. https://doi.org/10.1111%2F1467-9280.00483

International Civil Aviation Organization. (2018). *The ICAO UAS toolkit frequently used terms.* https://www.icao.int/safety/UA/UASToolkit/Pages/Frequently-Used-Terms.aspx

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual and Motor Skills, 37*(3), 967-979. https://doi.org/10.1177/003151257303700360

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory, 18*(4), 394-412. https://doi.org/10.1080/09658211003702171

Jahncke, H. (2012). Open-plan office noise: The susceptibility and suitability of different cognitive tasks for work in the presence of irrelevant speech. *Noise and Health*, *14*(61), 315. https://doi.org/10.4103/1463-1741.104901

Jaussi, J. A., & Hoffmann, H. O. (2018). Manned versus unmanned aircraft accidents, including causation and rates. *International Journal of Aviation, Aeronautics, and Aerospace*, *5*(4), 3. https://doi.org/10.15394/ijaaa.2018.1262

Jerison, H. J. (1959). Effects of noise on human performance. *Journal of Applied Psychology*, *43*(2), 96–101. https://doi.org/10.1037/h0042914

Johansson, R., Holmqvist, K., Mossberg, F., & Lindgren, M. (2012). Eye movements and reading comprehension while listening to preferred and non-preferred study music. *Psychology of Music*, *40*(3), 339-356. https://doi.org/10.1177/0305735610387777

Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews*, *84*(2), 541-577. https://doi.org/10.1152/physrev.00029.2003

Jorna, P. G. A. M. (1993). Heart rate and workload variations in actual and simulated flight. *Ergonomics*, *36*(9), 1043-1054. https://doi.org/10.1080/00140139308967976

Kaber, D. B. (2018). Issues in human–automation interaction modeling: Presumptive aspects of frameworks of types and levels of automation. *Journal of Cognitive Engineering and Decision Making*, *12*(1), 7-24. https://doi.org/10.1177/1555343417737203

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. *Theoretical Issues in Ergonomics Science*, *5*(2), 113-153. https://doi.org/10.1080/1463922021000054335

Kämpfe, J., Sedlmeier, P., & Renkewitz, F. (2011). The impact of background music on adult listeners: A meta-analysis. *Psychology of Music, 39*(4), 424-448. https://doi.org/10.1177/0305735610376261

Kantowitz, B. H., & Campbell, J. L. (1996). Pilot workload and flightdeck automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 117–136). Lawrence Erlbaum Associates, Inc.

Kawabata, N. (1976). Mathematical analysis of the visual illusion. *IEEE Transactions on Systems, Man, and Cybernetics*, (12), 818-824. https://doi.org/10.1109/TSMC.1976.4309478

Ke, J., Du, J., & Luo, X. (2021). The effect of noise content and level on cognitive performance measured by electroencephalography (EEG). *Automation in Construction*, 130, 103836. https://doi.org/10.1016/j.autcon.2021.103836

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

Keuss, P. J. G., Van der Zee, F., & Van den Bree, M. B. M. (1990). Auditory accessory effects on visual processing. *Acta Psychologica*, *75*(1), 41-54. https://doi.org/10.1016/0001-6918(90)90065-N

Kiger, D. M. (1989). Effects of music information load on a reading comprehension task. *Perceptual and Motor Skills*, *69*, 531-534. https://doi.org/10.2466/pms.1989.69.2.531

Kilmer, K. J., Bateman, R., & Malzahn, D. (1988, October). Techniques of subjective assessment: A comparison of the SWAT and modified Cooper-Harper scales. In *Proceedings of the human factors society annual meeting 32*(2), 155-159. SAGE Publications. https://doi.org/10.1177/154193128803200233

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*(4), 352–358. https://doi.org/10.1037/h0043688

Kjellberg, A. (1990). Subjective, behavioral and psychophysiological effects of noise. Scandinavian Journal of Work, *Environment & Health*, 16, 29-38. https://doi.org/10.5271/sjweh.1825

Kjellberg, A., Muhr, P., & Skoldstrom, B. (1998). Fatigue after work in noise-an epidemiological survey study and three quasi-experimental field studies. *Noise and Health*, *1*(1), 47.

Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology*, 578. https://doi.org/10.3389/fpsyg.2013.00578

Knoll, A., Wang, Y., Chen, F., Xu, J., Ruiz, N., Epps, J., & Zarjam, P. (2011, September). Measuring cognitive workload with low-cost electroencephalograph. In *13th international conference on human-computer interaction (INTERACT)* (No. Part IV, pp. 568-571). Springer. https://doi.org/10.1007/978-3-642-23768-3_84

Koppen, C., & Spence, C. (2007a). Assessing the role of stimulus probability on the Colavita visual dominance effect. *Neuroscience Letters, 418*(3), 266-271. https://doi.org/10.1016/j.neulet.2007.03.032

Koppen, C., & Spence, C. (2007b). Audiovisual asynchrony modulates the Colavita visual dominance effect. *Brain Research, 1186*, 224-232. https://doi.org/10.1016/j.brainres.2007.09.076

Koppen, C., & Spence, C. (2007c). Spatial coincidence modulates the Colavita visual dominance effect. *Neuroscience Letters, 417*(2), 107-111. https://doi.org/10.1016/j.neulet.2006.10.069

Koo, T. T., Molesworth, B. R., Dunn, M. J., Lodewijks, G., & Liao, S. (2022). Trust and user acceptance of pilotless passenger aircraft. *Research in Transportation Business & Management,* 100876. https://doi.org/10.1016/j.rtbm.2022.100876

Kraft, C. L. (1978). A psychophysical contribution to air safety: Simulator studies of visual illusions in night visual approaches. In *Psychology: From research to practice* (pp. 363-385). Springer, Boston, MA. https://doi.org/10.1007/978-1-4684-2487-4_21

Krigolson, O. E., Williams, C. C., Norton, A., Hassall, C. D., & Colino, F. L. (2017). Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. *Frontiers in Neuroscience, 11*, 109. https://doi.org/10.3389/fnins.2017.00109

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121. https://doi.org/10.1037/0022-3514.77.6.1121

Lacey, J. I. (1959). Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome. In E. A. Rubinstein & M. B. Parloff (Eds.), *Research in psychotherapy* (pp. 160–208). American Psychological Association. https://doi.org/10.1037/10036-010

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): still no identification without attention. *Psychological Review*, *111*(4), 880. https://doi.org/10.1037/0033-295X.111.4.880

Landry, S. J. (2009). Flight deck automation. In *Springer handbook of automation* (pp. 1215-1239). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78831-7_68

Landström, U. (1990). Noise and fatigue in working environments. *Environment International*, *16*(4-6), 471-476. https://doi.org/10.1016/0160-4120(90)90015-X

Leaver, A. M., & Rauschecker, J. P. (2010). Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *Journal of Neuroscience*, *30*(22), 7604-7612. https://doi.org/10.1523/JNEUROSCI.0296-10.2010

Lee, J.D. & Seppelt, B.D. (2012). Human factors and ergonomics in automation design. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics.* https://doi.org/10.1002/9781118131350.ch59

Lee, Y. H., & Liu, B. S. (2003). Inflight workload assessment: Comparison of subjective and physiological measurements. *Aviation, Space, and Environmental Medicine*, 74(10), 1078-1084.

Lee, H. P., Kumar, S., Garg, S., & Lim, K. M. (2022). Assessment of in-cabin noise of wide-body aircrafts. *Applied Acoustics*, 194, 108809. https://doi.org/10.1016/j.apacoust.2022.108809

Lees, M. N., Cosman, J., Lee, J. D., Vecera, S. P., Dawson, J. D., & Rizzo, M. (2012). Cross-modal warnings for orienting attention in older drivers with and without attention impairments. *Applied Ergonomics*, *43*(4), 768-776. https://doi.org/10.1016/j.apergo.2011.11.012

Lewkowicz, R., Stróżak, P., Bałaj, B., Francuz, P., & Augustynowicz, P. (2018). Selective auditory attention and spatial disorientation cues effect on flight performance. *Aerospace Medicine and Human Performance, 89*(11), 976–984. https://doi.org/10.3357/amhp.5153.2018

Leylekian, L., Lebrun, M., & Lempereur, P. (2014). An overview of aircraft noise reduction technologies. *Aerospace Lab,* (6), p-1. https://dx.doi.org/10.12762/2014.AL07-01

Li, W., Li, R., Xie, X., & Chang, Y. (2022). Evaluating mental workload during multitasking in simulated flight. *Brain and Behavior*, *12*(4), e2489. https://doi.org/10.1002/brb3.2489

Lin, Y., & Cai, H. (2009). A method for building a real-time cluster-based continuous mental workload scale. *Theoretical Issues in Ergonomics Science*, *10*(6), 531-543. https://doi.org/10.1080/14639220902836547

Litovsky, R. (2015). Development of the auditory system. *Handbook of Clinical Neurology*, *129*, 55-72. https://doi.org/10.1016/B978-0-444-62630-1.00003-2

Liu, Y. C. (2001). Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems. *Ergonomics, 44*(4), 425-442. https://doi.org/10.1080/00140130010011369

Liu, D., Jaramillo, M., & Vincenzi, D. (2015). The effects of system reliability and task uncertainty on autonomous unmanned aerial vehicle operator performance under high time pressure. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *25*(5), 515-522. https://doi.org/10.1002/hfm.20565

Liu, D., Peterson, T., Vincenzi, D., & Doherty, S. (2016). Effect of time pressure and target uncertainty on human operator performance and workload for autonomous unmanned aerial system. *International Journal of Industrial Ergonomics*, 51, 52-58. https://doi.org/10.1016/j.ergon.2015.01.010

Ljung, R., Sörqvist, P., Kjellberg, A., & Green, A. M. (2009). Poor listening conditions impair memory for intelligible lectures: Implications for acoustic classroom standards. *Building Acoustics, 16*(3), 257-265. https://doi.org/10.1260/135101009789877031

Longo, L., & Orru, G. (2019). An evaluation of the reliability, validity and sensitivity of three human mental workload measures under different instructional conditions in third-level education. In *International conference on computer supported education* (pp. 384-413). Springer, Cham. https://doi.org/10.1007/978-3-030-21151-6_19

Luo, Y., Wang, J., Liang, H. N., Luo, S., & Lim, E. G. (2021, August). Monoscopic vs. stereoscopic views and display types in the teleoperation of unmanned ground vehicles for object avoidance. In *2021 30th IEEE international conference on robot & human interactive communication (RO-MAN)* (pp. 418-425). IEEE. https://doi.org/10.1109/RO-MAN50785.2021.9515455

Loewen, L. J., & Suedfeld, P. (1992). Cognitive and arousal effects of masking office noise. *Environment and Behavior*, 24(3), 381-395. https://doi.org/10.1177/0013916592243006

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746-748. https://doi.org/10.1038/264746a0

McIntire, J. P., Havig, P. R., Watamaniuk, S. N., & Gilkey, R. H. (2010). Visual search performance with 3-d auditory cues: Effects of motion, target location, and practice. *Human Factors*, *52*(1), 41-53. https://doi.org/10.1177/0018720810368806

McIntosh, R. D., Fowler, E. A., Lyu, T., & Della Sala, S. (2019). Wise up: Clarifying the role of metacognition in the Dunning-Kruger effect. *Journal of Experimental Psychology: General*, *148*(11), 1882–1897. https://doi.org/10.1037/xge0000579

McKendrick, R., Shaw, T., de Visser, E., Saqer, H., Kidwell, B., & Parasuraman, R. (2014). Team performance in networked supervisory control of unmanned air vehicles: Effects of automation, working memory, and communication content. *Human Factors*, *56*(3), 463-475. https://doi.org/10.1177%2F0018720813496269

McLane, R. C., & Wierwille, W. W. (1975). The influence of motion and audio cues on driver performance in an automobile simulator. *Human Factors*, *17*(5), 488-501. https://doi.org/10.1177/001872087501700508

MacDonald, W. (2003). The impact of job demands and workload on stress and fatigue. *Australian Psychologist*, *38*(2), 102-117. https://doi.org/10.1080/00050060310001707107

Manjarrez, E., Mendez, I., Martinez, L., Flores, A., & Mirasso, C. R. (2007). Effects of auditory noise on the psychophysical detection of visual signals: cross-modal stochastic resonance. *Neuroscience Letters*, *415*(3), 231-236. https://doi.org/10.1016/j.neulet.2007.01.030

Mansikka, H., Simola, P., Virtanen, K., Harris, D., & Oksama, L. (2016). Fighter pilots' heart rate, heart rate variation and performance during instrument approaches. *Ergonomics*, *59*(10), 1344-1352. https://doi.org/10.1080/00140139.2015.1136699

Mansikka, H., Virtanen, K., & Harris, D. (2019). Comparison of NASA-TLX scale, modified Cooper–Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks. *Ergonomics*, *62*(2), 246-254. https://doi.org/10.1080/00140139.2018.1471159

Maragos P., Gros, P., Katsamanis, A., & Papandreou, G. (2008) Cross-modal integration for performance improving in multimedia: A review. In Maragos, P., Potamianos, A., Gros, P. (Eds.), *Multimodal Processing and Interaction*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-76316-3_1

Marquis-Favre, C., Premat, E., & Aubrée, D. (2005). Noise and its effects–a review on qualitative aspects of sound. Part II: Noise and annoyance. *Acta Acustica United with Acustica*, *91*(4), 626-642.

Marucci, M., Di Flumeri, G., Borghini, G., Sciaraffa, N., Scandola, M., Pavone, E. F., Babiloni, F., Betti, V., & Aricò, P. (2021). The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence. *Scientific Reports*, *11*(1), 1-15. https://doi.org/10.1038/s41598-021-84196-8

Mateeff, S., Hohnsbein, J., & Noack, T. (1985). Dynamic visual capture: apparent auditory motion induced by a moving visual target. *Perception*, *14*(6), 721-727. https://doi.org/10.1068/p140721

Mastoropoulou, G., Debattista, K., Chalmers, A., & Troscianko, T. (2005, August). The influence of sound effects on the perceived smoothness of rendered animations. In *Proceedings of the 2nd symposium on applied perception in graphics and visualizatio*n (pp. 9-15). https://doi.org/10.1145/1080402.1080404

Mazur, L. M., Mosaly, P. R., Hoyle, L. M., Jones, E. L., & Marks, L. B. (2013). Subjective and objective quantification of physician's workload and performance during radiation therapy planning tasks. *Practical Radiation Oncology*, *3*(4), e171-e177. https://doi.org/10.1016/j.prro.2013.01.001

Mehler, B., Reimer, B., Coughlin, J. F., & Dusek, J. A. (2009). Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record*, *2138*(1), 6-12. https://doi.org/10.3141/2138-02

Mehta, R., Zhu, R., & Cheema, A. (2012). Is noise always bad? Exploring the effects of ambient noise on creative cognition. *Journal of Consumer Research*, *39*(4), 784-799. https://doi.org/10.1086/665048

Mekdeci, B., & Cummings, M. L. (2009, September). Modeling multiple human operators in the supervisory control of heterogeneous unmanned vehicles. In *Proceedings of the 9th workshop on performance metrics for intelligent systems* (pp. 1-8). https://doi.org/10.1145/1865909.1865911

Melamed, S., & Bruhis, S. (1996). The effects of chronic industrial noise exposure on urinary cortisol, fatigue, and irritability: a controlled field experiment. *Journal of Occupational and Environmental Medicine*, *38*(3), 252-256. https://doi.org/10.1097/00043764-199603000-00009

Meredith, C., & Edworthy, J. (1995). Are there too many alarms in the intensive care unit? An overview of the problems. *Journal of Advanced Nursing, 21*(1), 15-20. https://doi.org/10.1046/j.1365-2648.1995.21010015.x

Miller, M., & Holley, S. (2017, July). SHELL Revisited: Cognitive Loading and Effects of Digitized Flight Deck Automation. In *International conference on applied human factors and ergonomics* (pp. 95-107). Springer, Cham. https://doi.org/10.1007/978-3-319-60642-2_9

Molesworth, B. R. C., Burgess, M., & Kwon, D. (2013). The use of noise cancelling headphones to improve concurrent task performance in a noisy environment. *Applied Acoustics, 74*(1), 110-115. https://doi.org/10.1016/j.apacoust.2012.06.015

Molesworth B. R. C., Burgess. M., Gunnell B., Löffler D., Venjakob A. (2014). The effect on recognition memory of noise cancelling headphones in a noisy environment with native and non-native speakers. *Noise Health, 16*(71), 240-247. https://doi.org/10.4103/1463-1741.137062

Molesworth, B. R. C., Burgess, M., & Zhou, A. (2015). The effects of noise on key workplace skills. *The Journal of the Acoustical Society of America, 138*(4), 2054-2061. https://doi.org/10.1121/1.4929741

Molesworth, B. R.C., (2016). Contextual factors impacting on aviation communication. In Estival, D., Farris, C. & Molesworth, B. (Eds.), *Aviation English* (pp. 111-139). Routledge.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, *27*, 84-92. https://doi.org/10.1016/0001-6918(67)90048-0

Moore, B. C. (2012). *An introduction to the psychology of hearing*. Brill.

Moore, D. S., McCabe, G. P., & Craig, B. A. (2009). *Introduction to the practice of statistics* (6th ed.). https://doi.org/10.1017/S0025557200003399

Moshkov, P. A. (2020). Acoustic characteristics of aircraft piston engines. *Russian Aeronautics*, *63*(4), 568-574. https://doi.org/10.3103/S1068799820040029

Ministry of Transport. (2018). *State of Israel Ministry of Transport and Road Safety, safety investigation report, serious incident file* No. 81-18. https://aviation-safety.net/reports/2018/20180814_R44_4X-BCR.pdf

Møller, A. (1983). *Auditory physiology*. New York: Academic Press.

Mukerji, N. (1957). Is "illusion" illusion?. *The Journal of General Psychology*, *57*(2), 209-212. https://doi.org/10.1080/00221309.1957.9920363

Müller, R., & Schneider, J. (2017). Noise exposure and auditory thresholds of German airline pilots: a cross-sectional study. *BMJ Open*, *7*(5), e012913. https://doi.org/10.1136/bmjopen-2016-012913

Murata, A. (2005). An attempt to evaluate mental workload using wavelet transform of EEG. *Human Factors*, *47*(3), 498-508. https://doi.org/10.1518/001872005774860096

Muzammil, M., & Hasan, F. (2004). Human performance under the impact of continuous and intermittent noise in a manual machining task. *Noise & Vibration Worldwide*, *35*(7), 10-15. https://doi.org/10.1260/0957456041589836

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence' in the auditory cortex. *Trends in Neurosciences*, *24*(5), 283-288. https://doi.org/10.1016/S0166-2236(00)01790-2

Namba, S., & Kuwano, S. (1988). Measurement of habituation to noise using the method of continuous judgment by category. *Journal of Sound and Vibration*, *127*(3), 507-511. https://doi.org/10.1016/0022-460X(88)90376-8

Narasimhan, S. G., & Nayar, S. K. (2002). Vision and the atmosphere. *International Journal of Computer Vision, 48*(3), 233-254. https://doi.org/10.1023/A:1016328200723

Naujoks, F., Mai, C., & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. *Advances in Human Aspects of Transportation*, *7*(Part I), 431.

Nava, E., & Pavani, F. (2013). Changes in sensory dominance during childhood: Converging evidence from the Colavita effect and the sound-induced flash illusion. *Child Development*, *84*(2), 604-616. https://doi.org/10.1111/j.1467-8624.2012.01856.x

Neiss, R. (1988). Reconceptualizing Arousal: Psychobiological States in Motor Performance. *Psychological Bulletin*, *103*(3), 345-366. https://doi.org/10.1037/0033-2909.103.3.345

NTSB. (2010). *National Transportation Safety Board aviation accident report, report number: NTSB/AAR-10/01 PB2010-910401*. https://www.ntsb.gov/investigations/AccidentReports/Reports/AAR1001.pdf

NTSB. (2014). *National Transportation Safety Board aviation accident report, report number: NTSB/AAR-14/01 PB2014-105984.* https://www.ntsb.gov/investigations/accidentreports/reports/aar1401.pdf

NTSB. (2017). *National Transportation Safety Board aviation accident report, report number:* DCA17IA202A. https://web.archive.org/web/20190223033039/https://app.ntsb.gov/pdfgenerator/ReportGeneratorFile.ashx?EventID=20170922X54600&AKey=1&RType=HTML&IType=IA

Norman, G. (2013). Working memory and mental workload. *Advances in Health Sciences Education*, *18*(2), 163-165. https://doi.org/10.1007/s10459-013-9451-y

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors*, *33*(1), 17-33. https://doi.org/10.1177/001872089103300102

Ogden, G. D., Levine, J. M., & Eisner, E. J. (1979). Measurement of workload by secondary tasks. *Human Factors*, *21*(5), 529-548.

Orban, G. A., Janssen, P., & Vogels, R. (2006). Extracting 3D structure from disparity. *Trends in Neurosciences, 29*(8), 466-473. https://doi.org/10.1016/j.tins.2006.06.012

Olmos, O., Wickens, C. D., & Chudy, A. (2000). Tactical displays for combat awareness: An examination of dimensionality and frame of reference concepts and the application of cognitive engineering. *The International Journal of Aviation Psychology*, *10*(3), 247-271. https://doi.org/10.1207/S15327108IJAP1003_03

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human Factors, 56*(3), 476-488. https://doi.org/10.1177/0018720813501549

Oskarsson, P. A., Eriksson, L., & Carlander, O. (2012). Enhanced perception and performance by multimodal threat cueing in simulated combat vehicle. *Human Factors*, *54*(1), 122-137. https://doi.org/10.1177%2F0018720811424895

Owen, R. S. (1991). Clarifying the simple assumption of the information load paradigm. *ACR North American Advances*, *18*, 552-557.

Ozcan, H.K., & Nemlioglu, S. (2006). In-cabin noise levels during commercial aircraft flights. *Canadian Acoustics, 34*(4), 31-35.

Paes, D., Kim, S., & Irizarry, J. (2017). Human factors considerations of first person view (FPV) operation of unmanned aircraft systems (UAS) in infrastructure construction and inspection environments. In *2017 CSCE annual conference and general meeting* (pp. CON252-1-CON252-10).

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230-253. https://doi.org/10.1518/001872097778543886

Park, J., & Jung, W. (2006). A study on the validity of a task complexity measure for emergency operating procedures of nuclear power plants—comparing with a subjective workload. *IEEE Transactions on Nuclear Science*, *53*(5), 2962-2970. https://doi.org/10.1109/TNS.2006.882149

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, *8*(5), 379-391. https://doi.org/10.1038/nrn2131

Pickles, J. O. (2015). Auditory pathways: anatomy and physiology. In *Handbook of clinical neurology* (Vol. 129, pp. 3-25). Elsevier.

Pick, H. L., Warren, D. H., & Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, *6*(4), 203-205. https://doi.org/10.3758/BF03207017

Pinna, B. (2013). Illusion and illusoriness: New perceptual issues and new phenomena. In Albertazzi, L. (Eds.), *Handbook of experimental phenomenology: Visual perception of shape, space and appearance.* 317-341. Wiley-Blackwell. https://doi.org/10.1002/9781118329016.ch13

Politi, E., I. Panagiotopoulos, I. Varlamis, and G. Dimitrakopoulos. (2021). A survey of UAS technologies to enable beyond visual line of sight (BVLOS) operations. *Proceedings of the 7th international conference on vehicle technology and intelligent transport systems* (VEHITS 2021) (pp. 505-512). https://doi.org/10.5220/0010446905050512

Posner, M., Nissen, M., & Klein, R. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review, 83*(2), 157-171. https://doi.org/10.1037/0033-295X.83.2.157

Poulton, E. C. (1978). A new look at the effects of noise: A rejoinder. *Psychological Bulletin*,85(5). https://doi.org/10.1037/0033-2909.85.5.1068

Poulton, E. C. (1979). Composite model for human performance in continuous noise. *Psychological Review*, *86*, 361–375. https://doi.org/10.1037/0033-295X.86.4.361

Prablanc, C., Echallier, J. F., Komilis, E., & Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target. *Biological Cybernetics*, *35*(2), 113-124. https://doi.org/10.1007/BF00337436

Previc, F. H. (2004). Visual illusions in flight. In Previc, F.H., Ercoline, W.R., (Eds.), *Spatial disorientation in aviation*, (pp. 283-322). AIAA.

Prinzel Iii, L. J., Pope, A. T., & Freeman, F. G. (2002). Physiological self-regulation and adaptive automation. *The International Journal of Aviation Psychology*, *12*(2), 179-196. https://doi.org/10.1207/S15327108IJAP1202_5

Rausch, V. H., Bauch, E. M., & Bunzeck, N. (2014). White noise improves learning by modulating activity in dopaminergic midbrain regions and right superior temporal sulcus. *Journal of Cognitive Neuroscience*, *26*(7), 1469-1480. https://doi.org/10.1162/jocn_a_00537

Rauscher, F.H., Shaw, G.L., & Ky, C.N. (1993). Music and spatial task performance. *Nature, 365*(6447), 611-611. https://doi.org/10.1038/365611a0

Recanzone, G. H. (2003). Auditory influences on visual temporal rate perception. *Journal of Neurophysiology*, *89*(2), 1078-1093. https://doi.org/10.1152/jn.00706.2002

Reeve, J. (2009). *Understanding motivation and emotion* (5th ed.). Hoboken, NJ: Wiley.

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

Reimer, B., Mehler, B., Coughlin, J. F., Godfrey, K. M., & Tan, C. (2009, September). An on-road assessment of the impact of cognitive workload on physiological arousal in young adult drivers. In *Proceedings of the 1st international conference on automotive user interfaces and interactive vehicular applications* (pp. 115-118). https://doi.org/10.1145/1620509.1620531

Reimer, B., & Mehler, B. (2011). The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation. *Ergonomics*, *54*(10), 932-942. https://doi.org/10.1080/00140139.2011.604431

Renshaw, P. F., & Wiggins, M. W. (2020). The sensitivity of Cue utilization when learning to operate a visual line of sight multi-rotor remotely piloted aircraft. *International Journal of Industrial Ergonomics, 77*, 102953. https://doi.org/10.1016/j.ergon.2020.102953

Roach, N. W., Heron, J., & McGraw, P. V. (2006). Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1598), 2159-2168. https://doi.org/10.1098/rspb.2006.3578

Robertson, J. A., Broomfield, N. M., & Espie, C. A. (2007). Prospective comparison of subjective arousal during the pre-sleep period in primary sleep-onset insomnia and normal sleepers. *Journal of Sleep Research*, *16*(2), 230-238. https://doi.org/10.1111/j.1365-2869.2007.00579.x

Rojas, D., Kapralos, B., Cristancho, S., Collins, K., Hogue, A., Conati, C., & Dubrowski, A. (2012, January). Developing effective serious games: The effect of background sound on visual fidelity perception with varying texture resolution. In *Medicine meets virtual reality* (pp. 386-392). https://doi.org/10.3233/978-1-61499-022-2-386

Rojas, D., Cowan, B., Kapralos, B., Collins, K., & Dubrowski, A. (2014, January). The effect of contextual sound cues on visual fidelity perception. In *Medicine meets virtual reality* (pp. 346-352). IOS Press.

Rosebrock, L. E., Hoxha, D., Norris, C., Cacioppo, J. T., & Gollan, J. K. (2017). Skin conductance and subjective arousal in anxiety, depression, and comorbidity: Implications for affective reactivity. *Journal of Psychophysiology, 31*(4), 145–157. https://doi.org/10.1027/0269-8803/a000176

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex, 17*(5), 1147-1153. https://doi.org/10.1093/cercor/bhl024

Rothman, K. (1990). No adjustments are needed for multiple comparisons. *Epidemiology, 1*(1), 43-46.

Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, *53*(1), 61-86. https://doi.org/10.1111/j.1464-0597.2004.00161.x

Ruiz, J. J., Escalera, M. A., Viguria, A., & Ollero, A. (2015a). A simulation framework to validate the use of head-mounted displays and tablets for information exchange with the UAV safety pilot. In *2015 Workshop on research, education and development of unmanned aerial systems (RED-UAS)* (pp. 336-341). IEEE. https://doi.org/10.1109/RED-UAS.2015.7441025

Ruiz, J. J., Viguria, A., Martinez-de-Dios, J. R., & Ollero, A. (2015b). Immersive displays for building spatial knowledge in multi-UAV operations. In *2015 international conference on unmanned aircraft systems (ICUAS)* (pp. 1043-1048). IEEE. https://doi.org/10.1109/ICUAS.2015.7152395

Sanbonmatsu, D. M., & Kardes, F. R. (1988). The effects of physiological arousal on information processing and persuasion. *Journal of Consumer Research*, *15*(3), 379-385. https://doi.org/10.1086/209175

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, *53*(1), 61-97. https://doi.org/10.1016/0001-6918(83)90016-1

Scerra, V. E., & Brill, J. C. (2012, September). Effect of task modality on dual-task performance, response time, and ratings of operator workload. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (Vol. 56, No. 1, pp. 1456-1460). SAGE Publications. https://doi.org/10.1177%2F1071181312561409

Schacter, D., Gilbert, D., Wegner, D., & Hood, B. M. (2011). *Psychology: European edition*. Macmillan International Higher Education.

Schellenberg, E. G., Nakata, T., Hunter, P. G., & Tamoto, S. (2007). Exposure to music and cognitive performance: Tests of children and adults. *Psychology of Music*, *35*(1), 5-19. https://doi.org/10.1177/0305735607068885

Schlüter, S. (2016, August). Determination of aircraft engine speed based on acoustic measurements. In *INTER-NOISE and NOISE-CON congress and conference proceedings* (Vol. 253, No. 4, pp. 4366-4373). Institute of Noise Control Engineering.

Scholey, A. B., Harper, S., & Kennedy, D. O. (2001). Cognitive demand and blood glucose. *Physiology & Behavior*, *73*(4), 585-592. https://doi.org/10.1016/S0031-9384(01)00476-0

Saeki, T., Fujii, T., Yamaguchi, S., & Harima, S. (2004). Effects of acoustical noise on annoyance, performance and fatigue during mental memory task. *Applied Acoustics*, *65*(9), 913-921. https://doi.org/10.1016/j.apacoust.2003.12.005

Secoli, R., Milot, M. H., Rosati, G., & Reinkensmeyer, D. J. (2011). Effect of visual distraction and auditory feedback on patient effort during robot-assisted movement training after stroke. *Journal of Neuroengineering and Rehabilitation*, *8*(1), 1-10. https://doi.org/10.1186/1743-0003-8-21

Self, B. P., Ercoline, W. R., Olson, W. A., & Tvaryanas, A. P. (2006). 10. Spatial disorientation in uninhabited aerial vehicles. In *Human factors of remotely operated vehicles* (Vol. 7, pp. 133-146). Emerald Group Publishing Limited. https://doi.org/10.1016/S1479-3601(05)07010-4

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport, 16*(17), 1923-1927. https://doi.org/10.1097/01.wnr.0000187634.68504.bb

Shen, J. (2003). On the foundations of vision modeling: I. Weber's law and Weberized TV restoration. *Physica D: Nonlinear Phenomena*, *175*(3-4), 241-251. https://doi.org/10.1016/S0167-2789(02)00734-0

Sheridan, T. B. (2018). Comments on "Issues in human–automation interaction modeling: presumptive aspects of frameworks of types and levels of automation" by David B. Kaber. *Journal of Cognitive Engineering and Decision Making*, *12*(1), 25-28. https://doi.org/10.1177/1555343417724964

Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. Massachusetts Inst of Tech Cambridge Man-Machine Systems Lab.

Shimizu, T., Makishima, K., Yoshida, M., & Yamagishi, H. (2002). Effect of background noise on perception of English speech for Japanese listeners. *Auris Nasus Larynx*, *29*(2), 121-125. https://doi.org/10.1016/S0385-8146(01)00133-X

Shively, R. J., Hobbs, A., Lyall, B., & Rorie, C. (2015, March). Human performance considerations for Remotely Piloted Aircraft Systems (RPAS). In *ICAO RPAS symposium (*pp.1-68*).* Montreal, Quebec, Canada.

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic Bulletin & Review*, *20*(1), 21-53. https://doi.org/10.3758/s13423-012-0333-8

Simpson, B., Bolia, R., & Draper, M. (2013). Spatial audio display concepts supporting situation awareness for operators of unmanned aerial vehicles. *Human Performance, Situation Awareness, and Automation: Current Research and Trends HPSAA II*, 2, 61-65.

Slutsky, D. A., & Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *NeuroReport*, *12*(1), 7-10. https://doi.org/10.1097/00001756-200101220-00009

Smith, R. P. (1981). Boredom: A review. *Human Factors*, *23*(3), 329-340. https://doi.org/10.1177/001872088102300308

Smith, A. P. (1985). The effects of noise on the processing of global shape and local detail. *Psychological Research*, *47*(2), 103-108. https://doi.org/10.1007/BF00309124

Smolyanskiy, N., & Gonzalez-Franco, M. (2017). Stereoscopic first person view system for drone navigation. *Frontiers in Robotics and AI*, 11. https://doi.org/10.3389/frobt.2017.00011

Snell, R. S., & Lemp, M. A. (2013). The eyeball. In *Clinical Anatomy of the Eye* (pp. 132-213). John Wiley & Sons. https://doi.org/10.1002/9781118690987

Söderlund, G., Sikström, S., & Smart, A. (2007). Listen to the noise: noise is beneficial for cognitive performance in ADHD. *Journal of Child Psychology and Psychiatry*, *48*(8), 840-847. https://doi.org/10.1111/j.1469-7610.2007.01749.x

Solomons, L. M. (1900). A new explanation of Weber's Law. *Psychological Review*, *7*(3), 234. https://doi.org/10.1037/h0065919

Sörqvist, P. (2010). The role of working memory capacity in auditory distraction: A review. *Noise and Health*, *12*(49), 217. https://doi.org/10.4103/1463-1741.70500

Spence, C. (2009). Explaining the Colavita visual dominance effect. *Progress in Brain Research*, *176*, 245-258. https://doi.org/10.1016/S0079-6123(09)17615-X

Staggers, N., & Norcio, A. F. (1993). Mental models: concepts for human-computer interaction research. *International Journal of Man-Machine Studies*, *38*(4), 587-605. https://doi.org/10.1006/imms.1993.1028

Stanton, N. A., Chambers, P. R., & Piggott, J. (2001). Situational awareness and safety. *Safety Science, 39*(3), 189-204. https://doi.org/10.1016/S0925-7535(01)00010-8

Stearman, O. (1997). Aircraft damage detection from acoustic and noise impressed signals found by a cockpit voice recorder. *The Journal of the Acoustical Society of America, 101*(5), 3085. https://doi.org/10.1121/1.418804

Stein, B. E. (1998). Neural mechanisms for synthesizing sensory information and producing adaptive behaviors. *Experimental Brain Research*, *123*(1), 124-135. https://doi.org/10.1007/s002210050553

Stenfelt, S., & Rönnberg, J. (2009). The Signal-Cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian Journal of Psychology*, *50*(5), 385-393. https://doi.org/10.1111/j.1467-9450.2009.00748.x

Stevenson, J. D., O'Young, S., & Rolland, L. (2015a). Assessment of alternative manual control methods for small unmanned aerial vehicles. *Journal of Unmanned Vehicle Systems*, *3*(3), 73-94. https://doi.org/10.1139/juvs-2015-000

Stevenson, J. D., O'Young, S., & Rolland, L. (2015b). Beyond line of sight control of small unmanned aerial vehicles using a synthetic environment to augment first person video. *Procedia Manufacturing, 3*, 960-967. https://doi.org/10.1016/j.promfg.2015.07.146

Stevenson, R. A., Schlesinger, J. J., & Wallace, M. T. (2013). Effects of divided attention and operating room noise on perception of pulse oximeter pitch changes: a laboratory study. *Anesthesiology*, *118*(2), 376-381. https://doi.org/10.1097/ALN.0b013e31827d417b

Stone, H., & Bosley, J. J. (1965). Olfactory discrimination and Weber's Law. *Perceptual and Motor Skills*, *20*(2), 657-665. https://doi.org/10.2466/pms.1965.20.2.657

Stough, C., Kerkin, B., Bates, T., & Mangan, G. (1994). Music and spatial IQ. *Personality and Individual Differences*, *17*(5), 695. https://doi.org/10.1016/0191-8869(94)90145-7

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643. https://doi.org/10.1037/h0054651

Sun, X., Chen, P. H., & Rau, P. L. P. (2022). Do congruent auditory stimuli facilitate visual search in dynamic environments? An experimental study based on multisensory interaction. *Multisensory Research*, *35*, 391-405. https://doi.org/10.1163/22134808-bja10075

Swiss Transportation Safety Investigation Board. (2019). *Rapport sommaire, HB-ZYZ.* https://www.sust.admin.ch/inhalte/AV-berichte/HB-ZYZ_F.pdf

Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., & Eggemeier, F. T. (2004). Effects of Sensory Modality and Task Duration on Performance, Workload, and Stress in Sustained Attention. *Human Factors*. https://doi.org/10.1518/hfes.46.2.219.37334

Szalma, J. L., & Hancock, P. A. (2011). Noise effects on human performance: a meta-analytic synthesis. *Psychological Bulletin*, *137*(4), 682. https://doi.org/10.1037/a0023987

Tabachnick, B. G., & Fidell, L.S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Tang, N. K., & Harvey, A. G. (2004). Effects of cognitive arousal and physiological arousal on sleep perception. *Sleep*, *27*(1), 69-78. https://doi.org/10.1093/sleep/27.1.69

Tannen, R. S., Nelson, W. T., Bolia, R. S., Warm, J. S., & Dember, W. N. (2004). Evaluating adaptive multisensory displays for target localization in a flight task. *The International Journal of Aviation Psychology*, *14*(3), 297-312. https://doi.org/10.1207/s15327108ijap1403_5

Tao, D., Tan, H., Wang, H., Zhang, X., Qu, X., & Zhang, T. (2019). A systematic review of physiological measures of mental workload. *International Journal of Environmental Research and Public Health*, *16*(15), 2716. https://doi.org/10.3390%2Fijerph16152716

Teigen, K. H. (1994). Yerkes-Dodson: A law for all seasons. *Theory & Psychology*, *4*(4), 525-547. https://doi.org/10.1177/0959354394044004

Thayer, R. E. (1978). Toward a psychological theory of multidimensional activation (arousal). *Motivation and Emotion*, *2*(1), 1-34. https://doi.org/10.1007/BF00992729

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis & Prevention*, *35*(3), 381-391. https://doi.org/10.1016/S0001-4575(02)00014-3

Tjolleng, A., Jung, K., Hong, W., Lee, W., Lee, B., You, H., Son, J., & Park, S. (2017). Classification of a Driver's cognitive workload levels using artificial neural network on ECG signals. *Applied Ergonomics, 59*, 326-332. https://doi.org/10.1016/j.apergo.2016.09.013

Transportation Safety Board of Canada. (2021). *Air transportation safety investigation A21O0069*. https://bit.ly/3EjCoQ5

Tredici, T. J. (1980). Visual illusions as a probable cause of aircraft accidents. In *AGARD conference proceedings, spatial disorientation in flight: Current problems*, (pp B51-B55). AGARD.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242-248. https://doi.org/10.1080/17470216008416732

Treisman, A. (1964). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, 3(6), 449-459. https://doi.org/10.1016/S0022-5371(64)80015-3

Tremblay, S., Nicholls, A. P., Alford, D., & Jones, D. M. (2000). The irrelevant sound effect: Does speech play a special role?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(6), 1750. https://doi.org/10.1037/0278-7393.26.6.1750

Tsang, P. S., & Velazquez, V. L. (1996). Diagnosticity and multidimensional subjective workload ratings. *Ergonomics*, *39*(3), 358-381. https://doi.org/10.1080/00140139608964470

Tuccio, W., Schuster, B., & Gregor, J. (2017, June). Deriving Engine Power from a Cockpit Voice Recording in an Accident Investigation. In *Audio Engineering Society conference: 2017 AES international conference on audio forensics*. Audio Engineering Society.

Tvaryanas, A. P., Thompson, W. T., & Constable, S. H. (2006). Human factors in remotely piloted aircraft operations: HFACS analysis of 221 mishaps over 10 years. *Aviation, Space, and Environmental Medicine*, *77*(7), 724-732.

Ünal, A. B., de Waard, D., Epstude, K., & Steg, L. (2013). Driving with music: Effects on arousal and performance. *Transportation Research Part F: Traffic Psychology and Behaviour*, *21*, 52-65. https://doi.org/10.1016/j.trf.2013.09.004

United States Air Force (January, 17, 2007a). *Aircraft accident investigation MQ-1B Predator, Deployed Location,* S/N 03-3112. https://bit.ly/3V4aRJc

United States Air Force (March 26, 2007b). *Aircraft accident investigation MQ-1B Predator, Kandahar Airbase, Afghanistan, MQ-1B,* S/N 03-3104. https://bit.ly/3UXGtjI

United States Air Force (April, 2012a). *Aircraft accident investigation MQ-9A, Victoria, Republic of Seychelles, MQ-1B,* S/N 05-000102. https://bit.ly/3UQTzzI

United States Air Force (October, 2012b). *Abbreviated Aircraft Accident Investigation, Board Report, Jalalabad Air Base, Afghanistan, MQ-1B*, T/N 07-3201. https://bit.ly/3ED0BkL

United States Air Force. (February 14, 2013). *Abbreviated Aircraft Accident Investigation Board Report, Douglas County, Nevada, MQ-9,* S/N 09-004065. https://bit.ly/3EqmhR0

United States Air Force. (July, 2015). *United States Air Force aircraft accident investigation board report, MQ-9A* , T/N 04-4011. https://bit.ly/3hUjOGL

United States Air Force. (July, 2016). *United States Air Force aircraft accident investigation board report, MQ-1B* , T/N 05-3136. https://bit.ly/3Atj9ma

United States Air Force. (March, 2017). *United States Air Force aircraft accident investigation board report*, *MQ-9A*, T/N 10-4113. https://bit.ly/3ULwM89

Van der Groen, O., & Wenderoth, N. (2016). Transcranial random noise stimulation of visual cortex: stochastic resonance enhances central mechanisms of perception. *Journal of Neuroscience*, *36*(19), 5289-5298. https://doi.org/10.1523/JNEUROSCI.4519-15.2016

Van der Stoep, N., Van der Stigchel, S., Nijboer, T. C. W., & Spence, C. (2017). Visually induced inhibition of return affects the integration of auditory and visual information. *Perception*, *46*(1), 6-17. https://doi.org/10.1177/0301006616661934

VandenBos, G.R. (Ed.)(2015). *APA dictionary of psychology, 2nd ed*., American Psychological Association, Washington, DC.

Verwey, W. B., & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology: Applied*, *2*(3), 270. https://doi.org/10.1037/1076-898X.2.3.270

Volante, W. G., Merz, M., Stowers, K., & Hancock, P. A. (2016, September). Sleep, workload and boredom: subject matter expert insights. In *Proceedings of the Human Factors and Ergonomics Society annual meeting* (pp. 1833-1837). Sage CA: Los Angeles. https://doi.org/10.1177/1541931213601418

Wada, Y., Kitagawa, N., & Noguchi, K. (2003). Audio–visual integration in temporal perception. *International Journal of Psychophysiology*, *50*(1-2), 117-124. https://doi.org/10.1016/S0167-8760(03)00128-4

Wagner, S., Bareiß, R., & Guidati, G. (1996). Noise and its effects. In *Wind turbine noise* (pp. 13-24). Springer Science & Business Media.

Walker, L. A. (1997). *Flight testing the X-36: The test pilot's perspective* (No. NASA-CR-198058). https://ntrs.nasa.gov/citations/19970031950

Wang, L., Gao, S., Tan, W., & Zhang, J. (2022). Pilots' mental workload variation when taking a risk in a flight scenario: a study based on flight simulator experiments. *International Journal of Occupational Safety and Ergonomics*, 1-10. https://doi.org/10.1080/10803548.2022.2049101

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 183–200). Lawrence Erlbaum Associates, Inc.

Weibel, R. E., & Hansman, R. J. (2005). *Safety considerations for operation of UAV in the national airspace system* (No. ICAT-2005-1). Cambridge, MA: MIT International Center for Air Transportation.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*(3), 638. https://doi.org/10.1037/0033-2909.88.3.638

White, T. L., & Hancock, P. A. (2020). Specifying advantages of multi-modal cueing: Quantifying improvements with augmented tactile information. *Applied Ergonomics*, *88*, 103146. https://doi.org/10.1016/j.apergo.2020.103146

Whitlock, C. (2014a). Stop saying 'uh-oh' while you're flying': Drone crash pilot quotes unveiled. *The Washington Post.* https://wapo.st/3tDf8aX

Whitlock, C. (2014b). When drones fall from the sky. *The Washington Post.* Retrieved from Washington Post website: https://wapo.st/3Ohb2Pa

Wickens, C.D. (1984). Processing resources in attention. In Parasuraman, R., & Davies., D.R. (Eds.), *Varieties of attention* (pp. 63-102). London: Academic Press.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*(2), 159-177. https://doi.org/10.1080/14639220210123806

Wickens, C. D. (2005). Attentional tunneling and task management. In *2005 international symposium on aviation psychology* (pp. 812-817). Wright State University: CORE Scholar.

Wickens, C. (2018). Automation stages & levels, 20 years after. *Journal of Cognitive Engineering and Decision Making*, *12*(1), 35-41. https://doi.org/10.1177/1555343417727438

Wickens, C. D., & Alexander, A. L. (2009). Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology*, *19*(2), 182-199. https://doi.org/10.1080/10508410902766549

Wickens, C. D., Lee, J., Liu, Y. D., & Gordon-Becker, S. (2004). *An introduction to human factors engineering (2nd ed.).* Upper Saddle River N.J.: Pearson Prentice Hall.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2021). *Engineering psychology and human performance (Fifth edition)*. Routledge. https://doi.org/10.4324/9781003177616

Wickens, C., & Tsang, P. S. (2015). Workload. In D. A. Boehm-Davis, F. T. Durso, & J. D. Lee (Eds.), *APA handbook of human systems integration* (pp. 277–292). American Psychological Association. https://doi.org/10.1037/14528-018

Widyanti, A., Johnson, A., & de Waard, D. (2013a). Adaptation of the rating scale mental effort (RSME) for use in Indonesia. *International Journal of Industrial Ergonomics*, *43*(1), 70-76. https://doi.org/10.1016/j.ergon.2012.11.003

Widyanti, A., de Waard, D., Johnson, A., & Mulder, B. (2013b). National culture moderates the influence of mental effort on subjective and cardiovascular measures. *Ergonomic*s, *56*(2), 182-194. https://doi.org/10.1080/00140139.2012.748219

Wierwille, W. W., & Casali, J. G. (1983, October). A validated rating scale for global mental workload measurement applications. In *Proceedings of the Human Factors Society annual meeting 27*(2), 129-133. SAGE Publications.

Wierwille, W. W., & Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors*, *35*(2), 263-281. https://doi.org/10.1177/001872089303500205

Williams, K. W. (2004). *A summary of unmanned aircraft accident/incident data: Human factors implications*. (No. DOT/FAA/AM-04/24). https://apps.dtic.mil/dtic/tr/fulltext/u2/a460102.pdf

Williams, M. A., LaMarche, J. A., Alexander, R. W., Stanford, L. D., Fielstein, E. M., & Boll, T. J. (1996). Serial 7s and Alphabet Backwards as brief measures of information processing speed. *Archives of Clinical Neuropsychology, 11*(8), 651-659. https://doi.org/10.1016/S0887-6177(96)80002-3

Xiao, Y. M., Wang, Z. M., Wang, M. Z., & Lan, Y. J. (2005). The appraisal of reliability and validity of subjective workload assessment technique and NASA-task load index. *Chinese Journal of Industrial Hygiene and Occupational Diseases*, *23*(3), 178-181.

Yang, J., Sasikumar, P., Bai, H., Barde, A., Sörös, G., & Billinghurst, M. (2020). The effects of spatial auditory and visual cues on mixed reality remote collaboration. *Journal on Multimodal User Interfaces*, *14*(4), 337-352. https://doi.org/10.1007/s12193-020-00331-1

Yerkes R.M., & Dodson J. D., (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*. *18*(5): 459–482. https://doi.org/10.1002/cne.920180503

Young, M. S., & Stanton, N. A. (2004). Mental workload. In *Handbook of human factors and ergonomics methods* (pp. 416-426). CRC Press.

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

Zaehle, T., Wüstenberg, T., Meyer, M., & Jäncke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. *European Journal of Neuroscience*, *20*(9), 2447-2456. https://doi.org/10.1111/j.1460-9568.2004.03687.x

Zatorre, R. J., Bouffard, M., Ahad, P., & Belin, P. (2002). Where is 'where' in the human auditory cortex?. *Nature Neuroscience*, *5*(9), 905. https://doi.org/10.1038/nn904

Zeier, H. (1994). Workload and psychophysiological stress reactions in air traffic controllers. *Ergonomics*, *37*(3), 525-539. https://doi.org/10.1080/00140139408963668

Zhou, Y., He, J., Wang, Z., & Fu, S. (2014, April). Study on the correlation between physiological and subjective workload measurements. In *2014 international conference on information science, electronics and electrical engineering* (Vol. 1, pp. 308-312). IEEE. https://doi.org/10.1109/InfoSEEE.2014.6948121

Zmarz, A., Rodzewicz, M., Dąbski, M., Karsznia, I., Korczak-Abshire, M., & Chwedorzewska, K. J. (2018). Application of UAV BVLOS remote sensing data for multi-faceted analysis of Antarctic ecosystem. *Remote Sensing of Environment*, *217*, 375-388. https://doi.org/10.1016/j.rse.2018.08.031

APPENDIX 1: Experiment 1 Participant Information Sheet & Consent Form

School of Aviation	UNSW AUSTRALIA
PARTICIPANT INFORMATION STATEMENT AND CONSENT FORM Remotely Piloted Aircraft: A Human Factors Investigation into Human Performance Limitations Dr. Brett Molesworth (Chief Investigator) & Mr. Matthew Dunn (Student Investigator) HC Number: HC180410	

What is the research study about?

You are invited to take part in this research study. The aim of the proposed research is to investigate the effect of different visual displays and auditory information on the operational performance of a remotely piloted aircraft (RPA) operator.

Who is conducting this research?

The study is being carried out by Mr. Matthew Dunn (Student Investigator - SI) under the supervision of Dr. Brett Molesworth (Chief Investigator) from the School of Aviation at UNSW Sydney.

Inclusion/Exclusion Criteria

Before you decide to participate in this research project, we need to ensure that it is okay for you to take part. The research study is looking to recruit 2nd Year and 3rd Year Bachelor of Aviation (Flying) students who meet the following criteria:

Pilot (cleared to fly solo)

No previous Remotely Piloted Aircraft (RPA) / drone flying experience

This research project is not related to the Bachelor of Aviation (Flying) course curriculum, and is voluntary.

Do I have to take part in this research study?

Participation in any research project is voluntary. If you do not want to take part, you do not have to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.

If you decide you want to take part in the research study, you will be asked to:

- Read the information carefully (ask questions if necessary; take the form with you);
- Sign the consent form;
- Complete a series of short questionnaires; and
- Complete a series of flights using a remotely piloted aircraft (RPA).

What does participation in this research require, and are there any risks involved?

If you decide to take part in the research study, the research team will ask you to complete the following tasks:

- Complete demographics questionnaire (3 minutes)
- Complete one pre-task motion-sickness questionnaire and one after each flight (average time to complete each questionnaire – 1 minutes).

- Participate in a briefing about an RPA and live demonstration of its functions and controls (5 minutes).
- Complete pre-experiment test flights under three different visual conditions (minimum 10 minutes):
    - Visual Line of Sight (VLOS);
    - Beyond Visual Line of Sight (BVLOS) using First Person View – Goggle (FPV-G);
    - Beyond Visual Line of Sight (BVLOS) using First Person View – Monitor (FPV-M).
- Complete six experimental tasks using the RPA under three different visual conditions and two different auditory conditions (20 minutes).

With added breaks as required and to transition between tasks, the research will take no longer than 1 hour and 30 minutes to complete. We don't expect the RPA flying task to cause any harm or discomfort, however if you experience feelings of distress as a result of participation in this study you can let the research team know and the research will be ceased. You will also be offered assistance.

Completion of Questionnaires

The research team will ask you to complete a series of short questionnaires. The first questionnaire relates to demographic information such as age, gender and flying experience. The following questionnaires are designed to examine motion-sickness, and hence you will be asked to complete one at the start of the research, and another after each time you have flown the RPA/drone. While we do not expect that you will suffer severe motion-sickness from operating the RPA/drone, it is important that if you do, that we are aware of this. If by the remote chance that you do experience severe motion-sickness, it is important that you tell us immediately, and we will cease the research. You will be provided an area to rest, as well as water until you recover.

Completion of RPA flights

The research also involves completing a number of flights using an RPA. Since flying an RPA will be new to you, we will provide you training. During the training, you will be provided sufficient time to feel comfortable flying. At the conclusion of the training, we will ask you to complete three small tasks, allowing us to assess your competency. After this, you will be asked to complete four additional flights. Each flight should last approximately 15 minutes. There is a small chance that you may lose control of the drone during training flights and hence it impacts you. In order to protect against this, we have taken a number of steps, including, placing a geofence around the area where the RPA will be flown; you are outside of this area, inserted barriers where possible, and installed guards on the propellers. During the experiment flights, you will be located inside a classroom adjacent to the hangar and will be protected by the walls and windows.

What are the possible benefits to participation?

We hope to use information we get from this research study to benefit others who will be operating RPA in the future. The technology is rapidly evolving, and we are using currently available technology to determine the most appropriate methods to control RPA in potential situations such as medical emergencies, search and rescue, humanitarian aid and post-disaster events. By understanding the limitations of human performance, new and safer autonomous systems can be developed in future which are constructed around the human operator in the instance of an emergency scenario.

What will happen to information about me?

By signing the consent form you consent to the research team collecting and using information about you for the research study. We will keep your data for 7 years in accordance with the National Ethics Guidelines. We will store information about you in a non-identifiable format securely within the UNSW School of Aviation Kensington campus. Your non-identifiable information will only be used for the purpose of this research study only and for statistical analysis. It will be disclosed in academic publications and presentations in the future.

How and when will I find out what the results of the research study are?

The research team intend to publish and report the results of the research study in a variety of ways. All information published will be done in a way that will not identify you.

If you would like to receive a copy of the results you can let the research team know by adding your email or postal address within the consent form. We will only use these details to send you the results of the research. The results will also be made available via the school's website, under the news tab. www.aviation.unsw.edu.au

What if I want to withdraw from the research study?

If you do consent to participate, you may withdraw at any time. You can do so by completing the 'Withdrawal of Consent Form' which is provided at the end of this document. Alternatively, you can ring the research team and tell them you no longer want to participate. If you decide to leave the research study, the researchers will not collect additional information from you. Any identifiable information about you will be withdrawn from the research project. Your decision not to participate will not affect your relationship with UNSW Sydney or the UNSW Aviation.

What should I do if I have further questions about my involvement in the research study?

The person you may need to contact will depend on the nature of your query. If you want any further information concerning this project or if you have any problems which may be related to your involvement in the project, you can contact the following member of the research team:

Research Team Contact Details

Name	Dr Brett Molesworth
Position	Chief Investigator
Telephone	+ 61 2 9385 6757
Email	b.molesworth@unsw.edu.au

Support Services Contact Details

If at any stage during the project, you become distressed or require additional support from someone not involved in the research please call:

Name/Organisation	Bankstown Hospital Medical Centre
Telephone	+ 61 2 9722 8453
Fax	+ 61 2 9722 8570

What if I have a complaint or any concerns about the research study?

If you have any complaints about any aspect of the project, the way it is being conducted, then you may contact:

Complaints Contact

Position	Human Research Ethics Coordinator
Telephone	+ 61 2 9385 6222
Email	humanethics@unsw.edu.au
HC Reference Number	HC180410

Consent Form – Participant providing own consent

Declaration by the participant

- ☐ I understand I am being asked to provide consent to participate in this research project;
- ☐ I have read the Participant Information Sheet, or someone has read it to me in a language that I understand;
- ☐ I understand the purposes, study tasks and risks of the research described in the project;
- ☐ I understand that the research team will collect personal information about me which will be de-identified for statistical purposes including age, sex, flying rating, endorsements and hours; I provide my consent for this to happen.
- ☐ I provide my consent for the information collected about me to be used for the purpose of this research study only.
- ☐ I have had an opportunity to ask questions and I am satisfied with the answers I have received;
- ☐ I freely agree to participate in this research study as described and understand that I am free to withdraw at any time during the project and withdrawal will not affect my relationship with any of the named organisations and/or research team members;
- ☐ I would like to receive a copy of the study results via email or post, I have provided my details below and ask that they be used for this purpose only;

Name: ___________________________________

Address: _________________________________

Email Address: ____________________________

I understand that I will be given a signed copy of this document to keep;

Participant Signature

Name of Participant (please print)	
Signature of Research Participant	
Date	

Declaration by Researcher*

I have given a verbal explanation of the research study, its study activities and risks and I believe that the participant has understood that explanation.

Researcher Signature*

Name of Researcher (please print)	
Signature of Researcher	
Date	

***An appropriately qualified member of the research team must provide the explanation of, and information concerning the research study.**

Note: All parties signing the consent section must date their own signature.

Form for Withdrawal of Participation

I wish to **WITHDRAW** my consent to participate in the research proposal described above and understand that such withdrawal **WILL NOT** affect my relationship with The University of New South Wales, or the UNSW School of Aviation. In withdrawing my consent, I would like any information which I have provided for the purpose of this research project withdrawn.

Participant Signature

Name of Participant (Please print)	
Signature of Research Participant	
Date	

The section for Withdrawal of Participation should be forwarded to:

CI Name:	Dr. Brett Molesworth
Email:	b.molesworth@unsw.edu.au
Phone:	02 9385 6757
Postal Address:	School of Aviation Old Main Building Second Floor, Room 205 The University of New South Wales Sydney, NSW Australia 2052

APPENDIX 2: The NASA-TLX Questionnaire

NASA Task Load Index

Participant	Task	Date

Mental Demand How mentally demanding was the task?

Very Low Very High

Physical Demand How physically demanding was the task?

Very Low Very High

Temporal Demand How hurried or rushed was the pace of the task?

Very Low Very High

Performance How successful were you in accomplishing what you were asked to do?

Very Low Very High

Effort How hard did you have to work to accomplish your level of performance?

Very Low Very High

Frustration How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low Very High

APPENDIX 3: Experiment 2 Participant Information Sheet & Consent Form

School of Aviation	UNSW AUSTRALIA

PARTICIPANT INFORMATION STATEMENT AND CONSENT FORM
Remotely Piloted Aircraft: A Human Factors Investigation into Human Performance Limitations
Dr. Brett Molesworth (Chief Investigator) & Mr. Matthew Dunn (Student Investigator)
HC Number: HC190808

What is the research study about?

You are invited to take part in this research study. The aim of the proposed research is to investigate the effect of auditory feedback and workload on the operational performance of a remotely piloted aircraft (RPA) operator.

Who is conducting this research?

The study is being carried out by Mr. Matthew Dunn (Student Investigator - SI) under the supervision of A/Prof. Brett Molesworth (Chief Investigator) from the School of Aviation at UNSW Sydney.

Inclusion/Exclusion Criteria

Before you decide to participate in this research project, we need to ensure that it is okay for you to take part. The research study is looking to recruit people who meet the following criteria:

Hold an Aviation Reference Number (ARN) from the Civil Aviation Safety Authority (CASA).

Do I have to take part in this research study?

Participation in any research project is voluntary. If you do not want to take part, you do not have to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage. Your decision to withdraw from the study will have no impact on your studies at UNSW Sydney.

If you decide you want to take part in the research study, you will be asked to:

- Read the information carefully (ask questions if necessary; take the form with you);
- Sign the consent form;
- Complete a series of short questionnaires; and
- Complete a series of flights using a remotely piloted aircraft (RPA).

What does participation in this research require, and are there any risks involved?

If you decide to take part in the research study, we will ask you to complete the following tasks:

- Complete one Demographics Questionnaire (5 minutes).
- Complete one Personality Questionnaire (10 minutes).
- Complete one pre-task and subsequent post-task Workload Questionnaires (1 minute per questionnaire; 5 minutes total).
- Participate in a briefing about an RPA and live demonstration of its functions and controls (5 minutes).
- Complete a training flight with the RPA until you feel confident flying (10 minutes).
- Successfully complete three pre-task test flights using the RPA to demonstrate proficiency (5 minutes).
- Complete four main experiment flights (5 minutes each – total time 20 minutes).

With added breaks as required and to transition between tasks, the research will take no longer than 1 hour to complete.

Completion of Questionnaires

The research team will ask you to complete a series of short questionnaires. The first questionnaire relates to demographic information such as age, gender and flying experience. The second questionnaire relates to personality and measures personal traits such as extraversion, agreeableness and conscientiousness. The third questionnaire will be completed initially for baseline information and then at the completion of each flying task and relates to your perceived workload of the task.

Completion of RPA/Drone flights

The research also involves completing a number of flights using an RPA/drone. You will receive a briefing and will be trained and provided a demonstration before undertaking the main experimental tasks. During the training, you will be provided sufficient time for you to feel comfortable flying. Following this training we will ask you to complete three short tasks to demonstrate proficiency before proceeding to the main experiment. You will then be asked to complete four additional flights. Each flight should last approximately 5 minutes. There is a very small chance that you may lose control of the drone during training flights. In order to protect against this, we have taken a number of steps, including securing the area where the RPA/drone will be flown and restricting outside access, inserting barriers where possible and signage for the perimeter exclusion zone, installing guards on the propellers, and most importantly operating the RPA behind a translucent barrier. Additionally, the Student Investigator who will be training you holds a Remote Pilot Licence (RePL) issued by CASA.

What are the possible benefits to participation?

We hope to use information we get from this research study to benefit others who will be operating RPAS in the future. The technology is rapidly evolving, and we are using currently available technology to determine the most appropriate methods to control RPAS in potential situations such as medical emergencies, search and rescue, humanitarian aid and post-disaster events. By understanding the limitations of human performance, new and safer human-centred autonomous systems can be developed.

What will happen to information about me?

By signing the consent form, you consent to the research team collecting and using information about you for the research study. We will keep your data for 7 years after the completion of the study in accordance with the National Ethics Guidelines. We will store information about you in a non-identifiable format securely within the UNSW School of Aviation Kensington campus. Your information will only be used for statistical analysis and will be disclosed in academic publications and presentations in the future.

How and when will I find out what the results of the research study are?

The research team intend to publish and report the results of the research study in a variety of ways. All information published will be done in a way that will not identify you. If you would like to receive a copy of the results you can let the research team know by adding your email or postal address within the consent form. We will only use these details to send you the results of the research.

What if I want to withdraw from the research study?

If you do consent to participate, you may withdraw at any time. You can do this by leaving the study site. If you withdraw from the research, we will destroy any information that has been collected. Once you have submitted the questionnaires however, we will not be able to withdraw your responses as the questionnaires are anonymous.

What should I do if I have further questions about my involvement in the research study?

The person you may need to contact will depend on the nature of your query. If you want any further information concerning this project or if you have any problems which may be related to your involvement in the project, you can contact the following member/s of the research team:

Research Team Contact Details

Name	Associate Professor Brett Molesworth
Position	Chief Investigator
Telephone	+ 61 2 9385 6757
Email	b.molesworth@unsw.edu.au

Support Services Contact Details

If at any stage during the project you become distressed or require additional support from someone not involved in the research, please call:

Name/Organisation	UNSW Counselling and Psychological Services
Telephone	+ 61 2 9385 5418
Email	counselling@unsw.edu.au

What if I have a complaint or any concerns about the research study?

If you have any complaints about any aspect of the project, the way it is being conducted, then you may contact:

Complaints Contact

Position	Human Research Ethics Coordinator
Telephone	+ 61 2 9385 6222
Email	humanethics@unsw.edu.au
HC Reference Number	TBA

Consent Form – Participant providing own consent

Declaration by the participant

- ☐ I understand I am being asked to provide consent to participate in this research project;
- ☐ I have read the Participant Information Sheet, or someone has read it to me in a language that I understand;
- ☐ I understand the purposes, study tasks and risks of the research described in the project;
- ☐ I understand that the research team will collect personal information about me which will be de-identified for statistical purposes including age, sex, flying rating, endorsements and hours; I provide my consent for this to happen.
- ☐ I provide my consent for the information collected about me to be used for the purpose of this research study only.
- ☐ I have had an opportunity to ask questions and I am satisfied with the answers I have received;
- ☐ I freely agree to participate in this research study as described and understand that I am free to withdraw at any time during the project and withdrawal will not affect my relationship with any of the named organisations and/or research team members;
- ☐ I would like to receive a copy of the study results via email or post, I have provided my details below and ask that they be used for this purpose only;

Name: ______________________________________

Address: ____________________________________

Email Address: _______________________________

I understand that I will be given a signed copy of this document to keep;

Participant Signature

Name of Participant (please print)	
Signature of Research Participant	
Date	

Declaration by Researcher*

I have given a verbal explanation of the research study, its study activities and risks, and I believe that the participant has understood that explanation.

Researcher Signature*

Name of Researcher (please print)	
Signature of Researcher	
Date	

***An appropriately qualified member of the research team must provide the explanation of, and information concerning the research study.**

Note: All parties signing the consent section must date their own signature.

Form for Withdrawal of Participation

I wish to **WITHDRAW** my consent to participate in the research proposal described above and understand that such withdrawal **WILL NOT** affect my relationship with The University of New South Wales, or the UNSW School of Aviation. In withdrawing my consent, I would like any information which I have provided for the purpose of this research project withdrawn.

Participant Signature

Name of Participant (Please print)	
Signature of Research Participant	
Date	

The section for Withdrawal of Participation should be forwarded to:

CI Name:	Brett Molesworth
Email:	b.molesworth@unsw.edu.au
Phone:	02 9385 6757
Postal Address:	School of Aviation The University of New South Wales Sydney, NSW Australia 2052

APPENDIX 4: Experiment 3 Participant Information Sheet & Consent Form

School of Aviation	UNSW AUSTRALIA

PARTICIPANT INFORMATION STATEMENT AND CONSENT FORM

Measured effects of workload and dynamic auditory feedback on remote pilot perception and decision-making

A/Prof. Brett Molesworth (Chief Investigator) & Mr. Matthew Dunn (Student Investigator)

HC Number: 210309

What is the research study about?

You are invited to take part in this research study. The aim of the proposed research is to investigate the effects of dynamic auditory feedback under changing workload levels on task performance by the remotely piloted aircraft (RPA) operator.

Who is conducting this research?

The study is being carried out by Mr. Matthew Dunn (Student Investigator - SI) under the supervision of A/Prof. Brett Molesworth (Chief Investigator) from the School of Aviation at UNSW Sydney.

Inclusion/Exclusion Criteria

There is no inclusion/exclusion criterion to participate in this research.

Do I have to take part in this research study?

Participation in any research project is voluntary. If you do not want to take part, you do not have to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage. Your decision to withdraw from the study will have no impact on your studies at UNSW Sydney.

If you decide you want to take part in the research study, you will be asked to:

- Read the information carefully (ask questions if necessary; take the form with you);
- Sign the consent form;
- Complete a series of short questionnaires; and
- Complete a series of remotely piloted aircraft (RPA) flight monitoring and decision-making tasks.

What does participation in this research require, and are there any risks involved?

If you decide to take part in the research study, we will ask you to complete the following tasks:

- Complete one Demographics Questionnaire (5 minutes).
- Complete thirteen Workload Questionnaires (1 minute per questionnaire; 13 minutes total).
- Complete a familiarisation task (3 minutes)
- Complete twelve main experiment tasks (5 minutes each – total time 60 minutes).

With added breaks as required and to transition between tasks, the research will take no longer than 1.5 hours to complete.

Completion of Questionnaires

The research team will ask you to complete a series of short questionnaires. The first questionnaire relates to demographic information such as age and gender. The second questionnaire will be completed initially for baseline information and then at the completion of each flying task and relates to your perceived workload of the task.

Auditory Feedback

During the research, you will be exposed to auditory feedback from the RPA propellers. You will be asked to set the volume level of the feedback to a level that is comfortable. As part of the research, we will increase this level by 10 dBA for some tasks. Noise levels above 85 dBA for 8 hours or more are known to cause hearing damage. Based on our previous research, the level of the auditory feedback is typically around 60 dBA. Hence, it is unlikely that the noise level with reach 85 dBA and as a result no damage is expected (plus the study lasts no longer than 1.5 hrs). Nonetheless, we will ensure it does not exceed 85 dBA and as part of this, we will show you the sound level meter reading to put your mind at ease. If at any point you are uncomfortable with the sound level, please notify the student investigator, and the research will be ceased.

What are the possible benefits to participation?

We hope to use the information we get from this research study to benefit others who will be operating RPAS in the future. The technology is rapidly evolving, and we are using currently available technology to determine the most appropriate methods to control RPAS in potential situations such as medical emergencies, search and rescue, humanitarian aid and post-disaster events. By understanding the limitations of human performance, new and safer human-centred autonomous systems can be developed.

What will happen to information about me?

By signing the consent form, you consent to the research team collecting and using information about you for the research study. We will keep your data for 7 years after the completion of the study in accordance with the National Ethics Guidelines. We will store information about you in a non-identifiable format securely within the UNSW School of Aviation Kensington campus. Your information will only be used for statistical analysis and will be disclosed in academic publications and presentations in the future.

How and when will I find out what the results of the research study are?

The research team intend to publish and report the results of the research study in a variety of ways. All information published will be done in a way that will not identify you. If you would like to receive a copy of the results you can let the research team know by adding your email or postal address within the consent form. We will only use these details to send you the results of the research.

What if I want to withdraw from the research study?

If you do consent to participate, you may withdraw at any time and all data collected from you will be destroyed. You can do so by completing the 'Withdrawal of Consent Form' which is provided at the end of this document. Alternatively, you can ring the research team and tell them you no longer want to participate. Your decision not to participate will not affect your relationship with UNSW Sydney or the UNSW School of Aviation.

What should I do if I have further questions about my involvement in the research study?

The person you may need to contact will depend on the nature of your query. If you would like any further information concerning this project or if you have any problems which may be related to your involvement in the project, you can contact the following member/s of the research team:

Research Team Contact Details

Name	Associate Professor Brett Molesworth
Position	Chief Investigator
Telephone	+ 61 2 9385 6757
Email	b.molesworth@unsw.edu.au

Support Services Contact Details

If at any stage during the project you become distressed or require additional support from someone not involved in the research, please call:

Name/Organisation	UNSW Counselling and Psychological Services
Telephone	+ 61 2 9385 5418
Email	counselling@unsw.edu.au

What if I have a complaint or any concerns about the research study?

If you have any complaints about any aspect of the project, the way it is being conducted, then you may contact:

Complaints Contact

Position	Human Research Ethics Coordinator
Telephone	+ 61 2 9385 6222
Email	humanethics@unsw.edu.au
HC Reference Number	TBA

Consent Form – Participant providing own consent

Declaration by the participant

- ☐ I understand I am being asked to provide consent to participate in this research project;
- ☐ I have read the Participant Information Sheet, or someone has read it to me in a language that I understand;
- ☐ I understand the purposes, study tasks and risks of the research described in the project;
- ☐ I understand that the research team will collect personal information about me which will be de-identified for statistical purposes including age and sex; I provide my consent for this to happen;
- ☐ I provide my consent for the information collected about me to be used for the purpose of this research study only;
- ☐ I have had an opportunity to ask questions and I am satisfied with the answers I have received;
- ☐ I freely agree to participate in this research study as described and understand that I am free to withdraw at any time during the project and withdrawal will not affect my relationship with any of the named organisations and/or research team members;
- ☐ I would like to receive a copy of the study results via email or post, I have provided my details below and ask that they be used for this purpose only;

Name: ____________________________________

Address: __________________________________

Email Address: _____________________________

I understand that I will be given a signed copy of this document to keep.

Participant Signature

Name of Participant (please print)	
Signature of Research Participant	
Date	

Declaration by Researcher*

I have given a verbal explanation of the research study, its study activities and risks, and I believe that the participant has understood that explanation.

Researcher Signature*

Name of Researcher (please print)	
Signature of Researcher	
Date	

***An appropriately qualified member of the research team must provide the explanation of, and information concerning the research study.**

Note: All parties signing the consent section must date their own signature.

Form for Withdrawal of Participation

I wish to **WITHDRAW** my consent to participate in the research proposal described above and understand that such withdrawal **WILL NOT** affect my relationship with The University of New South Wales, or the UNSW School of Aviation. In withdrawing my consent, I would like any information which I have provided for the purpose of this research project withdrawn.

Participant Signature

Name of Participant (Please print)	
Signature of Research Participant	
Date	

The section for Withdrawal of Participation should be forwarded to:

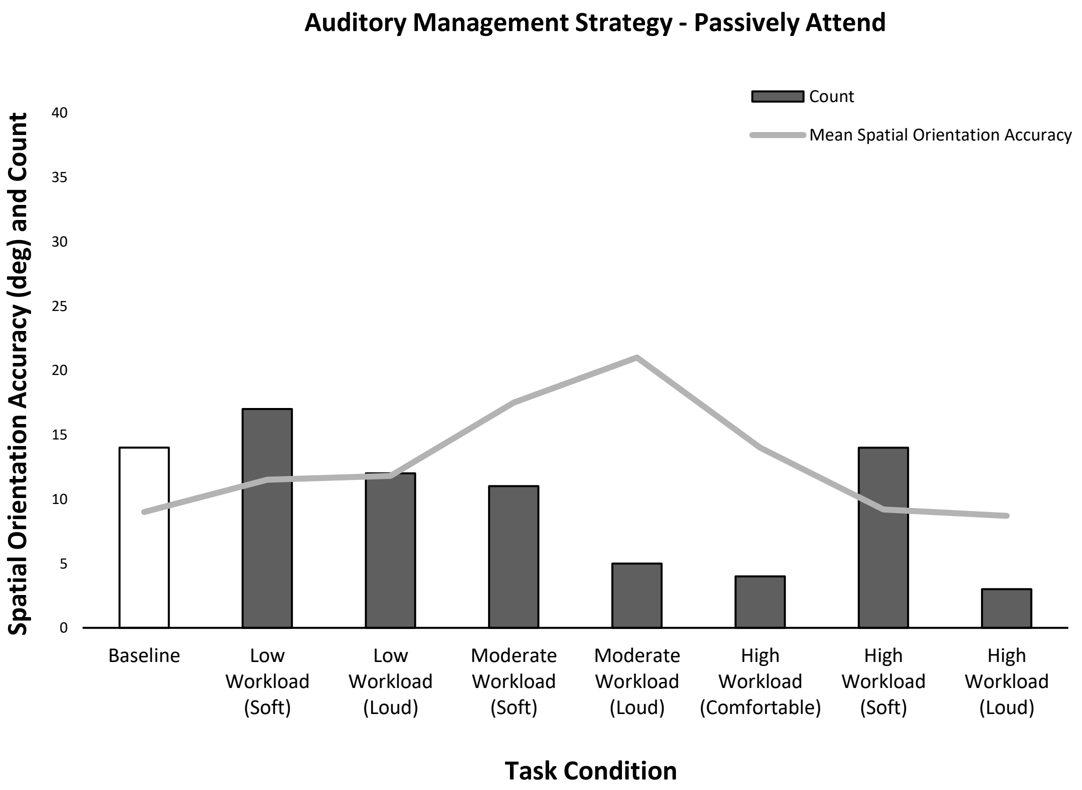
CI Name:	Brett Molesworth
Email:	b.molesworth@unsw.edu.au
Phone:	02 9385 6757
Postal Address:	School of Aviation The University of New South Wales Sydney, NSW Australia 2052

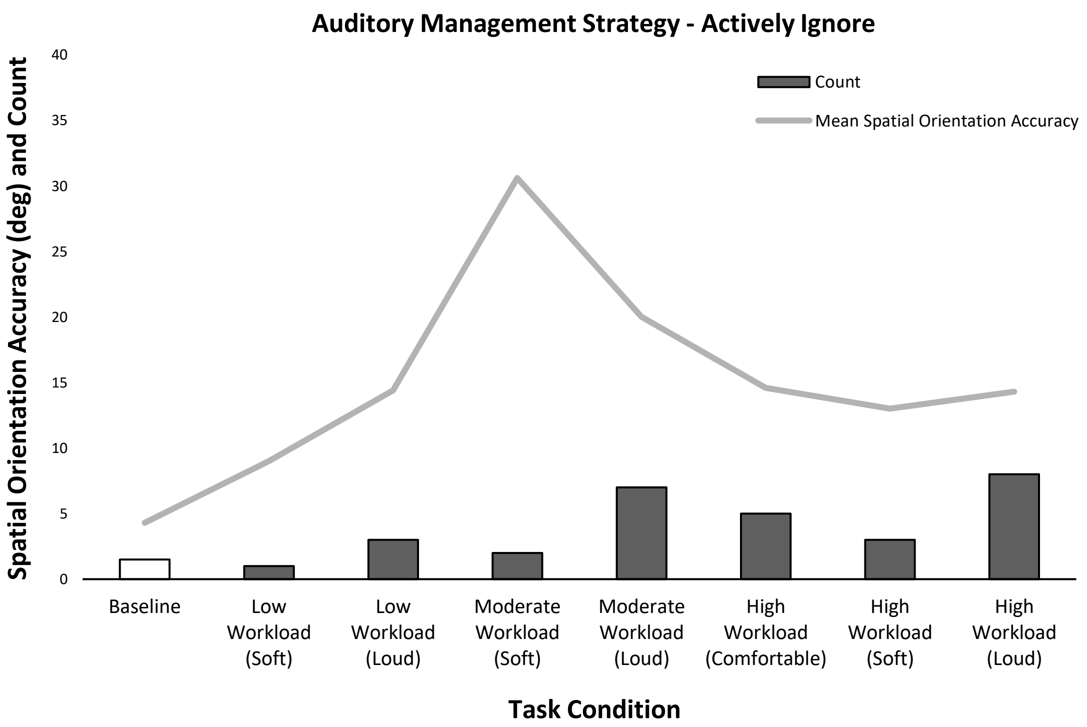
APPENDIX 5: Volume Reference Table (Experiment 3)

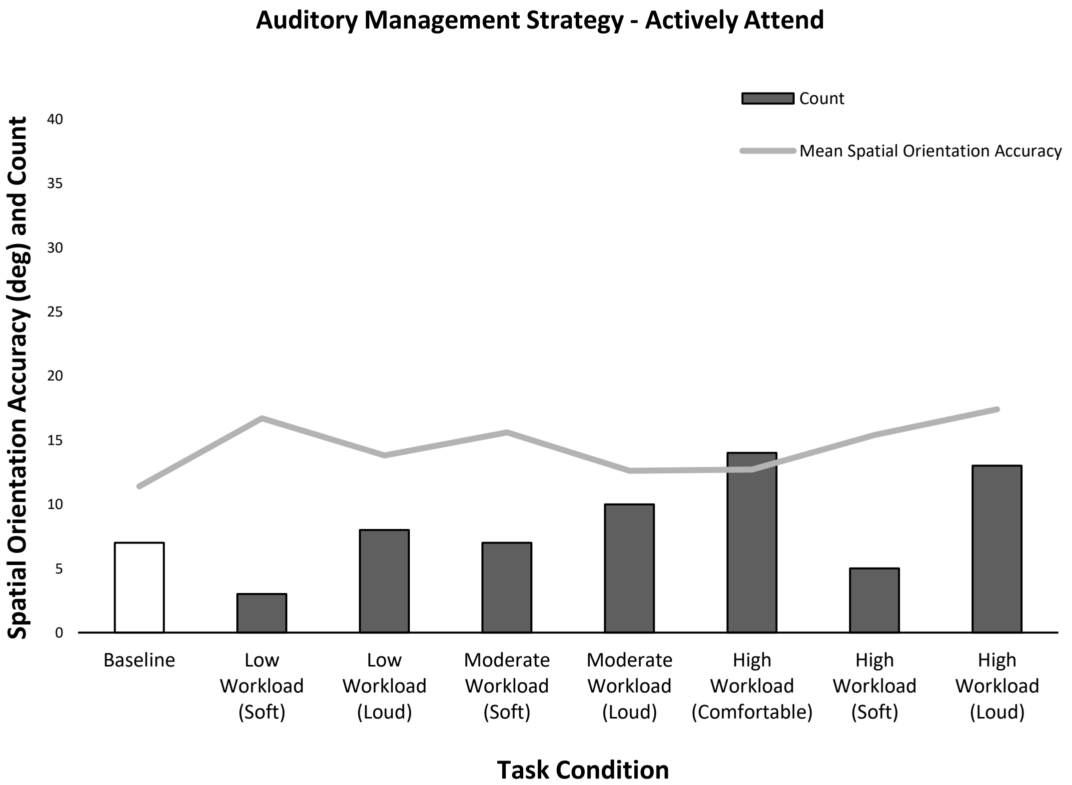
Volume % Level	dB(a)	+10 dB(a) Vol %	-10 dB(a) Vol %
10	42.1		
11	42.1		
12	42.3		
13	42.3		
14	42.9		
15	43.6		
16	44.3		
17	44.9		
18	45.8		
19	46.9		
20	48.6		
21	49.8		
22	50.9	32	10
23	51.9	34	11
24	53.9	36	15
25	54.9	38	17
26	55.8	39	18
27	56.7	40	19
28	57.5	41	19
29	58.2	42	20
30	59.1	44	21
31	59.9	45	21
32	60.7	47	22
33	61.5	48	23
34	62.3	50	23
35	63.1	51	24
36	63.8	53	24
37	64.4	54	25
38	**65.1**	**55**	**25**
39	66.3	57	27
40	67.0	59	27
41	67.6	61	28
42	68.3	62	29
43	68.7	63	29
44	69.3	65	30
45	69.8	66	31
46	70.3	68	31
47	70.9	69	32
48	71.4	70	33
49	71.9	71	33
50	72.4	73	34
51	72.9	74	35
52	73.4	75	35
53	73.9	77	36
54	74.7	80	37
55	75.2	82	38
56	75.6	83	38
57	76.1	85	39
58	76.5	87	39
59	76.9	89	40
60	77.3	90	41
61	77.8	92	41
62	78.1	94	42
63	78.5	96	42
64	78.9	97	43
65	79.3	99	44
66	79.7	101	45
67	80.1	104	46
68	80.4	105	46
69	81.1	110	47
70	81.5		

*Using a Brüel & Kjær Integrating-averaging Sound Level Meter Type 2240 (calibrated 31 May 2021, +/- 0.2 dB). Using the LAeq i.e. the equivalent continuous ‘A’ frequency-weighted sound pressure level (linear average of all varying sound pressure level detected during a measurement; not time-weighted).
**Audio balance on computer: L = 75%, R=100% to get equal dB(a) reading (91.0 dB(a) at 100% volume value for R side).
*** Values obtained using flight 1, average reading from mm:ss = 02:00 to 02:30.

APPENDIX 6: Large-Sized Figures Comprising Figure 16







Auditory Management Strategy - Actively Attend
Count
Mean Spatial Orientation Accuracy
Spatial Orientation Accuracy (deg) and Count
40
35
30
25
20
15
10
5
0
Baseline
Low Workload (Soft)
Low Workload (Loud)
Moderate Workload (Soft)
Moderate Workload (Loud)
High Workload (Comfortable)
High Workload (Soft)
High Workload (Loud)
Task Condition

Auditory Management Strategy - Did Not Notice
Count
Mean Spatial Orientation Accuracy
Spatial Orientation Accuracy (deg) and Count
40
35
30
25
20
15
10
5
0
Baseline
Low Workload (Soft)
Low Workload (Loud)
Moderate Workload (Soft)
Moderate Workload (Loud)
High Workload (Comfortable)
High Workload (Soft)
High Workload (Loud)
Task Condition