# An investigation of the role of spatial ability in problem solving in engineering education

Doctoral Thesis

Submitted by

**Clodagh Reid**

to

Faculty of Engineering and Informatics

Technological University of the Shannon: Midlands Midwest,

Athlone Campus, Ireland

Completed under the supervision of:

Dr Rónán Dunbar, Technological University of the Shannon: Midlands Midwest**,** Athlone, Ireland

Dr Jeffrey Buckley, Technological University of the Shannon: Midlands Midwest**,** Athlone, Ireland & KTH Royal Institute of Technology, Stockholm, Sweden

Dr Niall Seery, Technological University of the Shannon: Midlands Midwest**,** Athlone, Ireland

Submitted to Technological University of the Shannon: Midlands Midwest.

November 2021

# ABSTRACT

Problem solving is a fundamental component of engineering practice and as such, the development of problem-solving skills is a core goal of engineering education. Problem-orientated pedagogical approaches are used in educational practice to facilitate the development of this key skill, however, there are reports that the problem-solving skills of engineering graduates are underdeveloped. It is necessary to investigate and understand factors which support problem solving to contribute towards addressing this skills gap. General cognitive abilities are outlined as an important factor underpinning problem solving. Of particular interest in the context of engineering problem solving is the general cognitive ability, spatial ability. Engineering is a spatially and visually oriented discipline where spatial ability has been associated with retention and success of individuals in the field. Although there is a body of correlational evidence to indicate the contribution of spatial ability to success in STEM disciplines, there is a significant gap in understanding the causal relationship between spatial ability and associated success.

This thesis aims to contribute towards the understanding of the causal association between spatial ability and success in engineering through investigating the role of spatial ability in problem solving. As spatial ability is a cognitive factor, cognitive load is considered as a pertinent factor in this research as it is theorised that spatial ability may influence the cognitive load experience during problem solving. A convergent mixed method study was conducted to address the research aim with engineering students at the initial and concluding stages of their engineering education to attend to the research question and objectives.

The results of this thesis establish that complex problem-solving performance did not differ across levels of engineering expertise (as determined through progression), showing a tentative misalignment between educational goals and outcomes. The findings indicated that spatial ability levels did not differ across groups of expertise. However, findings suggested that higher levels of spatial ability contribute to (a) problem solving performance (b) cognitive load during problem solving and (c) behaviours demonstrated during problem solving. This research contributes towards understanding of the causal association between spatial ability and success in engineering by demonstrating that irrespective of levels of engineering expertise, spatial ability can influence problem-solving performance which is an inherent component of contemporary engineering education approaches. The findings of this thesis are discussed relative to their potential implications for engineering education practice and advancement of understanding the causal theory of spatial ability and success in engineering.

*Key Words: Spatial ability, Problem solving, Engineering education, Cognitive load*

## RELATED WORK

The following peer-reviewed publications have resulted from this work to date:

Reid, C., Buckley, J. & Dunbar, R. (Under Review). An approach to support systematic inquiry into cognitive abilities used when problem solving in technological education.

Reid, C., Dunbar, R., Seery, N. & Buckley, J. (Under Review). Evaluating the Capacity of a Convergent Mixed Method to Gain Insight into the Effect of Cognitive Load and Behaviors on Problem-Solving Performance.

Reid, C., Buckley, J., & Dunbar, R. (2021). Investigating the effect of engineering student's spatial ability and expertise on general complex problem solving. *Techne serien- Forskning i slöjdpedagogik och slöjdvetenskap*, 28(2), 72-81.

Reid, C., Keighrey, C., Murray, N., Dunbar, R., & Buckley, J. (2020). A novel mixed methods approach to synthesize EDA data with behavioral data to gain educational insight. *Sensors*, 20(23), 6857.

Reid, C., Keighrey, C., Dunbar, R., Murray, N., & Buckley, J. (2019). Developing a methodological approach to measure cognitive load during complex problem solving. In *PATT37 Conference 2019: Developing a knowledge economy through technology and engineering education*.

Reid, C., Dunbar, R., & Buckley, J. (2018). A preliminary model of problem categorisation to explore the cognitive abilities required for problem solving in engineering education. In *PATT36 International Conference: Research and Practice in Technology Education: Perspectives on Human Capacity and Development. 36th International Pupils' Attitudes Towards Technology Conference*. TERG.

Reid, C., & Dunbar, R. (2018). Spatial Ability in Engineering Education: Working Towards an Integrated Learning Approach. In *46th SEFI Annual Conference: Creativity, Innovation and Entrepreneurship for Engineering Education Excellence*.

## DECLARATION

I hereby declare that the work contained within this thesis, submitted to Technological University of the Shannon: Midland Midwest for the degree of Doctorate of Philosophy, has not been accepted for the award of any other degree, in any other higher education institution, and is entirely my own work and to the best of my knowledge contains no work previously written or published by another party, except in the case of reference material.

Clodagh Reid

19 / 11 / 2021

Clodagh Reid

Date

## ACKNOWLEDGEMENTS

I would like to express my sincere thanks to all the people that have supported me in completing this thesis.

First of all, I would like to thank my supervisors, Dr Rónán Dunbar, Dr Jeffrey Buckley, and Dr Niall Seery for their knowledge and guidance throughout this process. To my primary supervisor Rónán who has been there since the beginning, thank you for all the time, mentorship, help, and encouragement to pursue opportunities throughout the last four years. Jeff, as my co-supervisor, you have always been there to offer advice, to encourage me, and to support me in challenging my own perspectives for the benefit of the work. Niall, as my newest co-supervisor, thank you for your advice, perspectives, and wealth of experience that you have brought to this process. Your contribution has been invaluable to this work.

I would like to thank my friends and colleagues in the research hub, you have always been there for me whether it was for a chat, call or coffee to unwind.

I would like to thank the TERG members for all their support at conferences and for helping me to expand my research understandings. You have all contributed to my growth as a researcher.

Thank you to the technology and engineering education research communities for being so welcoming and providing a friendly space to discuss my research.

I would like to thank all of the students that volunteered to participate and staff in the engineering faculty. Without your involvement I would not have been able to complete this work.

I would like to thank Dr Dawne Irving-Bell and Professor Tom Collins for acting as examiners for my transfer process. You challenged me to consider different aspects of my work which has helped me in clarifying the perspectives I have taken.

I would like to thank my boyfriend, Damien, for everything he has done to help me in reaching this point. Your continuous support and encouragement to pursue my goals has made all of this possible.

Finally, I would like to thank my brother Shaun, sisters Heather and Aisling, my parents Aidan and Alice, and grandmothers Alice and Dympna, for always being there for me in everything that I do. Your unwavering support has made me who I am today, and I know that this wouldn’t have been possible without you all.

**TABLE OF CONTENTS**

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF APPENDICES

## NOMENCLATURE

CDIO – Conceive, Design, Implement, Operate

CHC – Cattell Horn and Carroll

CLT – Cognitive Load Theory

EDA – Electrodermal activity

IR 4.0 – Fourth industrial revolution

MRT – Mental Rotation Test

PBL – Problem based learning

PFT – Paper Folding Test

PSVT:R – Purdue Spatial Visualisation Test

SDT – Surface Development Test

STEM – Science Technology Engineering Mathematics

TOH – Tower of Hanoi

VAS – Visual Analogue Scale

# 1. Introduction

## 1.1. Context

### *1.1.1. Contemporary engineering education*

The desired attributes of an engineer are ever evolving in an era of rapidly advancing societal and technical needs (Engineers Ireland, 2021; SEFI, 2016; UNESCO, 2010). In line with these advancements in the engineering profession, the provision of engineering education has been required to move beyond the development of discipline-specific knowledge and skills towards equipping engineering graduates with a broader professional skillset to face the challenges of engineering today and address industry needs (Caeiro-Rodriguez et al., 2021; Litzinger et al., 2011; SEFI, 2016; UNESCO, 2010). Discipline knowledge and skills, also referred to as technical knowledge and skills, incorporate factors such as science and engineering fundamentals, applications, and practice (Caicedo et al., 2014; Crawley et al., 2014; Lantada et al., 2014; Passow & Passow, 2017; Saleh & Lamsali, 2020; Shuman et al., 2005; UNESCO, 2010). Professional skills are also commonly referred to as “transversal”, “generic”, “soft”, “non-technical” or “domain-general” skills in the literature (Balaji & Somashekar, 2009; Caeiro-Rodriguez et al., 2021; Caicedo et al., 2014; Colman & Willmot, 2016; Flening et al., 2021; Kamaruzaman et al., 2019; Lantada et al., 2014; Leandro Cruz et al., 2020; Saleh & Lamsali, 2020; Tricot & Sweller, 2014). These skills will be referred to as both *professional skills* and *domain-general skills* throughout this thesis as the term professional skills is common to engineering education while the term domain-general skills is used in human intelligence literature (Anwar & Richards, 2018; Caicedo et al., 2014; Flening et al., 2021; Hambrick & Meinz, 2011; Oswald et al., 2014; Shuman et al., 2005; Tricot & Sweller, 2014). These are skills that do not relate directly to a specific discipline but can be applied to a broad range of problems and situations in diverse settings (Tricot & Sweller, 2014). In the context of engineering professional skills include communication, critical-thinking, ethics, teamwork, and problem solving (Balaji & Somashekar, 2009; Caeiro-Rodriguez et al., 2021; Kamaruzaman et al., 2019; Nair et al., 2009; Passow & Passow, 2017; Saleh & Lamsali, 2020; SEFI, 2016). Contemporary engineering education programmes aim to balance the development of both discipline-specific competences (engineering knowledge and skill which manifests as engineering expertise) and domain general competences (which manifest as professional engineering skills). This core component of engineering education is represented in Figure 1. The significance of the professional skill of complex problem solving and cognitive capacity of spatial ability will be unpacked in sections 1.1.2 through 1.1.4.

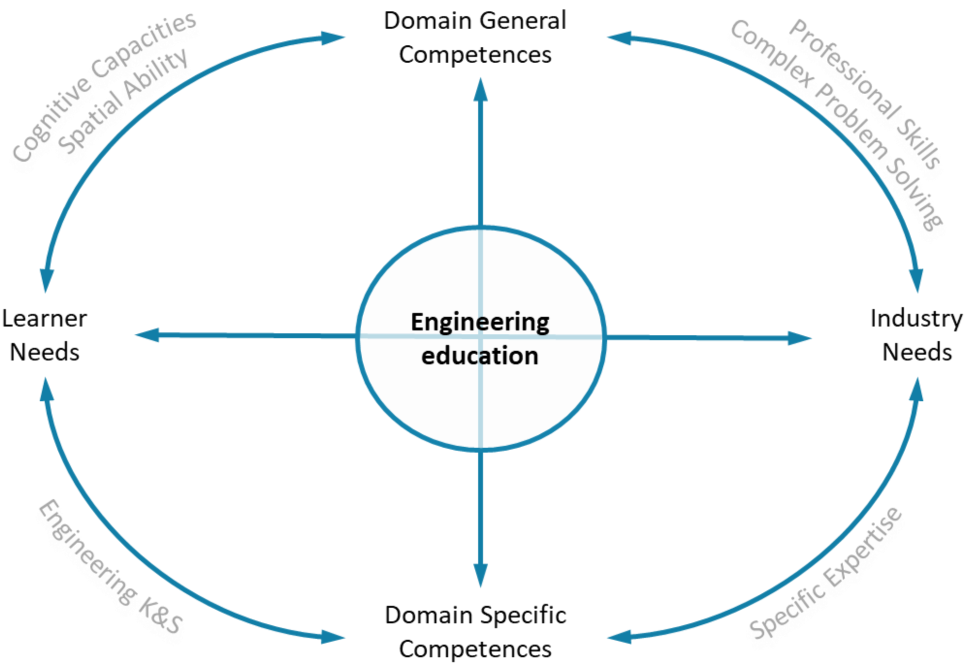


*Figure 1. Core components of contemporary engineering education programmes.*

In balancing the development of competences, the goal is to address learner needs and equip students with the capacity to readily adapt to solve the novel and complex problems that they will experience throughout their engineering careers (Caeiro-Rodriguez et al., 2021; Litzinger et al., 2011; SEFI, 2016). Therefore, there has been a significant drive over the last number of years to establish alternatives to traditional higher engineering education approaches and to adapt curricula to facilitate the development of a modern engineering skillset which strikes the balance between industry and learner needs (Edström & Kolmos, 2014; Faber & Benson, 2017), demonstrated by the learner needs-industry needs dichotomy in . This has seen a rise in the use of problem-orientated engineering education approaches such as problem/project-based learning (PBL) and conceive-design-implement-operate (CDIO) (Crawley et al., 2014; Edström & Kolmos, 2014; Hanney & Savin-Baden, 2013; Savin-baden, 2014). The premise of these educational approaches is for students to be provided with effective learning experiences whereby students develop a deep understanding around key concepts and principles, develop both discipline knowledge and skills and professional skills, and apply their knowledge and skills to problems representative of real-world engineering problems (Crawley et al., 2014; Edström & Kolmos, 2014; Litzinger et al., 2011).

### *1.1.2. Problem solving in engineering education*

Problem solving is argued to be one of the most fundamental requirements of a professional engineer. Jonassen (2015, p. 103) details that “learning to solve workplace problems is an essential learning outcome for any engineering graduate. Every engineer is hired, retained, and rewarded for his or her ability to solve problems.” Sheppard et al. (2008, p. 3) outline the importance of problem solving in engineering noting that “engineering practice is, in its essence, problem solving”. Passow and Passow (2017) describe problem solving as “the core of engineering practice” and from 60 samples consisting of 14,429 respondents (practicing engineers, undergraduate alumni, and engineering faculty), they identified problem solving in the top level of importance of competencies that should be emphasised in undergraduate engineering programmes.

In another systematic review of engineering skills, Kamaruzaman et al., (2019) evidence problem-solving skills as the fourth most preferred non-technical engineering skill required by accrediting bodies for engineering in 18 countries under the Washington Accord. More specifically they identified *complex problem-solving skills* as the most important skills need in relation to the contemporary requirements of Industry 4.0. Complex problem solving is broadly described as “reviewing related information to develop and evaluate options and implement solutions” (World Economic Forum, 2020, p. 153).

However, despite the emphasis being placed on the development of problem-solving skills through engineering curricula, recent research examining the employability of engineering graduates has highlighted that problem solving is an area that they are underperforming in, according to their employers (Nair et al., 2009; Valentine et al., 2017; Ward & Thiriet, 2010). Given the prevalence of problem solving in engineering education, and the importance placed on problem solving from both national and international organisations, it is necessary that efforts are made towards understanding the components underpinning problem-solving skills to address the documented skills gap and in turn, optimise engineering education provision.

### *1.1.3. Problem solving as a cognitive process*

Problem solving is described as one of the fundamental human cognitive processes and detailed as a higher-order cognitive process (Y. Wang & Chiew, 2010). From a cognitive perspective, when a problem is identified, problem solving can be considered as a search process in an individual's memory to find a relationship between goals to reach a solution and a set of alternative paths (Carlson & Bloom, 2005; Mayer & Wittrock, 2006; Y. Wang, 2007; Y. Wang & Chiew, 2010). The characteristics of a successful problem solver include: a capacity to correctly identify the goals of a problem, persistence, ability to adopt search strategies that are efficient, regulate one's actions, and a capacity to trace back to previous points in the solution process (Litzinger et al., 2010; Y. Wang & Chiew, 2010). Complex problem solving may require an individual to employ their discipline knowledge, cognitive strategies (e.g., heuristics), and cognitive abilities (Hambrick et al., 2012; Litzinger et al., 2010). Depending on the type of problem the individual is solving, they may require a combination of each of these components. Cognitive strategies are used to categorise problems, focus attention, integrate information, and facilitate knowledge transfer (Litzinger et al., 2010; Schraw et al., 2006). When engaging with a novel problem or problem situation, where an individual has not acquired *expertise* relative to the problem and has not acquired problem-solving strategies, individuals may rely on cognitive abilities to circumvent limits in their knowledge to process information and successfully solve the problem (Hambrick et al., 2012, 2016; Hambrick & Meinz, 2011). Expertise in an area is defined as "the possession of a large body of knowledge and procedural skills" (Chi et al., 1981, p. 2). Throughout this thesis, the term 'expertise' will be used to refer to different levels of educational experience. Through the circumvention-of-limits hypothesis, it is proposed that individuals with low levels of expertise but high levels of pertinent cognitive abilities to solve a problem can perform to a similar standard as those with high levels of domain knowledge (Hambrick et al., 2012). Of particular interest in the context of engineering is the broad cognitive ability of visual processing, commonly referred to as spatial ability in the literature (Buckley et al., 2018a).

### *1.1.4. Spatial ability as a factor of problem-solving performance*

Within the Cattell-Horn-Carroll (CHC) theory of intelligence, *spatial ability* is described as “the ability to make use of simulated mental imagery to solve problems- perceiving, discriminating, manipulating, and recalling non-linguistic images in the “mind’s eye”” (Schneider & McGrew, 2018, p. 125). There is a significant body of empirical evidence indicating that spatial ability is a predictor of student success in Science, Technology, Engineering, and Maths (STEM) education (Lubinski, 2010; Stieff & Uttal, 2015; Uttal & Cohen, 2012; Wai et al., 2009). It is evidenced to contribute to the retention of students in STEM education, while also attributed to the underrepresentation of females, individuals from lower socioeconomic backgrounds, and minority groups in these disciplines (Ball et al., 2019; Blums et al., 2017; Sorby et al., 2013; Sorby & Veurink, 2019; M.-T. Wang & Degol, 2017). Spatial ability is described as an enhancer to learning with individuals with higher spatial ability having more cognitive resources available to build referential connections between verbal and visual representations of information (Mayer & Sims, 1994). As such, the malleability of spatial ability has been explored through the literature, with a body of evidence indicating that it can be developed through intervention and contribute to increased performance in STEM subjects (Martin-Dorta et al., 2009; Martín-Gutiérrez et al., 2015; Martín-Gutiérrez & González, 2017; Metz et al., 2016; N. Newcombe, 2017; Roca-González et al., 2017; Sorby, 1999, 2007, 2009; Sorby & Baartmans, 1996, 2000; Uttal, Meadow, et al., 2013; Uttal & Cohen, 2012). Given the described role of spatial ability in problem solving, its relevance to success in STEM education, and malleability, it is a necessary factor to be considered when seeking to gain understanding of the elements that contribute to problem-solving performance. Furthermore, in viewing spatial ability as a cognitive factor which may support complex problem solving, which is detailed to make more cognitive resources available to learners, it is pertinent to consider the potential of spatial ability to influence the cognitive load experienced during problem solving.

### *1.1.5. Cognitive load in problem solving*

*Cognitive load* is the total working memory resources necessary to carry out a learning task (P. A. Kirschner et al., 2018). In education, the goal is to optimise intrinsic cognitive load which is the load “imposed by the basic structure of the information that the learner needs to acquire for achieving learning goals irrespective of the instructional procedures used” (Sweller et al., 2011, p. 57). When this load is not optimised, where there is too great of a demand being placed on working memory capacity, it can hinder an individual’s capacity to learn or successfully perform the task (O. Chen & Kalyuga, 2020; Paas et al., 2004; Paas, Renkl, et al., 2003; Sweller et al., 2019). This is referred to as extraneous cognitive load, which is load that is “unnecessary and extraneous to the learning goals” (Sweller et al., 2011, p. 57). This load can influence the learner’s capacity to transfer knowledge relative to the experience to long-term memory and affect their willingness to engage with similar tasks in the future (Paas et al., 2005; Sweller et al., 2011, 2019). Together intrinsic and extraneous cognitive load determine the total cognitive load experienced which in turn, determines the working memory resources required to process the information (P. A. Kirschner et al., 2018; Sweller et al., 2011). Understanding the implications of cognitive load on learning, thinking, and problem solving “can provide us with a coherent, unifying base that can be used to generate instructional hypotheses and data” (Sweller et al., 2011, p. v) which may then be applied through educational approaches.

Where cognitive load is experienced during problem solving, it is hypothesised that spatial ability may support individuals in overcoming limitations in their processing capacity to successfully navigate the problem. This hypothesis forms the basis for the research aim outlined in the following section.

### 1.2. Overview of research context

The following points are a synopsis of the critical perspectives within the context of this research:

- The development of problem-solving skills is prioritised due to the key role that it plays within engineering education and practice.
- Employers claim that problem solving is an area where engineering graduates are underperforming.
- Discipline knowledge, cognitive strategies and general cognitive abilities are outlined as important factors which support problem solving.

- Spatial ability is a critical general cognitive ability to be considered given (1) the spatially and visually orientated nature of engineering and (2) the evidenced role of spatial ability in the retention and success of individuals in the discipline.

Although a body of correlational evidence indicates that spatial ability contributes to success in STEM disciplines, there is no clear body of empirical evidence documenting the causal relationship between spatial ability and this associated success.

### 1.3. Research aim

This research aimed to investigate the role of spatial ability in problem solving to contribute towards the understanding of the causal association between spatial ability and success in engineering education.

### 1.4. Research question

To contribute to addressing this aim, the following research question is attended to throughout this thesis:

> How does spatial ability's role in problem solving contribute to a broader causal theory between spatial ability and success in engineering education?

### 1.5. Research objectives

To respond to the research question the following research objectives were formulated:

- To examine performance on a domain general complex problem across levels of engineering student disciplined expertise.
- To investigate if levels of spatial ability vary between engineering students with different levels of disciplined expertise.
- To examine if spatial ability influences engineering students' performance in general complex problem solving.
- To investigate whether levels of spatial ability affect the cognitive load experienced during problem solving.

### 1.6. Thesis outline

Following the presentation of the underpinning context, aim, question and objectives of this thesis, the literature review section sets out the significant bodies of research which form the theoretical foundations of this work. Figure 2 provides an overview of the theoretical connections between the variables explored through this thesis. Discipline problem-solving capability is underpinned by both expertise and cognitive abilities, where cognitive abilities can circumvent limits in expertise to support successful problem solving. Acquiring expertise requires engagement in learning experiences that optimise intrinsic cognitive load which is theorised to be supported by an individual's level of cognitive ability, specifically spatial ability.

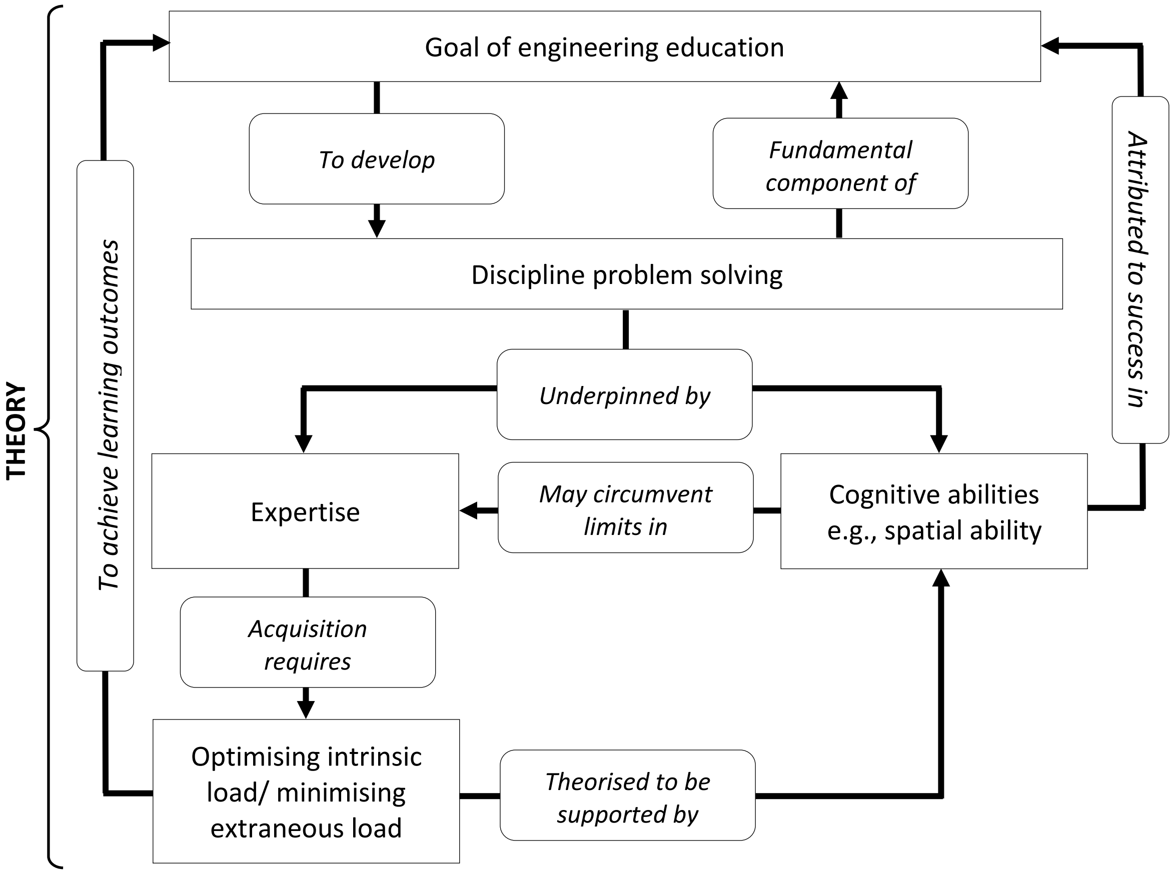


*Figure 2. The main theories underpinning the research conducted in this thesis.*

Through the methodology section, the research contributing to the methodological framework of this thesis is presented and an overview of the approach taken to addressing the research objectives is set out. The findings section sets out the analytic approach employed to evaluate the collected data and to facilitate answering of the research questions. The discussion section synthesises the findings from the research and details the potential implications for engineering education research and practice.

The conclusions of the research are then described with specific focus on the achievement of the research aim and objectives. Finally, some limitations of this research are described, and recommendations provided for continuation of future work in this space.

# 2. Literature review

## 2.1. Overview of literature review structure

As the work presented in this thesis is bound in the context of engineering education, there is a need to provide an overview of the evolution of engineering education and the core agendas of contemporary educational programmes. Problem solving is a central component of engineering practice, making the development of problem-solving skills a fundamental element of engineering education. Throughout this section problem solving in the context of engineering will be unpacked by exploring the various structures of problems, and the underpinning cognitive process of problem solving. Viewing problem solving through a cognitive lens and as a cognitive process requires consideration of cognitive factors that it may interact with. The structure and contemporary understandings of human cognitive ability, along with the contribution of factors of intelligence to problem-solving performance in engineering, will be examined. Spatial ability is a pertinent cognitive ability to be considered in the context of this thesis due to its documented role in success in engineering education and problem solving. An exploration of the position of spatial ability within the empirical framework of human intelligence, its significance to engineering, and approaches to measure and develop this cognitive ability are presented. Considering the interactions of cognitive processes towards problem solving, it is pertinent to reflect upon the management of learner's cognitive resources during this process. Therefore, a broad description of the influence of cognitive load on processing during learning will be presented and approaches for measuring cognitive load during a learning experience explored.

## 2.2. Epistemic practices and pedagogy in engineering education

### *2.2.1. Positioning the epistemic practices of engineering education*

There is a viewpoint that engineering education can be a catalyst for interconnected Science, Technology, Engineering, and Mathematics (STEM) education, and thus increase the ability to achieve holistic STEM education (Simarro & Couso, 2021). Engineering design has been viewed as a context for the application of scientific and mathematical concepts (Barak, 2013). Some authors assert that engineering education can improve students learning in science and mathematics, increase technological literacy and knowledge of engineering, thus stimulating interest in engineering as a career option (Simarro & Couso, 2021). Other authors believe that there is pedagogical value in using engineering design approaches in other discipline areas to promote students' meaningful engagement in the learning experience (Cunningham & Kelly, 2017). While it is important to consider the value that engineering education can bring to other STEM disciplines, it is also important to reflect on the epistemic practices and processes of the discipline which set it apart from other STEM disciplines. However, minimal research has been conducted exploring the epistemology of engineering disciplines (Cunningham & Kelly, 2017). Simarro and Couso, (2021) differentiate the epistemic process of engineering and science in terms of their use of models, whereby science disciplines use models as reasoning artefacts and engineering uses models as a mechanism of evaluation. The epistemic practices of engineering and science are further differentiated in the literature in terms of their goals. In science, there are goals of generating new and verifiable knowledge, whereas engineering goals are often to solve specific problems (Cunningham & Kelly, 2017). Through problem solving, engineering transcends the need of specific knowledge, instead requiring a combination of knowledge types and reasoning to reach a solution (Cunningham & Kelly, 2017; Litzinger et al., 2010). A review of the literature and practices in engineering and engineering education (Cunningham & Kelly, 2017) identified 16 epistemic practices of engineering which were classified into four categories:

1. Engineering in social contexts
2. Uses of data and evidence to make decisions
3. Tools and strategies for problem solving
4. Finding solutions through creativity and innovation

These practices highlight important aspects of engineering and learning to become an engineer, but it is necessary to reflect on how engineering education has evolved to incorporate these practices.

### *2.2.2. Evolution of engineering education pedagogy*

Since the formal inception of engineering education in the 19th century, tensions between theory and practice have permeated the field with demand for reforms in engineering education intensifying around the 1970's-1980's, continuing to this day (Crawley et al., 2014; Felder, 2012). The agenda for engineering education has evolved over time being driven by the changeable emphasis placed on practical orientation versus theoretical priorities such as a scientific orientation (Crawley et al., 2014; Harwood, 2006; Reynolds & Seely, 1993). Schools of engineering offered radically different strategies with some providing practical training, contrasting others emphasising engineering fundamentals favouring laboratory settings, while others endeavoured to implement both strategies (Reynolds & Seely, 1993). Minimal efforts, if any, were made to standardise and regulate engineering education programmes until the early 20th century (Prados et al., 2005). Countries initially moved to standardise and regulate their own engineering education programmes before this culminated in international agreements such as the Bologna declaration, the Washington Accord, and criteria set out by the Accreditation Board for Engineering and Technology (Akera, 2017; Anwar & Richards, 2018; Dixit & Pathak, 2012; Dyrenfurth & Murphy, 2006; McGrath, 2000; Patil et al., 2007; Prados et al., 2005; Shearman, 2007). While these agreements set out the learning objectives and content to be delivered, they do not specify educational approaches which should be implemented to achieve them. Crawley et al. (2014, p. 250) note that reforms in engineering education approaches have been driven by a new awareness of the contribution of science to innovation and the use of technology in context. Additionally, reforms in engineering education have emphasised the need to engage students in problem solving and project work simulating real-world engineering practice (Crawley et al., 2014; Edström & Kolmos, 2014). It is intended that through such an approach engineering education may bridge the gap between “disciplinary knowledge of the technical sciences and social sciences and the practical domains of engineering with their unique knowledge and routines that integrate the social, practical, and technical aspects of technology at work” (Crawley et al., 2014, p. 250). This presents a significant and complex challenge for contemporary engineering education programmes to ensure they are providing students with a holistic experience to acquire the necessary knowledge, skills, and awareness’s to succeed in a rapidly advancing engineering space.

### *2.2.3. Contemporary pedagogical approaches in engineering education*

The role of an engineer has evolved extensively over the last number of years in conjunction with societal requirements and beginnings of IR 4.0 (Engineers Ireland, 2021; Kamaruzaman et al., 2019; SEFI, 2016; UNESCO, 2010). Throughout these changes, engineering educators and providers have been working to respond to industry skill needs and equip engineering graduates with a broad skillset to adapt to the demands of their ever-evolving future role (Flening et al., 2021). Over the last number of years, the value of using applied educational approaches has been explored as they provide students with experiences similar to those in the engineering profession (Edström & Kolmos, 2014). This has been outlined as a fundamental need to be addressed in engineering education programmes by practicing engineers (Jonassen et al., 2006). PBL and CDIO are examples of applied contemporary educational approaches that are implemented in third-level engineering programmes throughout the world (J. Chen et al., 2021; Crawley et al., 2014; Edström, 2016; Edström & Kolmos, 2014; Savin-Baden, 2008; Strobel & Barneveld, 2015; van Barneveld & Strobel, 2011). PBL and CDIO are based on a learner-centred pedagogy whereby students are active participants in their own learning (Crawley et al., 2014; van Barneveld & Strobel, 2011). Active learning encompasses a broad range of activities including action-orientated lectures, cooperative and collaborative learning, and problem-based activities and projects which require the highest level of student activity (Hadgraft & Kolmos, 2020). Through learner-centred approaches, also referred to as student-centred, there is a transition away from traditional approaches where students are lectured by an academic towards a more engaging and involving approach where students actively contribute to the direction of their own learning within a given framework (Hadgraft & Kolmos, 2020) e.g., PBL or CDIO. These approaches are documented to have positive effects on learning outcomes and promote sustained knowledge and skills development (Hadgraft & Kolmos, 2020; van Barneveld & Strobel, 2011).

The CDIO approach is underpinned by twelve standards, with seven of these described as essential for a minimal approach to develop a CDIO programme (Crawley et al., 2014; Edström & Kolmos, 2014). The premise of the approach is to provide students with a comprehensive and holistic learning experience (Edström & Kolmos, 2014). It is intended that students will develop a deep understanding of key concepts and principles, develop both technical and professional skills, and apply their knowledge and skills to problems representative of real-world engineering problems

(Crawley et al., 2014; Edström & Kolmos, 2014; Litzinger et al., 2011). A strength of the CDIO approach is the existence of a rubric for rating programmes (CDIO, 2010). This rubric provides dimensions for both implementing the CDIO approach and for systematically monitoring the programmes development (Edström & Kolmos, 2014; Malmqvist, 2012; Malmqvist, Edström, et al., 2006; Malmqvist, Östlund, et al., 2006). In contrast however, there are different approaches to implementing PBL which continue to evolve over time with regards to content and educational methods (Edström & Kolmos, 2014). Strobel and van Barneveld (2009) present PBL as an approach where students are presented with ill-defined problems which encourage them to generate multiple thoughts on both the cause and solution to the problem. This approach is noted to be more common at the latter stages of engineering programmes whereby the use of PBL is to afford students the opportunity to transfer the skills they have acquired through their education, and to practice and apply their professional skills (Mitchell & Smith, 2008; van Barneveld & Strobel, 2011). Edström and Kolmos (2014) outline three core principles to PBL: cognitive learning, collaborative learning, and content in the curriculum. Within the cognitive learning principle are problem, project, experience, and context. In contrast to the perspective of Strobel and van Barneveld (2009), it is noted that with the problem orientation "learning starts by analysing and defining problems, be they open and ill-defined, or well-defined" (Edström & Kolmos, 2014, p. 544). The contrasting views of PBL between (Edström & Kolmos, 2014; Strobel & van Barneveld, 2009) are just one example of the different perceptions of PBL across engineering education research literature. In a review of the literature, J. Chen et al. (2021) highlight that the lack of consensus on a specific approach to PBL has been attribute to challenges in the implementation of the approach.

Irrespective of the differences between the structures of these prevalent engineering education approaches, central to each is the goal to support students in acquiring  expertise to succeed in engineering (Caicedo et al., 2014; Crawley et al., 2014; Edström & Kolmos, 2014; Leandro Cruz et al., 2020; Litzinger et al., 2011; Passow & Passow, 2017). Expertise are expert knowledge or skills that an individual has in a particular field and are organised around key concepts in that space e.g., concepts of mass, force, and other general principles in engineering (Litzinger et al., 2011). How expertise are acquired and the role of deliberate practice in expertise and task performance has been the subject of debate among researchers for a number of years (Hambrick et al., 2016). One view is that experts are “born” and while training is required to become an expert in a particular field, an individual’s innate ability limits the performance that they can reach in a domain (Hambrick et al., 2016). The contrasting perspective is that experts are “made” arguing that where an individual may have an innate ability, this can be surpassed through deliberate practice in a domain (Ericsson, 2006, 2008; Keith & Ericsson, 2007). Although positive effect sizes have been found between deliberate practice and performance on discipline-specific tasks, the proportion of performance that deliberate practice can explain is less than the amount that is unexplained (Hambrick et al., 2016; Macnamara et al., 2014). This indicates that to support graduates in acquiring engineering expertise and capacity to solve complex real-world engineering problems, there is a need to go beyond technical knowledge and procedures which can be acquired through deliberate practice to consider other factors which may underpin performance.

### 2.3. Problem solving in the context of engineering education

To address the problem-solving skills gap highlighted by engineering employers (Nair et al., 2009; Valentine et al., 2017; Ward & Thiriet, 2010), and the skills need for effective problem solvers through IR 4.0 (Kamaruzaman et al., 2019), it is necessary to understand the factors and variables underpinning engineering problem solving performance and examine their potential development through current educational approaches. Problem solving is considered as a search process through memory to identify relationships between problem goals and alternative paths which can be used to reach a solution (Carlson & Bloom, 2005; Mayer & Wittrock, 2006; Y. Wang, 2007; Y. Wang & Chiew, 2010). When engaging in problem solving, either consciously or sub-consciously, individuals use strategies to categorise problems, focus attention, integrate information, and facilitate knowledge transfer (Litzinger et al., 2010). To facilitate the development of students’ capacity to solve a wide range of problems, engineering education programmes incorporate problems of different structures to familiarise students with adapting to different problem scenarios (Crawley et al., 2014; Edström & Kolmos, 2014).

### *2.3.1. The varying structure of problems*

Problems are typically defined by (Jonassen, 2010; Jonassen & Hung, 2008):

- Context/domain
- Type
- Process
- Solution.

The context of a problem for example might be bound to a field of study, to the knowledge of the problem, or knowledge for the solution (Barlex & Steeg, 2017). Knowledge of the problem relates to an understanding of the situation in which the problem is embedded (Barlex & Steeg, 2017). For instance, if an individual was tasked with designing and making an aid to improve the quality of life of elderly individuals, they would require knowledge of the issues that these individuals experience in their day-to-day living. Knowledge for the solution relates to an understanding or awareness of a component/s of the solution regardless of the situation or task they arise in, such as the role or function of gears which remains the same irrespective of working as part of a toy or a car (Barlex & Steeg, 2017). Depending on the context of the problem, different knowledge may be required by the problem solver. The knowledge required and the relative level of knowledge of the problem solver thus influences the problem type or structure, depicted in Figure 3. Problem structure can vary depending on the degree to which the problem is well-defined or ill-defined and whether it is open-ended or closed (Dörner & Funke, 2017; Jaarsveld & Lachmann, 2017; Jonassen & Hung, 2015).

The degree of problem definition describes the specifications of the problem that are made available to the problem solver. Well-defined problems set out clear constraints whereas ill-defined problems are not evidently specified (Dörner & Funke, 2017; Jaarsveld & Lachmann, 2017; Jonassen, 1997, 2000; Schraw et al., 1995). The degree to which a problem is open-ended or closed describes the solution options available to the problem solver. Closed problems have a limited number of solutions which can be determined with absolute certainty through the implementation of a distinct set of specific procedures (Dörner & Funke, 2017; Jaarsveld & Lachmann, 2017; Jonassen, 2000; Schraw et al., 1995). In contrast, the number of possible solutions in open-ended problems is not restricted and the procedure to solve the problem may not be apparent or predictable (Dörner & Funke, 2017; Jaarsveld & Lachmann, 2017; Jonassen, 1997). Authentic problems in technological disciplines are often more open-ended and ill-defined (de Vries, 2016), such as engineering design problems (Gómez Puente et al., 2015). However, in educational settings such as engineering education, it is difficult for educators to simulate these authentic ill-defined open problems that engineering graduates will experience in their future work as the emphasis is placed on facilitating the acquisition of their expertise to address such problems (Flening et al., 2021).

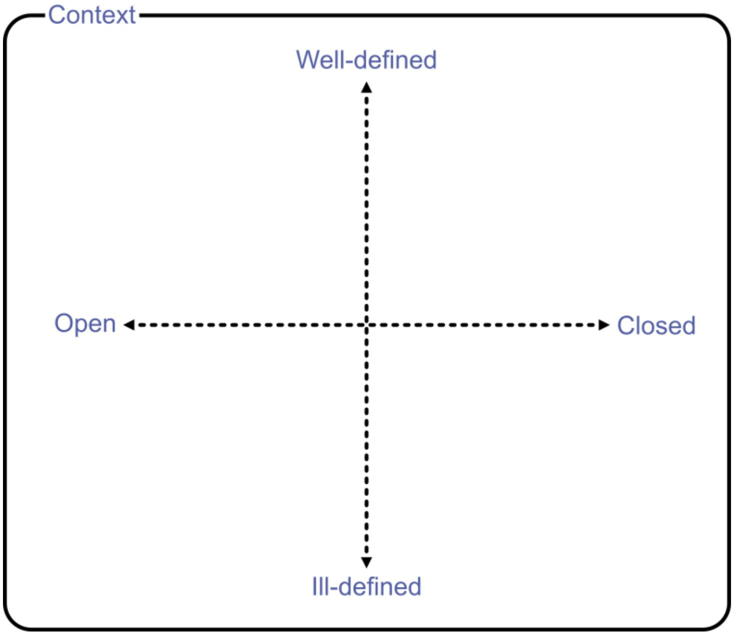


*Figure 3. Well-defined vs ill-defined and open vs closed ended dichotomies where problems can be situated and bound in a specific context.*

As an individual works to solve a problem, regardless of structure, their interactions with the problem may also vary depending on the degree to which they are operating inside or outside the head (Kimbell et al., 2004) and whether they are working individually or collaboratively (P. A. Kirschner et al., 2018), as demonstrated through Figure 4. As an individual engages with a problem they can repeatedly transition from considering and conceiving ideas in their mind to realising and modelling ideas with their hands (Kimbell et al., 2004). It is worth noting that these modes of designerly activity are not mutually exclusive, whereby an individual, when working with their hands are also likely to be considering elements of the process in the mind. Within this process, there are iterations of ideas in an individual's mind and these ideas can be expressed in the external world through appropriate models varying from discussion (verbal modelling) to graphic and physical models (Kimbell et al., 1991). Depending on how the endedness of the problem is defined this can happen once or be repeated several times supporting continuous iterations of ideas and models, as is the case when solving design problems (Kimbell et al., 2004; Ramey & Uttal, 2017). In design problems these iterations of ideas and models can be used to generate and test solutions which may also be presented as elements of the solution to the problem (Ramey & Uttal, 2017).

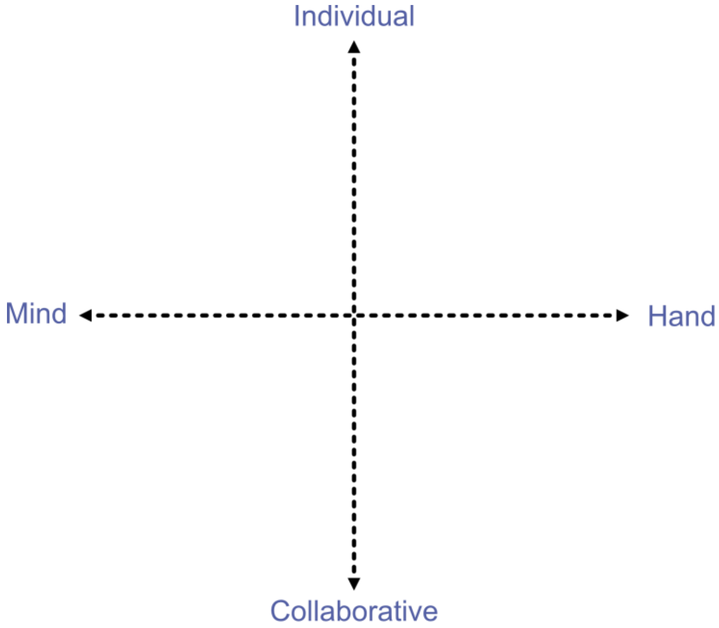


*Figure 4. Individual vs collaborative and mind vs hand dichotomies where an individual's interaction with the problem can vary between.*

The necessity to model ideas in the mind or hands would also vary depending on the individual or group of individuals solving the problem. Working collaboratively is of particular importance in STEM disciplines as some problems are extremely complex and often require contributions of team members with specific expertise (Crawley et al., 2014; Stieff, 2007; Wai et al., 2009). When problem solving collaboratively, tasks may be divided amongst learners in the group resulting in some modelling in the mind and others modelling with their hands. This division of tasks can result in a reduction in the demands of cognitive processing capacity as the load can be shared and distributed among the group (P. A. Kirschner et al., 2018). For instance, if a learner in a group has a higher level of knowledge or skill in a particular area in comparison to another, the deficits of one can be compensated by the strengths of the other (P. A. Kirschner et al., 2018). Therefore, it is important to consider that different members involved in collaborative problem solving could require the use of different additional cognitive processes to different degrees in comparison to problem solving at an individual level. In measuring or assessing the relative success of problem solving, irrespective of it being an individual or group task, is also dependent on the parameters of each individual problem. For instance, measuring success in solving a design problem may look at the suitability of the solution to address the problem brief, design decision skills, or being able to justify the solution (Dixon & Johnson, 2011; Gómez Puente et al., 2013; Gomez Puente et al., 2014; Savin-baden, 2014). While measuring success with a more structured complex problem may evaluate performance in terms of the use of a specific procedure and efficiency in applying that procedure e.g., using specific strategies (Greiff & Funke, 2009; Lotz et al., 2016; Rudolph et al., 2018; Schiff & Vakil, 2015). The capacity to solve these complex problems, irrespective of structure, is underpinned by knowledge relative to the context of the problem and an individual's problem-solving skill proficiency (Litzinger et al., 2010; Mainali, 2012).

### *2.3.2. Problem solving process*

Problem solving is recognised as being a higher-order cognitive process (Barak, 2013; Mainali, 2012; Y. Wang & Chiew, 2010). Within the Layered Reference Model of the Brain (LRMB) (detailed in Figure 5), which encompasses 37 cognitive processes at six layers, problem solving is categorised within Layer 6- Higher cognitive functions along with factors such as imagery, comprehension, learning, reasoning, and decision making (Y. Wang, 2007; Y. Wang et al., 2006; Y. Wang & Chiew, 2010). Efforts to understand and advance individuals' capacity to solve problems have been a research interest for several years for researchers from multiple discipline areas. From a human intelligence perspective, problem solving is noted to be a basic life function, the most complex intellectual function, and related to an individual's capacity to perceive and learn (Barak, 2013; Y. Wang & Chiew, 2010). Being a higher-order cognitive process, problem solving interacts with various other cognitive processes including abstraction, analysis, decision making, inference, learning and synthesis based on internal knowledge representation (Y. Wang & Chiew, 2010). During problem solving, cognitive and metacognitive strategies act together to reach a solution (Litzinger et al., 2010). An individual's cognitive strategies are used for task execution, whereas their metacognitive strategies support understanding of how the task is performed and regulate actions (Litzinger et al., 2010; Schraw, 2001). The characteristics which contribute to an individual being a successful problem solver include: the capacity to correctly identify problem goals, persistence, ability to adopt efficient search strategies, regulation of actions, and capacity to trace back to previous points in the solution process (Litzinger et al., 2010; Y. Wang & Chiew, 2010).

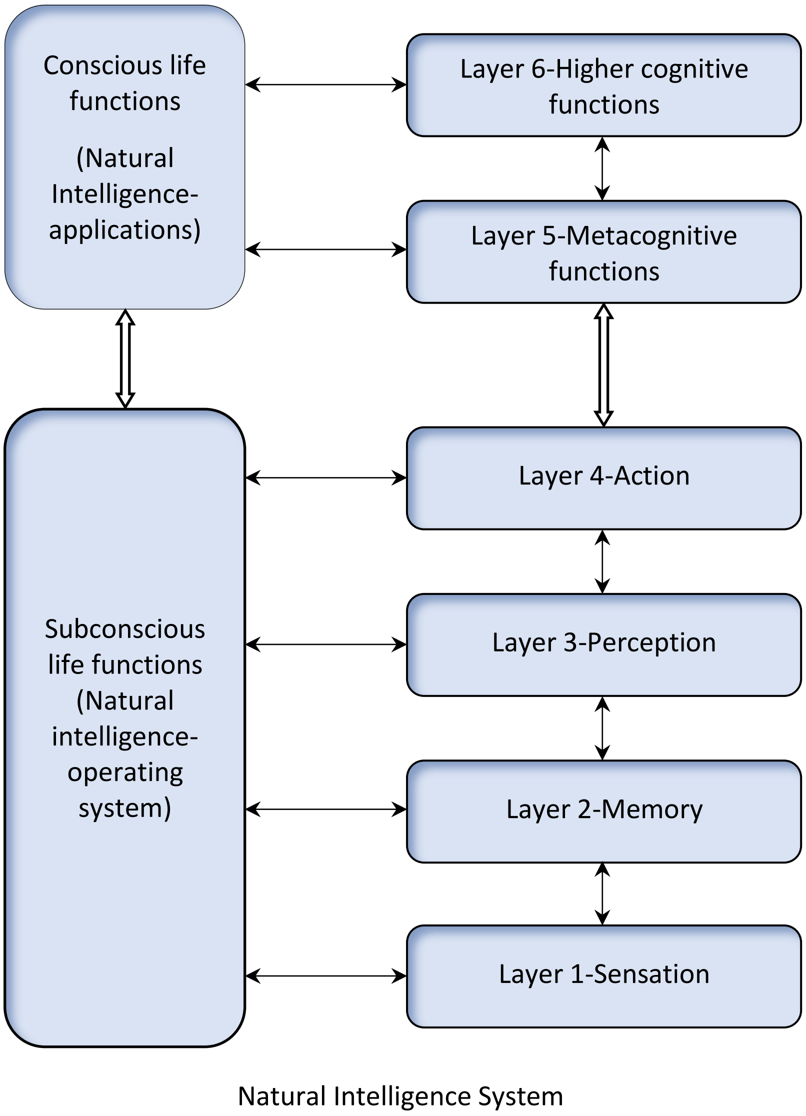


*Figure 5. The Layered Reference Model of the Brain (LRMB) presented by* (Y. Wang et al., 2006).

Problems are considered to exist when there is a goal to be reached but an individual does not know how to reach this goal (Buckley et al., 2018a; Fischer et al., 2011; Schoenfeld, 1983). When faced with a problem an individual has to recourse to thinking (Fischer et al., 2011). Mainali (2012) describes problem solving as a component of higher-order thinking, whereby the nature of this thinking involves:

- open ended problem solving where multiple solution options are possible,

- the path to solve the problem is not specified in advance,
- there is significant mental energy invested in solving the problem,
- there are subtle decisions about strategies to solve the problem not obvious to the problem solver,
- there is not always a clear starting point to the problem,
- there is an element of metacognition and self-awareness,
- development and/or application of new theories to sets of facts and problems.

In essence, higher order thinking through this description is problem solving. The characteristics set out by Mainali are similar to those set out by others as the characteristics of complex problem solving (Dörner & Funke, 2017; Funke, 2012). Funke (2012) describes complex problem solving as taking place to reduce the barrier between a specific start state and intended goal state which can be supported with the help of cognitive activities and behaviour. The start state, goal state, and barriers to reach a solution differentiate complex problem solving from simple problem solving. In a complex problem these elements can change dynamically over time and be partially unclear (Funke, 2012). Comparatively, when solving simple problems, the exact start state, goal state, and barriers are known.

Problem solving is initiated by constructing internal representations of the external problem statement creating the “problem space” (Fischer et al., 2011). The problem space is all of the possible goals and paths which may be used to reach a solution which are known to the problem solver (Y. Wang & Chiew, 2010). Forming this space requires reflection of various problem states given the initial problem state, applicable operators, and particular goal states (Fischer et al., 2011; Y. Wang et al., 2006). Operators are possible actions which can be implemented to achieve the goals of the desired solution (Y. Wang & Chiew, 2010). The operators which may be applicable can vary between problem solvers depending on factors such as different levels of expertise or intelligence (Fischer et al., 2011; Newell & Simon, 1972). When an individual has formed an internal representation of the problem, they begin to consider methods which may be used to reach the goal state (Fischer et al., 2011). In situations where the problem solver does not have expertise relative to the problem context, or to reduce the effort and difficulty of the search, they can rely on heuristics such as “means-end-analysis” or “take-the-best” to support them in identifying methods to solve the problem (Fischer et al., 2011; Shah & Oppenheimer, 2008; Y. Wang et al., 2006). Heuristics can be classified as search, stop, or decision-making rules (Gigerenzer, 2001). Employing heuristics, the intention is not to find the optimal approach to solving the problem, rather to satisfice (Gigerenzer, 2008). When engaging with a complex problem individuals rely on a combination of both conceptual knowledge and cognitive strategies (e.g., heuristics) (Litzinger et al., 2010). The cognitive strategies that an individual has acquired are used to categorise problems, focus attention, integrate information, and facilitate knowledge transfer (Litzinger et al., 2010; Schraw et al., 2006). Some scholars believe that complex problem solving is a facet of general intelligence, although others view it as distinct from intelligence (Rudolph et al., 2018). Within contemporary theories of human intelligence, problem solving is not outlined as a factor of intelligence. However, successful problem-solving performance can be supported by an individual’s relative level of intelligence and ability within specific cognitive factors. Where an individual has not acquired the necessary expertise to solve the problem, they may utilise factors of their intelligence to circumvent the limits in their domain knowledge to reach a solution (Hambrick et al., 2018; Hambrick & Meinz, 2011; Kulasegaram et al., 2013; Meinz et al., 2012). The premise of the circumvention-of-limits hypothesis is reflected by a Knowledge x Ability interaction whereby factors of intelligence, or cognitive abilities, are most important for individuals with low levels of domain knowledge but not individuals with high levels of domain knowledge (Hambrick et al., 2012; Meinz et al., 2012). This interaction is demonstrated in Figure 6 where individuals with high levels of domain knowledge rely less on cognitive abilities to solve the problem than those with lower levels of expertise. Within the theory, the concept is that as individuals engage in direct practice, relative to the discipline area, they acquire cognitive strategies such as heuristics and problem-solving schemas that support them in solving problems without having to rely significantly on cognitive abilities (Kulasegaram et al., 2013). However, domain-general abilities relate to performance for experts in some situations such as dealing with a novel topic or task (Hambrick & Meinz, 2011). Additionally, through the circumvention-of-limits hypothesis, it is proposed that individuals with low levels of expertise but high levels of pertinent cognitive abilities to solving the problem can perform to a similar standard as those with high levels of domain knowledge (Hambrick et al., 2012). As such, it is important to consider theories of intelligence when seeking to understand problem solving given the contribution of cognitive abilities to solving complex and novel problems.

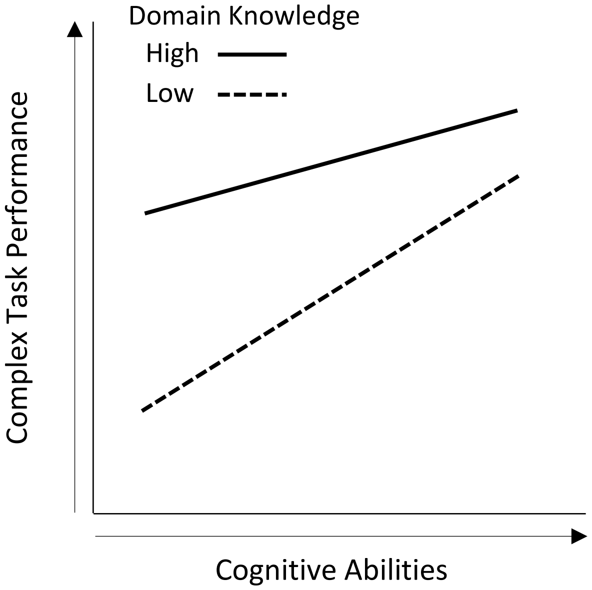


*Figure 6. The interaction between cognitive abilities and complex task performance for individuals with different levels of domain knowledge proposed through the circumvention-of-limits hypothesis.*

## 2.4. Theory of human intelligence

### 2.4.1. Evolution of human intelligence theory

Factors associated with human intelligence have been an area of research focus for a significant period. A single general cognitive ability (*g*) was originally represented through the work of Spearman (1904) where he examined the association between responses to sensory stimuli and educational performance through factor analyses. Advancing on the work of Spearman, Thurstone (1938) developed new factor analysis methods which identified 13 group factors, classifying seven of these as primary mental abilities, also referred to as second-order factors. These second-order factors were space, perceptual speed, number facility, verbal relations, word fluency, memory, and induction (Thurstone, 1938). No general factor was identified through Thurstone's work. Following these empirical investigations of intelligence factors, other psychologists began to develop theories of intelligence. Through advancement of the theories, hierarchical models began to emerge such as that of Burt (1949) who presented a model with various levels of dichotomy, the first of which was between intellectual (*g*) and practical/behavioural characteristics. While Vernon (1950) theorised an alternative hierarchical model where *g* was the primary factor and all other factors derived from *g*. This was followed by the theory that *g* consisted of two separate factors, fluid intelligence (Gf) and crystallised intelligence (Gc), which was conceptualised by Cattell, (1943). Fluid intelligence was first described as having "the character of a purely general ability to discriminate and perceive relations between any fundamentals, new or old" and crystallised ability as consisting of "discriminatory habits long established in a particular field, originally through the operation of fluid ability, but no longer requiring insightful perception for their successful operation" (Cattell, 1943, p. 178). In other words, fluid intelligence is considered as a general cognitive ability associated with identifying relations in any context, while crystallised intelligence describes learned knowledge and its application. The Gf-Gc theory was extended through the work of Cattell and Horn and posited as a hierarchical model with Gf and Gc represented above several lower order abilities (Cattell, 1963;Cattell & Horn, 1978; J. Horn, 1985, 1988; J. Horn & Cattell, 1966; Schneider & McGrew, 2012). Later, Carroll (1993) advanced the theory of the structure of human intelligence further by presenting a hierarchical three-stratum theory which was the result of a meta-analysis of more than 460 psychometric datasets. Carroll presented an empirically based taxonomy of human cognitive abilities with stratum three consisting of general cognitive ability (*g*), stratum two consisting of eight broad cognitive abilities, and stratum one consisting of several narrow abilities. A synthesis of Carroll's three-stratum theory and Gf-Gc theory led to the current most extensive organisation and structure of cognitive abilities, the Cattell-Horn-Carroll (CHC) theory of human intelligence (Schneider & McGrew, 2018).

### *2.4.2. Contemporary understandings of intelligence*

The CHC theory was intended to provide a connection between theory and practice in human intelligence research (McGrew, 2005, 2009). It evolved from the two-stratum structure of the Gf-Gc theory where general cognitive ability was omitted (McGrew, 2005), to the current model with cognitive abilities in a hierarchical three-stratum structure based on the extensive empirical work conducted by Carroll following the conception of the Gf-Gc theory (Carroll, 1993; Schneider & McGrew, 2012). The CHC theory expanded the broad cognitive abilities to sixteen from the original eight presented through Carroll's theory (McGill & Dombrowski, 2019; Schneider & McGrew, 2018). While there is contention in the literature about the order, inclusion, and exclusion of some of the abilities within the CHC model (McGill & Dombrowski, 2019), these concerns were noted by Carroll (1997) in earlier works in relation to his own findings where he acknowledged that through future research some errors may be revealed. Despite this contention, the CHC model presents the current most extensive organisation and structure of cognitive abilities and provides a suitable foundation for the consideration of the cognitive processes and abilities which underpin problem solving in the engineering domain.

Within the three-stratum structure of the CHC theory, several narrow abilities contribute to the structure of each of the broad cognitive abilities, detailed in Table 1, which subsequently represent elements of overall general cognitive ability (McGill & Dombrowski, 2019; Schneider & McGrew, 2018). The broad cognitive abilities outlined through the CHC model are not necessarily a list of abilities which will be used as an individual solves a problem. The broad abilities might be critical to performance to varying degrees in different disciplines or fields of study depending on the variability in problem parameters. The necessity of each of these abilities for an individual to solve a problem is dependent on the problem requirements. For example, fluid intelligence is likely to be of greater significance for solving novel problems than domain-specific knowledge, as fluid intelligence describes the ability to reason in novel situations (Oswald et al., 2014; Schneider & McGrew, 2012), and the domain-specific knowledge may be unrelated to the problem. Likewise, cognitive abilities can be viewed in terms of their necessity for a particular discipline; for example, a chef is likely to use their olfactory and tactile abilities frequently, whereas comparatively a mechanical engineer could rely more on fluid intelligence and visual processing due to the nature of the problems they are required to solve.

*Table 1. Descriptions of the broad cognitive abilities from the CHC theory of human intelligence (Schneider & McGrew, 2018).*

| **Broad cognitive abilities** | **Description** |
|---|---|
| Fluid reasoning | “The use of deliberate and controlled procedures (often requiring focused attention) to solve novel, “on-the-spot” problems that cannot be performed by using previously learned habits, schema, and scripts” |
| Working memory capacity | “The ability to maintain and manipulate information in active attention” |
| Learning efficiency | “The ability to learn, store, and consolidate new information over periods of time measured in minutes, hours, days, and years” |

| Broad cognitive abilities | Description |
|---|---|
| Retrieval fluency | "The rate and fluency at which individuals can produce selectively and strategically retrieve verbal and nonverbal information or ideas stored in long-term memory" |
| Processing speed | "The ability to control attention to automatically, quickly, and fluently perform relatively simple repetitive cognitive tasks. May also be described as attentional fluency or attentional speediness" |
| Reaction and decision speed | "The speed of making very simple decisions or judgements when items are presented one at a time" |
| Psychomotor speed | "The ability to perform skilled physical body motor movements (e.g., movement of fingers, hands, legs) with precision, coordination, fluidity, or strength" |
| Comprehension-knowledge | "The depth and breadth of knowledge and skills that are valued by one's culture. It includes the depth and breadth of both declarative and procedural knowledge, and skills such as language, words, and general knowledge developed through experience, learning and acculturation" |
| Domain-specific knowledge | "The depth, breadth, and mastery of specialised declarative and procedural knowledge (knowledge not all members of a society are expected to have)" |
| Reading and writing | "The depth and breadth of declarative and procedural knowledge and skills related to written language" |
| Quantitative knowledge | "The depth and breadth of declarative and procedural knowledge related to mathematics" |
| Visual processing | "The ability to make use of simulated mental imagery to solve problems- perceiving, discriminating, manipulating, and recalling non-linguistic images in the "mind's eye"" |
| Auditory processing | The "ability to discriminate, remember, reason, and work creatively (on) auditory stimuli, which may consist of tones, environmental sounds, and speech units" |
| Olfactory abilities | "The ability to detect and process meaningful information in odors" |
| Tactile abilities | "The abilities to detect and process meaningful information in haptic (touch) sensations. This domain includes perceiving, discriminating, and manipulating touch stimuli" |
| Kinaesthetic abilities | "The abilities to detect and process meaningful information in proprioceptive sensations. Proprioception refers to the ability to detect limb position and movement via proprioceptors (sensory organs in muscles and ligaments that detect stretching)" |
| Psychomotor abilities | "The abilities to perform physical body motor movements (e.g., movement of fingers, hands, legs) with precision, coordination, or strength" |

Through analysis of these abilities, it is evident that cognitive ability is a complex phenomenon (Carroll, 1993; Lubinski, 2004; McGrew, 2009; Schneider & McGrew, 2018). In addition to identifying important cognitive abilities which should be considered within education from the perspective of cognitive development, the relationship of these abilities to learning must also be considered (F. Kirschner et al., 2009; P. A. Kirschner et al., 2018; Sweller et al., 2011). In the development of expertise, novices can utilise cognitive abilities to compensate for a lack of discipline knowledge when addressing a problem, therefore circumventing the limits of their information processing capacity and knowledge (Hambrick et al., 2012; Hambrick & Meinz, 2011), as earlier outlined. Considering the role of cognitive abilities in circumventing limitations in domain-specific knowledge and viewing this in the context of engineering problem solving, spatial ability, referred to as visual processing in the CHC theory, is a pertinent factor to be considered. Spatial ability has previously been demonstrated to support problem solving in other spatially orientated STEM disciplines for individuals not familiar with the procedures to solve the problem (Hambrick et al., 2012). Given these findings and relative importance of spatial ability in engineering (Maeda et al., 2013; Metz & Sorby, 2016; Sorby et al., 2014; Uttal & Cohen, 2012; N. Veurink & Sorby, 2012; Wai et al., 2009), it is necessary to further explore this relevant cognitive factor.

## 2.5. Spatial ability

### 2.5.1. Defining spatial ability

Galton (1879, 1880) introduced the concept of spatial ability with Spearman and Thurstone providing quantitative evidence of its existence (Buckley et al., 2018a). Various definitions of spatial ability are offered through the literature. Carroll (1993) described spatial ability as the capacity of an individual to successfully mentally manipulate visual patterns. While Lohman, describes spatial ability as “the ability to generate, retain, retrieve, and transform well-structured visual images” (Lohman, 1994, p. 1000). These visual images can include rotation, to folding and unfolding of objects (Lohman, 1988, 1989). Meehl (2006) outlined verbal definitions of intelligence factors are not sufficient and that it should instead be defined based on empirical frameworks. Throughout this thesis, spatial ability will be viewed through the empirical taxonomy of intelligence.

Within the CHC theory of intelligence (based on the empirical work of Cattel, Horn, and Carroll (Carroll, 1993; R. B. Cattell & Horn, 1978; J. Horn, 1985, 1988; J. L. Horn & Cattell, 1966)), spatial ability is defined as “the ability to make use of simulated mental imagery to solve problems- perceiving, discriminating, manipulating, and recalling non-linguistic images in the “mind’s eye”” (Schneider & McGrew, 2018, p. 125). The term ‘visual processing’ is used to describe spatial ability through the CHC framework, where it is a second-order factor of intelligence. It is detailed to consist of 12 first-order factors, also referred to as narrow factors, through the most current version of the model, depicted in Figure 7 (Schneider & McGrew, 2018). These narrow factors have been identified through an extensive body of empirical work which began with the work of Galton (1879, 1880), and was advanced through the work of a range of other researchers (Carroll, 1993; J. Horn, 1988; Lohman, 1994, 1996; Schneider & McGrew, 2018; Thurstone, 1938). The definitions of these narrow factors are set out through Table 2.

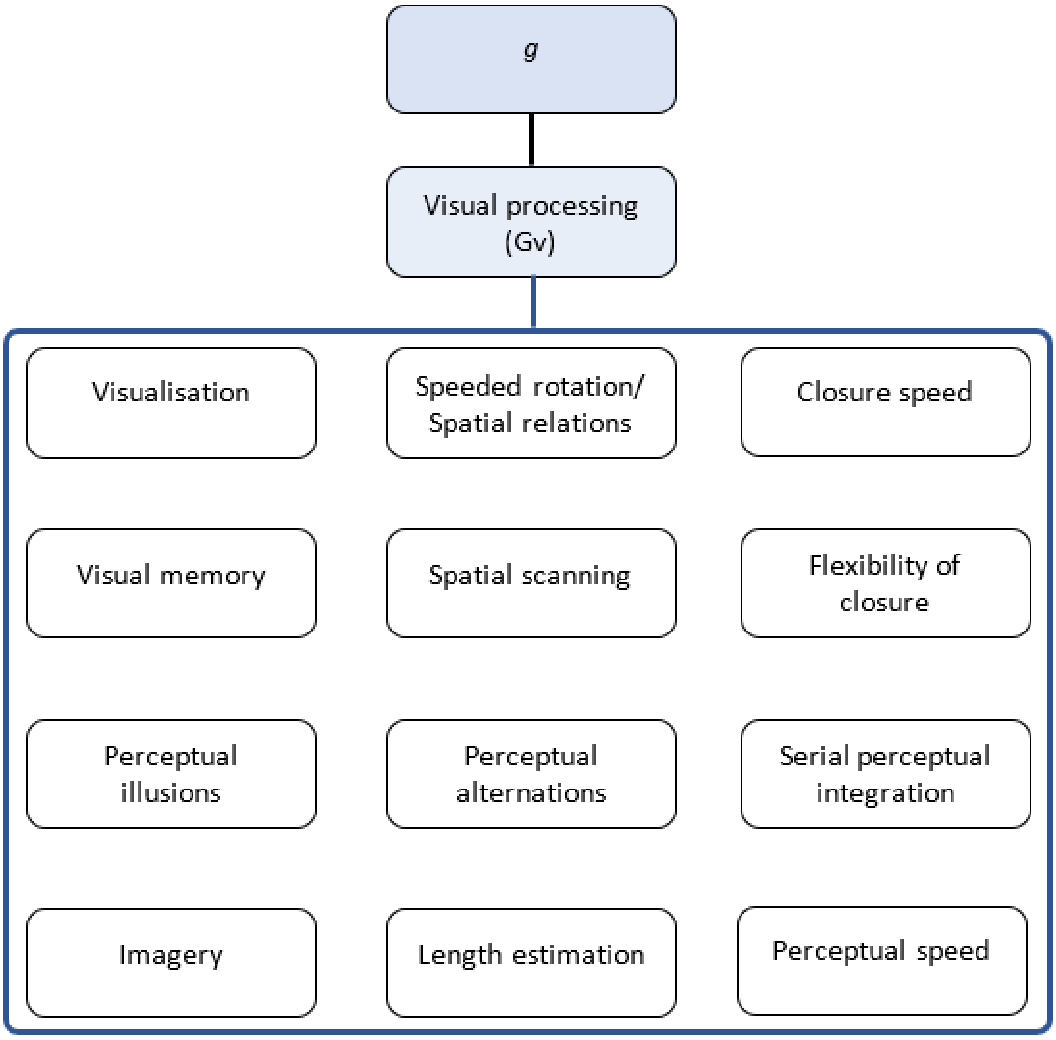


*Figure 7. Narrow factors of spatial ability (visual processing) within the CHC theory of human intelligence.*

*Table 2. Descriptions of the narrow factors of spatial ability from the CHC theory of human intelligence (Schneider & McGrew, 2018).*

| **Narrow factors of spatial ability** | **Description** |
|---|---|
| Visualisation | “The ability to perceive complex patterns and mentally simulate how they might look when transformed (e.g., rotated, twisted, inverted, changed in size, partially obscured)” |
| Speeded rotation/ spatial relations | “The ability to solve problems quickly by using mental rotation of simple images” |
| Closure speed | “The ability to quickly identify and access a familiar, meaningful visual object stored in long-term memory from incomplete or obscured (e.g., vague, partially obscured, disguised, disconnected) visual cues of the object, without knowing in advance what the object is” |
| Visual memory | “The ability to remember complex images over short periods of time (less than 30 seconds)” |
| Spatial scanning | “The ability to quickly and accurately survey (visually explore) a wide or complicated spatial field or pattern with multiple obstacles, and identify a target configuration or identify a path through the field to a target endpoint” |
| Flexibility of closure | “The ability to identify a visual figure or pattern embedded in a complex distracting or disguised visual pattern or array, when one knows in advance what the pattern is” |
| Perceptual illusions | “The ability not to be fooled by visual illusions” |
| Perceptual alternations | “Consistency in the rate of alternating between different visual perceptions” |
| Serial perceptual integration | “The ability to recognize an object after only parts of it are shown in rapid succession” |
| Imagery | “The ability to voluntarily mentally produce very vivid images of objects, people, or events that are not actually present” |
| Length estimation | “The ability to visually estimate the length of objects (without using measuring instruments)” |
| Perceptual speed | “The speed and fluency with which similarities or differences in visual stimuli can be distinguished” |

It is important to note that this is not an exhaustive list of all of the narrow spatial factors which exist. The factors documented here represent the validated factors within the current framework which are largely the outcome of bottom-up programmes of research which are predicated on developing tests for practical purposes such as prediction or diagnosis (Schneider & McGrew, 2018). Conceptualisations of static and dynamic spatial factors viewing spatial ability as a broader spatial thinking construct present a new typology to consider spatial ability within (Buckley et al., 2018a; Newcombe & Shipley, 2015; Schneider & McGrew, 2018). Recent conceptualisations of spatial ability under the static and dynamic constructs propose the 11 previous narrow factors documented through an earlier model of the CHC theory, and 16 possible additional narrow factors (Buckley et al., 2018a; Schneider & McGrew, 2012). However, validation of these factors is required through structural validity research (Schneider & McGrew, 2018).

### *2.5.2. Measuring and developing spatial ability*

There are various psychometric measures of spatial ability which load on the narrow spatial factors outlined in the CHC framework (Buckley et al., 2019). Visualisation is one of these narrow factors and in almost all studies showing the predictive capacity of spatial ability for performance outcomes, visualisation is used as a proxy measure for the broad factor spatial ability (Schneider & McGrew, 2018). This is due to visualisation being the closest factor to spatial ability and strongest loading factor in analytic models of multiple spatial factors (Buckley, 2020; Carroll, 1993). Psychometric tests such as the Paper Folding Test (PFT), Surface Development Test (SDT), Mental Rotation Test (MRT), and Purdue Spatial Visualisation Test: and Rotations (PSVT:R) all load on the visualisation factor of spatial ability (Bodner & Guay, 1997; Ekstrom et al., 1976; Guay, 1976; Vandenberg & Kuse, 1978). Within these tests, individuals are asked to identify what an object might look like when moved into a different position, after being cut or punched, and where lines of a development would appear when the object is folded up. These methods of analysis have been used in STEM research for several years (Kozhevnikov et al., 2007; Sorby, 1999; Sorby et al., 2013, 2014; Uttal, Miller, et al., 2013; Uttal & Cohen, 2012; Wai et al., 2009; Xu et al., 2016) and can be used to capture an extensive range of information regarding an individual’s spatial ability (Mohler, 2008).

Spatial ability is outlined as the innate ability that an individual has before any formal training has taken place, while spatial skills can be learned or acquired through training (Sorby, 1999). When examining the construct of spatial ability at a university level, it is almost impossible to distinguish between abilities and skills because there is not an understanding of whether any spatial training has occurred, therefore, the terms have been used interchangeably in these instances (Sorby, 1999). Spatial skills can be developed through activities that students engage in throughout their education (Kozhevnikov & Thornton, 2006; Metz et al., 2016; Newcombe, 2017; Ramey & Uttal, 2017). Olkun (2003) presents engineering drawing activities as a means of developing spatial ability in engineering education as it requires the visualisation of 2D and 3D solutions. Other approaches to developing spatial ability include interventions whereby the students engage in a series of spatial training activities to enhance their spatial skills such as semester-long programme, or video-game experiences (Martin-Dorta et al., 2009; Martín-Gutiérrez et al., 2015; Martín-Gutiérrez & González, 2017; Metz et al., 2016; Newcombe, 2017; Roca-González et al., 2017; Sorby, 1999, 2007, 2009; Sorby & Baartmans, 1996, 2000; Uttal, Meadow, et al., 2013; Uttal & Cohen, 2012). Measuring spatial ability can contribute to talent searches to identify individuals suitable to a career in STEM or to determine why and in what situations spatial ability contributes to success in the context of engineering and engineering education.

### *2.5.3. Influence of spatial ability in engineering education*

Significant evidence documents the gender differences in spatial ability, with males outperforming their female counterparts (Doyle et al., 2012; Feng et al., 2007; Metz & Sorby, 2016; Sorby, 1999; Sorby et al., 2013; Vandenberg & Kuse, 1978; Xu et al., 2016; Yilmaz, 2009). Although, given additional time on spatial tests/tasks this gender difference can disappear as females may implement an analytic or “piecemeal” approach to solving the problem, circumventing the requirement for spatial ability and invalidating the measure (Khooshabeh et al., 2011; Lippa et al., 2010; Maeda & Yoon, 2016; Voyer, 2011; M.-T. Wang & Degol, 2017). There are other areas of cognitive abilities where females outperform their male counterparts such as verbal ability (M.-T. Wang et al., 2013; M.-T. Wang & Degol, 2017). Differences in spatial ability are noted to begin emerging at approximately nine or ten years of age, however this may vary across populations (Sorby, 1999; Sorby et al., 2014; Uttal & Cohen, 2012; Yilmaz, 2009). A variety of explanations have been explored for the gender difference in spatial cognition, such as childhood experiences and perceptions (Doyle et al., 2012; Kell & Lubinski, 2013; Vandenberg & Kuse, 1978; Yilmaz, 2009). These gender differences have had far reaching affects throughout the world, including the entry of individuals into STEM disciplines (Ardies et al., 2015; Uttal & Cohen, 2012; Xu et al., 2016).

Spatial ability has been outlined as a predictor of success in STEM disciplines (Maeda et al., 2013; Metz et al., 2016; Olkun, 2003; Schneider & McGrew, 2018; Sorby et al., 2014; Uttal & Cohen, 2012; Veurink & Sorby, 2012; Wai et al., 2009). It has also been documented that an individual's pursuit of a career in STEM is influenced by their spatial ability and their belief in their ability (Maeda et al., 2013; Sorby et al., 2014).

Due to the nature of the engineering profession and advancements in both technology and the complexity of engineering problems, there is a greater importance for engineers to be able to visualise effectively in order to solve problems (Burgess et al., 2013; Y. S. Chang et al., 2016; Li & Fu, 2012; Ramey & Uttal, 2017; SEFI, 2016; Shuman et al., 2005). Disciplines of engineering are spatially demanding, with some, such as mechanical engineering, perceived as highly spatially orientated (Veurink & Sorby, 2012). An aim of education is to provide students with domain-specific knowledge and skills (Sweller, 2015) and therefore support them in acquiring the expertise to succeed. Information is often communicated through visual means in engineering and engineering education e.g., CAD and engineering drawings (Y. Chang, 2014; Y. S. Chang et al., 2016; Olkun, 2003). Therefore, spatial abilities, such as visualisation, retention, or mental manipulation (Carroll, 1993; Lohman, 1994; Schneider & McGrew, 2018), are necessary to understand the information presented to support the acquisition of expertise. Spatial ability is outlined as having a strong relationship with performance at all levels of expertise in STEM (Wai et al., 2009). Hambrick, et al. (2012), in the context of geology, determined that individuals with lower levels of expertise in an area and high levels of spatial ability can perform to a similar standard to those with high levels of expertise on an authentic problem. However, to date, such a study has not been carried out in engineering education to determine why spatial ability is indicative of success. Ramey and Uttal (2017) highlight this gap in the literature and through their work outline how spatial practices and processes affect learning and the types of activities that facilitate these practices and processes. Despite the importance of spatial abilities in STEM, they are outlined as neglected talents in both educational and occupational settings which, if incorporated to talent searches, could reduce the loss of talent in STEM disciplines (Kell & Lubinski, 2013; Wai et al., 2009).

### *2.5.4. Spatial ability for problem solving*

As contemporary engineering education programmes are increasingly employing problem-orientated approaches (Crawley et al., 2014; Edström & Kolmos, 2014), it is conceivable that the role of spatial ability as a component of success in engineering is due to its relationship with problem solving. Spatial ability plays an important role in problem solving (Hambrick & Meinz, 2011; Ramey & Uttal, 2017; Tzuriel & Egozi, 2010) and the communication of solutions (Y. Chang, 2014; Y. S. Chang et al., 2016; Olkun, 2003; Ramey & Uttal, 2017) which is a fundamental aspect of engineering (Jonassen et al., 2006). To solve a problem, an individual must build an understanding of its structure which is facilitated by forming a mental representation of the problem and reflecting on the perceived problem state and the desired outcome (Björklund, 2013). This can also be considered as the development of cognitive models which involves mentally “building and manipulating images so as to define, refine and communicate ideas and solutions” to problems (Gaughran, 2002, p. 3). Developing these cognitive models is underpinned by spatial ability with cognitive modelling also referred to as visualisation in intelligence research (Schneider & McGrew, 2012). Previous research has explored the differences in problem solving approaches implemented between individuals with different levels of spatial ability (Khooshabeh et al., 2011; Lin, 2016; Tzuriel & Egozi, 2010). Through these studies individuals with high levels of spatial abilities were found to use more holistic strategies for problem solving rather than more analytic or ‘piecemeal’ approaches which are less optimal. Holistic approaches are based on considering a model as a unit and dealing with this model as a whole unit whereas analytic or ‘piecemeal’ approaches break down elements of the model (Tzuriel & Egozi, 2010). Additionally, in working examining gender differences in problem-solving strategies, males were demonstrated to use more holistic strategies whereas females were found to use more analytic approaches which may be due to the noted gender differences in spatial ability (Lin, 2016; Tzuriel & Egozi, 2010). Holistic problem-solving strategies are demonstrated to be more efficient than analytic approaches (Tzuriel & Egozi, 2010). As a key component of engineering is efficiency in problem solving, spatial ability is a critical cognitive factor to be considered for supporting problem-solving performance. Given that spatial ability is a cognitive factor, a need emerges to consider the influence it may have on cognitive load experienced during problem solving.

## 2.6. Cognitive load

### *2.6.1. Defining cognitive load theory*

When the goal is learning (i.e., to invoke a change in students' long-term memory through the acquisition of knowledge), cognitive load needs to be managed (F. Kirschner et al., 2009; P. A. Kirschner et al., 2018). "Cognitive load theory aims to explain how the information processing load induced by learning tasks can affect students' ability to process new information and to construct knowledge in long-term memory" (Sweller et al., 2019, p. 261). The premise of the theory is that individuals limited working memory capacity, their capacity to temporarily hold and process information, can constrain cognitive processing (O. Chen & Kalyuga, 2020; Sweller et al., 2019). Cognitive load is a cognitive reaction that is increased when demands are placed on the cognitive system. When this load becomes too high, it can hinder an individual's capacity to learn and their motivation to engage in similar situations in the future (O. Chen & Kalyuga, 2020; Paas et al., 2004; Paas, Renkl, et al., 2003; Sweller et al., 2019). This load can be increased through insufficient instructional methods and unnecessary distractions (Sweller et al., 2019). The goal of cognitive load theory is for innovative and effective instructional procedures to be generated to manage the load imposed on working memory and optimise information processing capacity (O. Chen & Kalyuga, 2020). This requires an awareness of the different types of cognitive load that can be experienced, and which type of cognitive load should be optimised to facilitate learning.

### *2.6.1. Types of cognitive load*

There are two types of cognitive load agreed upon in the literature; intrinsic and extraneous cognitive load (F. Kirschner et al., 2009; P. A. Kirschner et al., 2018). The existence of a third type of cognitive load, germane cognitive load, is contended in the literature base. As germane cognitive load is presented as an element of intrinsic cognitive load through the expansion of its definition, it is contended that it is not a separate type of cognitive load (Kalyuga, 2011; Kalyuga & Singh, 2016; Leppink, 2017; Leppink et al., 2014; Sweller, 2010; Sweller et al., 2011). Germane cognitive load will be considered as a component of intrinsic cognitive load through this thesis. Intrinsic cognitive load refers to the interaction between the information to be learned and the expertise of the learner (F. Kirschner et al., 2009; P. A. Kirschner et al., 2018; Paas et al., 2004; Paas, Renkl, et al., 2003; Paas, Tuovinen, et al., 2003). The load is intrinsic if it is caused by the number of information elements in the assigned task and the interactivity between the information elements (F. Kirschner et al., 2009). Different tasks have various levels of interactivity and therefore have varying capacity to cause intrinsic cognitive load (Paas, Renkl, et al., 2003). To manage the effects of intrinsic cognitive load, tasks of increasing complexity may be used to 'build-up' the interactivity of the task (Paas, Renkl, et al., 2003). Germane cognitive load is also classified as effective cognitive load that is beneficial to the learning experience (F. Kirschner et al., 2009; P. A. Kirschner et al., 2018; Paas et al., 2004; Paas, Renkl, et al., 2003; Paas, Tuovinen, et al., 2003; Sweller et al., 2011; van Gog & Paas, 2008). Germane cognitive load is caused by information and activities that support the learning experience (F. Kirschner et al., 2009). This type of load can be managed through the design of the task (P. A. Kirschner et al., 2018; Paas, Renkl, et al., 2003). Extraneous cognitive load is also referred to as ineffective cognitive load (Paas et al., 2004; Paas, Renkl, et al., 2003; Sweller et al., 2011) and is caused by information and activities that do not support the learning experience (F. Kirschner et al., 2009; Paas et al., 2004; Sweller et al., 2011; van Gog & Paas, 2008). As extraneous cognitive load is caused by elements that do not support the learning approach, the goal is to manage it through the instructional approaches implemented for a task (P. A. Kirschner et al., 2018). The varying types of cognitive load which can be experienced by a learner can also have various effects on the learning experience.

### *2.6.2. Effects of cognitive load on learning*

Human cognitive architecture can broadly be viewed as having three core elements; amassing information, acquiring information, and interacting with the external environment (Sweller et al., 2011). Within the three core elements of human cognitive architecture, five principles of natural information processing systems are described; the information store principle, the borrowing and reorganising principle, the randomness as genesis principle, the narrow limits of change principle, and the environmental organising and linking principle (Sweller et al., 2011). These principles are described by Sweller, et al. (2011) as:

1. The information store principle- large amounts of organised knowledge can be stored in long-term memory and the information stored here allows humans to engage in activities automatically by recognition.
2. The borrowing and reorganising principle- the majority of the secondary knowledge that is stored in long-term memory is borrowed from others. This information can be reorganised for storage in long-term memory and transmitted to others.
3. The randomness as genesis principle- focuses on how information is initially created and how random generate and tests procedures are used as the source of novel information.
4. The narrow limits of change principle- is the collective term used to describe the characteristics of a natural information processing system in order to acquire information from the environment.
5. The environmental organising and linking principle- is the collective term used to describe the characteristics necessary for a system to perform appropriately within an environment. The final two principles provide the connection between natural information processing and systems and the environment they are in.

The interaction of an individual with the external environment forms the basis of cognitive load theory. Cognitive load theory focuses on individuals learning complex cognitive tasks and overcoming the cognitive strain experienced for meaningful learning to take place (Paas, Renkl, et al., 2003; Sweller, 1988). The effects of cognitive load on individuals and groups of individuals have been examined throughout the literature (F. Kirschner et al., 2009; P. A. Kirschner et al., 2018; Paas, Renkl, et al., 2003; Paas, Tuovinen, et al., 2003; Sweller, 1988; Sweller et al., 2011). Cognitive load can have several effects such as split-attention, redundancy, expertise reversal, and element interactivity with Sweller, et al. (2011) providing the following descriptions of these effects. The split-attention effect occurs when an individual's attention must split between at least two sources of information which are separated spatially or temporally. Redundancy effects are caused when several sources of information can be understood separately without mental integration being required. With the expertise reversal effect, numerous sources of information may be required for novices to support understanding while the same information may not be required for expert understanding. Element interactivity effect is caused by the level of interactivity between essential elements of information for understanding. To manage the effects of cognitive load on learning and ascertain whether it is being optimised, rather than hindering a learner's capacity to engage with a problem, requires a mechanism to suitably measure cognitive load.

### 2.6.3. Measuring cognitive load

Measures of cognitive load can be classified through two dimensions, causal relation (direct or indirect) and objectivity (subjective or objective) (Brünken et al., 2003). Subjective measures of cognitive load include self-reported invested mental effort (indirect) and self-reported difficulty of materials (direct). Difficulty is a direct measure of cognitive load given the causal relation between the measure and the phenomenon under investigation. There is a direct link between the difficulty of learning materials and the cognitive load experienced because difficulty is a direct result of the extraneous and intrinsic cognitive load imposed by the learning material (Brünken et al., 2003). Whereas mental effort is an indirect subjective measure as there is an indirect causal relationship between the measure and cognitive load. Objective measures of cognitive load include psychophysiological measures such as pupillometry or Electrodermal activity (EDA) which have an indirect causal relation with cognitive load, and brain activity and dual-task performance which have a direct causal relation with cognitive load. Table 3 outlines the classification of these approaches to measuring cognitive load.

*Table 3. The classification of approaches to measuring cognitive load with examples.*

| | **Causal Relationship** | |
|---|---|---|
| **Objectivity** | **Direct** | **Indirect** |
| Objective | Brain activity<br>Dual-task performance | Pupillometry<br>Electrodermal activity<br>Behavioural measures |
| Subjective | Self-reported difficulty level | Self-reported invested mental effort |

#### 2.6.3.1. Subjective measures

Subjective rating scales can be used to determine an individual's level of agreement with a statement or intensity of a feeling or emotion in response to an event (Cohen et al., 2018). In cognitive load measurement, there are two underlying assumptions in the use of subjective measures. It is assumed that individuals have an understanding of the terms “invested mental effort” and “difficulty of a task” and assumed that individuals have the metacognitive ability to evaluate how much mental effort they have invested (Ouwehand et al., 2021). The most associated subjective measures of cognitive load are self-reporting Likert scales and semantic differential scales (Kalyuga & Plass, 2017; Leppink et al., 2013; Leppink & van Merrienboer, 2015; Ouwehand et al., 2021; Paas, Renkl, et al., 2003; van Merriënboer & Sweller, 2005). The number of points on these self-reporting scales can vary. In certain approaches, seven-point scales such as the NASA Task Load Index (NASA-TLX) are implemented whilst others employ nine-point rating scales in cognitive load measurement (Leppink et al., 2014). NASA-TLX was developed with the goal of providing a sensitive summary of variations of workload (Hart & Staveland, 1988). The rationale and process of development of the scale are documented by Hart and Staveland (Hart & Staveland, 1988). NASA-TLX is a direct subjective measure of cognitive load (Kalyuga & Plass, 2017). Paas (1992) developed and validated a nine-point Likert-type item to evaluate the mental effort experienced by an individual as they performed a task. The numbers on the scale were assigned labels ranging from (1) very, very low mental effort to (9) very, very high mental effort. The difference between the two scales, apart from the number of points, is that the scale developed by Paas (1992) is a single item that solely measures mental effort. The NASA-TLX has multiple items that measure various factors which contribute to workload such as mental demand, effort, and frustration (Hart & Staveland, 1988). A Visual Analogue Scale (VAS) can also be used to measure cognitive load and presents numbers on a line continuum where individuals can move a bar between 0 and 100% to indicate the level of cognitive load that they have experienced (Ouwehand et al., 2021). A VAS collects a continuous and interval-level measurement and has been demonstrated to have high test-retest reliability and small measurement error (Ouwehand et al., 2021). Each approach offers a valid and reliable subjective measure of cognitive load (Ouwehand et al., 2021; Paas, Tuovinen, et al., 2003). The scales can be administered multiple times throughout an activity or once at the end of a series of activities (Leppink & van Merrienboer, 2015). Research findings have indicated that a single retrospective measure yields a higher response than the average of the multiple measures after each activity (Leppink & van Merrienboer, 2015). Rating scales, however, have some limitations. There is no assumption of equal intervals between each of the ratings. There is a tendency on five-point and seven-point scales for individuals to avoid selecting extreme values on the scale, and there is no way of knowing if the individual completing it wished to add a comment on what was being investigated (Cohen et al., 2018). It is also noted that frequent administration throughout a learning experience can be intrusive (Kalyuga & Plass, 2017). However, some of these issues can be addressed through method design, e.g., using a nine-point scale or adapting the standard format to include a comment section.

Subjective verbal qualitative approaches can also support investigations of cognitive load by gaining understanding and insight of experiences from individual perspectives. These include interviews such as stimulated recall interviews (Beers et al., 2008) and concurrent verbal protocols such as think-aloud (Jenkinson, 2009). Each of these approaches have both strengths and weaknesses associated with their application. Stimulated recall interviews have been used in cognitive load research to gain an insight of thought processes and how they relate to different types of loads experienced (Beers et al., 2008). Cohen et al. (2018) and Creswell and Creswell (2018) discuss the advantages and disadvantages of interview approaches in detail. They outline that interview approaches, not specifically for the purpose of cognitive load measurement, can be used to gather information directly relating to research questions or objectives, to test or generate hypotheses, or in conjunction with other methods to examine and validate the methods or investigate the motivations and responses of individuals. However, some weaknesses associated with interview approaches exist such as interviewer bias or interviewees becoming uneasy with a line of questioning.

Think-aloud protocols provide a method for studying both behavioural and cognitive processes during problem solving (Prokop et al., 2020). These protocols have been used with implicit measures of cognitive load such as eye tracking to inform a detailed account of performance and performance parameters (Prokop et al., 2020). However, despite these protocols informing the approach being implemented in real-time, there are some weaknesses to their implementation such as slowing participants down, which in turn makes tasks take longer and changes the participant’s interaction with a task (Barkaoui, 2011; Prokop et al., 2020). In addition, when used simultaneously with eye-tracking measurement, an objective measure of cognitive load, it can lead to an increased number of fixations (Prokop et al., 2020).

#### *2.6.3.2. Objective measures*

As detailed in Table 3, behavioural measures and physiological responses are examples of objective measurement of cognitive load. Previous work has examined the utility of gesturing behaviours to minimize the load on memory during information processing by temporarily off-loading working memory resources normally required for internally maintaining information by physically maintaining some information through gesturing (e.g., finger counting or pointing) (Eielts et al., 2020; Paas & van Merriënboer, 2020; Pouw et al., 2014). Research into associations between behaviours and cognitive processing has not yet advanced beyond the role of the gesturing behaviour. Insights to behaviour can be gained subjectively through approaches such as think-aloud protocols (Prokop et al., 2020), or objectively using approaches such as observation. Observations provide capacity for the researcher to capture situations such as events and behaviours as they occur and afford strong face validity through capturing rich contextual information (Cohen et al., 2018; Creswell & Creswell, 2018). Limitations in their use include the researcher being seen as intrusive and there may be problems in gaining rapport with certain participants (Creswell & Creswell, 2018). Video recording can also be used as a means of observation to circumvent limitations in building rapport. In video observations, the collection of footage by an observer may be disruptive to the participant or affect responses. However, recording equipment can be discretely setup to minimize intrusiveness. Using video also affords capacity to observe behaviours retrospectively (Reid et al., 2020). Audio-visual recordings can be used to support additional data collection through interview techniques, such as video-stimulated recall interviews, to provide an in-depth understanding of events (Paskins et al., 2017).

The measurement of physiological responses as an indicator of cognitive load is based on the premise that changes in cognitive load are reflected by physiological variables (Paas, Tuovinen, et al., 2003). Various physiological responses have been used as objective, but indirect, measures to investigate cognitive load experienced by individuals throughout activities (Reid et al., 2020). These include eye-tracking, pupillometry, electroencephalography (EEG), heart rate (HR) and EDA (Antonenko et al., 2010; Brünken et al., 2003; Paas, Tuovinen, et al., 2003; Setz et al., 2010). In using physiological responses as measures of cognitive load, it is necessary to observe additional variables to triangulate the measurement to evaluate whether it can be interpreted as an indication of cognitive load (Kalyuga & Plass, 2017). Physiological measures afford an objective measure of cognitive load during the completion of a task. However, they often require the use of invasive technologies, which themselves have been criticized due to the potential negative impact that they can have on primary task performance and therefore the ecological validity of a study (Paas, Tuovinen, et al., 2003; van Merriënboer & Sweller, 2005). However, recent advances in technology have provided capacity to measure physiological responses such as HR and EDA unobtrusively, e.g., wearable wristbands (Cain & Lee, 2020; Poh et al., 2010; Posada-Quintero & Chon, 2020). This increases the viability of implicitly measuring cognitive load through engagement with a task, as significant movement restrictions are no longer a concern (Reid et al., 2020).

### 2.7. Summary of literature review

Throughout this section the epistemic position and existing pedagogical practices of engineering education have been explored. Tools and strategies for problem solving are outlined as a core epistemic practice of engineering for education (Cunningham & Kelly, 2017; Jonassen, 2015; Passow & Passow, 2017; Sheppard et al., 2008). Further, complex problem-solving skills are outlined as the most important skill needed in engineering for IR 4.0 (Kamaruzaman et al., 2019), while also being highlighted as one of top competencies required of engineering graduates by practicing engineers, undergraduate alumni, and engineering faculty (Passow & Passow, 2017). However, despite the relative importance of problem-skills in engineering and significant efforts to develop these skills through problem-orientated pedagogical approaches (Crawley et al., 2014; Edström & Kolmos, 2014), engineering employers are reporting a lack of proficiency in engineering graduates problem-solving skills (Nair et al., 2009; Valentine et al., 2017; Ward & Thiriet, 2010). Therefore, it is necessary that research is conducted to gain insight to the factors underpinning complex problem-solving performance to contribute towards addressing this skills gap.

As a result of the contribution of spatial ability to success in engineering education (Buckley, et al., 2018, 2017; Maeda, et al., 2013; Olkun, 2003; Schneider & McGrew, 2018; Sorby & Metz, 2016; Sorby, et al., 2014; Uttal & Cohen, 2012; Veurink & Sorby, 2012; Wai, et al, 2009) and its described role in problem solving from an intelligence perspective (Schneider & McGrew, 2018), it is conceivable that spatial ability level may contribute to the complex problem-solving process for individuals. Therefore, it is a pertinent factor to be considered in seeking to understand the factors which underpin problem-solving performance. Exploring problem solving from a cognitive perspective as a fundamental human cognitive process (Y. Wang & Chiew, 2010) also requires consideration of cognitive load which is associated with the management of cognitive resources during engagement with a task (Chen & Kalyuga, 2020; Sweller, et al., 2019). As a study of this nature has not previously been conducted in the field of engineering education, it is likely that exploring problem solving in this way will contribute to both a new dimension of understanding of the problem-solving process and towards informing understanding of the causal association between spatial ability and success in engineering education.

# 3. Methodology

### 3.1. Research paradigm

In educational research there are several research paradigms that can be subscribed to such as the pragmatist, post-positivist, interpretivist, and critical paradigms which are highlighted as main worldviews through the literature (Feilzer, 2010; Kivunja & Kuyini, 2017). Research in the critical paradigm investigates economic and social issues to address problems such as social oppression (Kivunja & Kuyini, 2017). Interpretivist research works to understand human experience and relies on qualitative or mixed data (Mackenzie & Knipe, 2006). Research in the post-positivist paradigm holds a deterministic philosophy through which causes determine outcomes or effects (Creswell & Creswell, 2018). Research in this paradigm focuses on numerically measuring observations and studying human behaviours to verify a theory. Finally, work in the pragmatist paradigm places focus on solving practical “real world” problems and accepts that there are various realities that are open to inquiry (Feilzer, 2010). Within this paradigm the emphasis is placed on the research problem and question and the use of all available approaches to understand the problem (Creswell & Creswell, 2018).

The work presented in this thesis sought to examine how the role of spatial ability in complex problem solving contributes to a broader causal theory of the relationship between spatial ability and success in engineering. Considering the research problem there is a necessity to reflect on the approaches which may be used to investigate the core variables to gain understanding of the research problem. For example, it was essential that in this research the focal variable of complex problem-solving capability is isolated. Therefore, consideration was required surrounding the type of problem that would facilitate this. While discipline-specific complex problem solving is a core component of engineering practice, employing a discipline-specific problem in this research would introduce confounding variables impeding the isolation of the focal variable of interest. Using a domain-general complex problem affords the capacity to control for confounding variables while isolating the focal variable. Therefore, in the context of this research, a domain-general complex problem is most suitable for exploring complex problem-solving capability and addressing the research questions. The process of determining the most suitable approaches to explore the core variables is documented in section 3.4. Given the complexity of the research problem and questions being addressed through this research, it was necessary to not only consider all available approaches to understanding the phenomenon under investigation but also to reflect on the various qualitative and quantitative methods which may be used to obtain a comprehensive insight to the potential influence of spatial ability in domain-general complex problem solving. Therefore, the research presented in this thesis subscribes to the pragmatist research paradigm. Creswell and Creswell (2018, 36–37) outline the key assumptions made in pragmatist research as:

- "Pragmatism is not committed to any one system of philosophy and reality. This applies to mixed methods research in that inquirers draw liberally from both quantitative and qualitative assumptions when they engage in their research.
- Individual researchers have a freedom of choice. In this way, researchers are free to choose the methods, techniques, and procedures of research that best meet their needs and purposes.
- Pragmatists do not see the world as an absolute unity. In a similar way, mixed methods researchers look to many approaches for collecting and analysing data rather than subscribing to only one way (e.g., quantitative or qualitative).
- Truth is what works at the time. It is not based in a duality between reality independent of the mind or within the mind. Thus, in mixed methods research, investigators use both quantitative and qualitative data because they work to provide the best understanding of a research problem.
- The pragmatist researchers look to the *what* and *how* to research based on the intended consequences—where they want to go with it. Mixed methods researchers need to establish a purpose for their mixing, a rationale for the reasons why quantitative and qualitative data need to be mixed in the first place.
- Pragmatists agree that research always occurs in social, historical, political, and other contexts. In this way, mixed methods studies may include a postmodern turn, a theoretical lens that is reflective of social justice and political aims.
- Pragmatists have believed in an external world independent of the mind as well as that lodged in the mind. But they believe that we need to stop asking questions about reality and the laws of nature.

- Thus, for the mixed methods researcher, pragmatism opens the door to multiple methods, different worldviews, and different assumptions, as well as different forms of data collection and analysis."

## 3.2. Methodological approach

In this research there are three primary variables being considered: spatial ability, general complex problem solving capacity, and cognitive load. In this work, problem-solving performance and the management of cognitive load during problem solving are theorised to be supported by an individual's spatial ability (Figure 2). It is necessary that a comprehensive approach is taken to examining these variables to understand the connections that are theorised to exist. The measurement of problem-solving performance, spatial ability and cognitive load are commonly associated with quantitative data collection. However, this research also theorises that individuals may use different behaviours to manage cognitive resources during problem solving. Therefore, there is a requirement to document and examine qualitative observations to test this hypothesis. To robustly examine the role of cognitive load in problem-solving performance from this perspective requires the use of a mixed method research design.

There are three core mixed method research designs: exploratory sequential, explanatory sequential, and convergent designs (Creswell & Creswell, 2018). These designs can also be used in more complex mixed methods designs e.g., mixed method, case study designs or mixed method experimental designs. Through an exploratory sequential design, qualitative data is initially collected in one phase and analysed. This is then followed by quantitative data collection and analysis to explore findings from insights gained from the qualitative analysis (Cohen et al., 2018; Creswell & Creswell, 2018). The explanatory sequential design is the inverse of the exploratory sequential approach. In an explanatory sequential design, quantitative data is initially collected and analysed with qualitative data collected in a following phase to explain the quantitative findings in more detail (Cohen et al., 2018; Creswell & Creswell, 2018).

As a component of this research seeks to gain insight to the role of behaviours in managing cognitive resources throughout a problem-solving experience, neither sequential approach would be appropriate as through these mixed method research designs the quantitative and qualitative data is collected separately. To gain an insight of the role of behaviours in cognitive resource management throughout problem solving would necessitate the simultaneous collection of qualitative observational data and quantitative cognitive load data to determine whether behaviours can be inferred as an indicator of cognitive resource management during problem solving. A convergent mixed method design would facilitate this investigation and the determination of whether there is an association between externalised behaviours and the management of cognitive resources throughout a problem-solving episode. Through a convergent mixed method design, qualitative and quantitative data are collected simultaneously and merged to provide a comprehensive analysis of the research problem (Creswell & Creswell, 2018). A convergent mixed methods design was employed in this thesis, and the full methodological design will be outlined in the next section.

## 3.3. Design

To address the aim of this thesis, to investigate the role of spatial ability in problem solving from a cognitive perspective to contribute towards the understanding of a causal association between spatial ability and success in engineering education, two studies with convergent mixed method designs were conducted. These studies will be referred to as the preliminary study and the main study throughout this thesis. The studies were developed to attend to the research question:

> How does spatial ability's role in problem solving contribute to a broader causal theory between spatial ability and success in engineering education?

### *3.3.1. Preliminary study*

In the preliminary study, the intention was to examine the capacity of a methodology consisting of a video protocol and indirect physiological and self-report cognitive load measures and gain insight to the behaviours demonstrated by individuals during a problem-solving episode (Reid et al., 2020). As previous research has not incorporated the observation of behaviours with physiological objective measures of cognitive load it was necessary that a pilot test was conducted to ensure that such a methodology could appropriately contribute to addressing the research aim and question. Additionally, further to piloting methods, this study would facilitate comprehensive investigation of the observable behaviours demonstrated by individuals while solving a domain general complex problem through inductive thematic coding. From the outcomes of previous research where gesturing has been demonstrated to be representative of internal cognitive processes (Eielts et al., 2020), it is conceivable that other observable behaviours may also contribute to this process and as with gesturing, relate to individuals with different levels of spatial ability. To investigate these theories and the capacity of the methodology to inform understanding of the role of spatial ability in problem solving, the preliminary study design facilitated the collection of qualitative data in relation to performance and behaviours whilst quantitative data was simultaneously collected relative to cognitive load and performance in one session. In a separate session, spatial tests were administered to measure spatial ability. The convergent mixed method design of the study is illustrated in Figure 8 below.

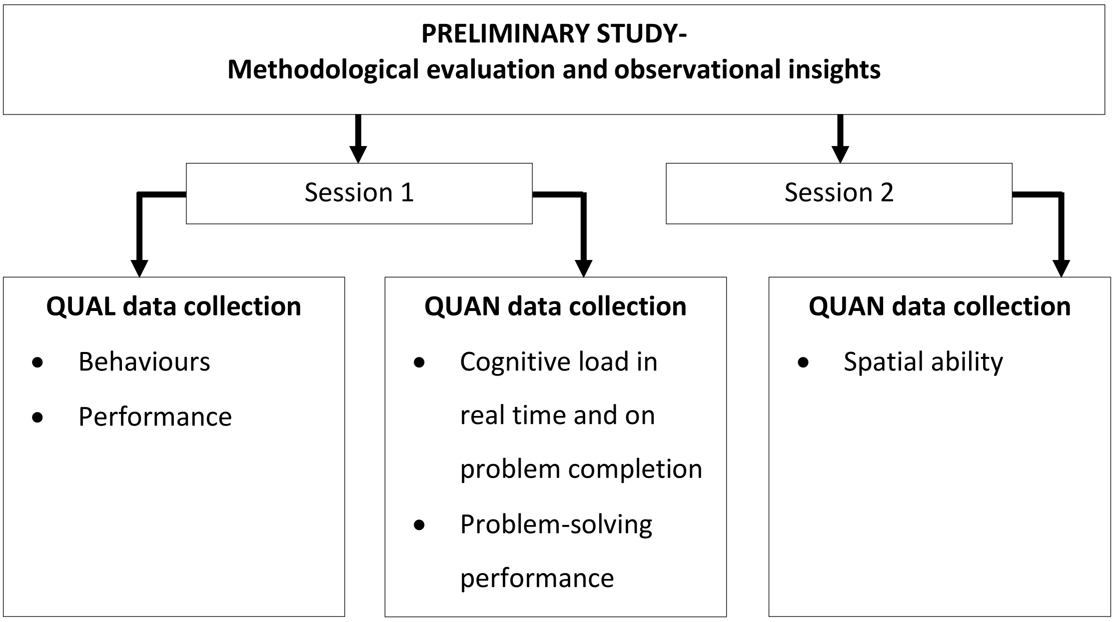


*Figure 8. Preliminary study convergent mixed method research design.*

### *3.3.2. Primary study*

From the insights gained in the preliminary study the methodological approach was advanced for the main research investigation, particularly in terms of the behaviours to be observed and measurement of cognitive load. Again, a convergent mixed method design was used (outlined in Figure 9) where in one session the qualitative observations were collected whilst simultaneously collecting quantitative performance data. Quantitative cognitive load data was also collected in this session. In a separate session, spatial tests were administered to obtain a measure of spatial ability. The specific approaches used to measure the constructs under investigation , which were treated as variables, are detailed in the following sub-sections.

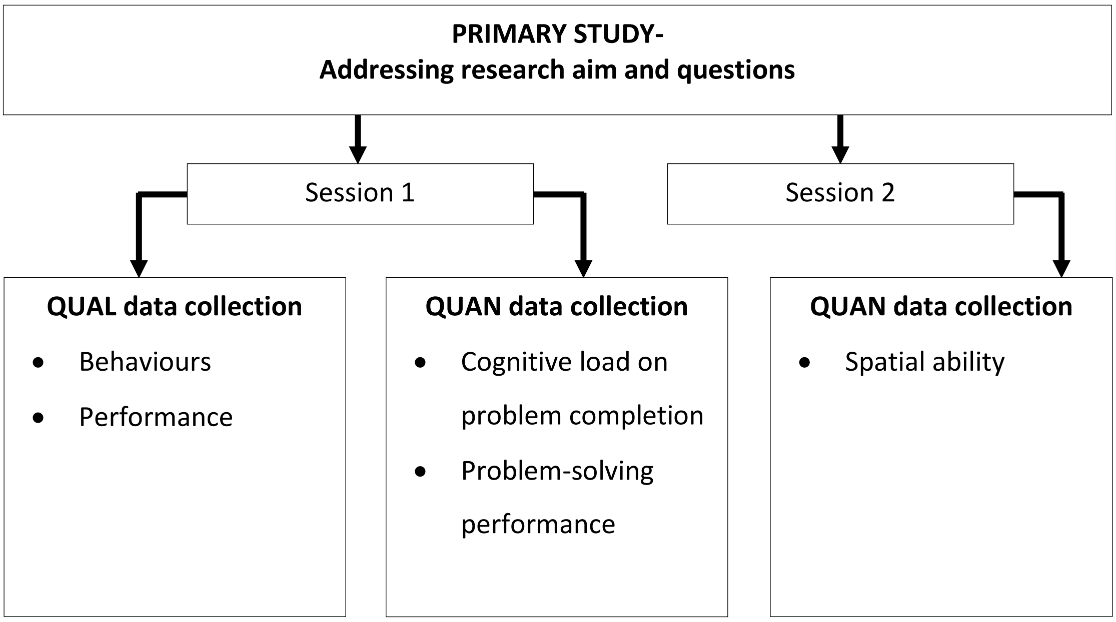


*Figure 9. Overview of the convergent mixed method design for the primary research study.*

## 3.4. Treatment and measurement of core constructs

### *3.4.1. Complex problem-solving performance*

Complex problem solving is described as reviewing information related to a problem to develop and evaluate options available to solve the problem and implement solutions (World Economic Forum, 2020). The characteristics of a successful problem solver include (Litzinger et al., 2010; Y. Wang & Chiew, 2010):

- the capacity to correctly identify problem goals
- persistence
- ability to adopt efficient search strategies
- regulation of actions and
- capacity to trace back to previous points in the solution process.

Given the complexity of problem solving, it is necessary to consider a measurement approach that can suitably and reliably measure these capacities. Commonly, complex problem-solving performance is associated with ill-defined open-ended problems such as design problems (Dörner & Funke, 2017; Funke, 2012; Jonassen, 2015; Murray et al., 2019). However, in seeking to measure a student's underlying capacity to solve complex problems the use of a discipline-specific complex problem, such as an engineering design problem, would not be suitable as the authentic and disciplined nature of such a problem would introduce confounding variables such as discipline knowledge making the isolation of the focal variable more difficult. Employing a domain-general complex problem would afford the capacity to control for confounding variables, such as discipline knowledge, whilst isolating the focal variable of complex problem-solving capability. Therefore, in the context of this research, a domain-general complex problem is most suitable for exploring complex problem-solving capability and addressing the research questions.

There has been little consensus on how to measure complex problem-solving performance on an individual level (Greiff & Funke, 2009). In previous work on complex problem solving in educational contexts, the computer based MicroDYN approach has been implemented (Lotz et al., 2016; Rudolph et al., 2018; Wüstenberg et al., 2012). The MicroDYN approach requires individuals to detect causal associations between elements of a system and in turn, control the system (Greiff & Funke, 2009). Previous research has supported the usability of this approach in terms of its internal consistency (Rudolph et al., 2018; Wüstenberg et al., 2012). This approach, or similar approaches, was not deemed appropriate for this research as a computerised problem reduces the possibility to explore the potential use of externalised behaviours by problem solvers as a method for cognitive resource management. Instead, a general physical problem was deemed appropriate.

The Tower of Hanoi (TOH) problem has been extensively used in studies examining complex problem solving, executive functioning, and gesturing to manage cognitive resources during problem solving (Beilock & Goldin-Meadow, 2010; Cooperrider et al., 2015; Eielts et al., 2020; Moreno & Guidetti, 2018; Pouw et al., 2016; Schiff & Vakil, 2015; Trofatter et al., 2015; Welsh et al., 1999; Welsh & Huizinga, 2005). Although, it has been criticised by some as being too simplistic, fully transparent, and static (Funke, 2010), others contend that the complexity of the problem lies within the identification and management of its sub-goals whereby the solver must recursively think ahead, consider the implications of their immediate and future actions, and have an awareness that counterintuitive moves are necessary to reach the goal state (Schiff & Vakil, 2015). This aligns with the adopted definition of complex problem solving in this thesis, where complexity relates to the identification and evaluation of solution pathways, and the aforementioned characteristics of successful problem solvers. The difficulty of the TOH problem can be increased for the user through the addition of more disks whereas the complexity can be increased through the addition of more pegs as more solution pathways would be present at any given time. The instructions for the TOH are simplistic and can be carefully controlled for identical instruction to each participant. This can theoretically normalise the amount of extraneous cognitive load experienced by participants due to the instructional design of the problem. Importantly, using the TOH to measure complex problem-solving performance negates the requirement of computers for the process.

In the context of this research, the TOH was deemed a suitable measure of complex problem-solving capacity as it requires the use of capacities outlined as key characteristics of successful problem solvers (Litzinger et al., 2010; Schiff & Vakil, 2015; Y. Wang & Chiew, 2010). A physical model of the TOH was used to maintain the possible affordances of physical external behaviours such as gesturing. The version adopted had three pegs to control for complexity and two variations of the problem were used, the three- and more difficult four-disk TOH. It is standard in studies that use the TOH as a measure of complex problem-solving performance to use both versions of the problem (Eielts et al., 2020; Moreno & Guidetti, 2018; Schiff & Vakil, 2015; Welsh & Huizinga, 2005). The premise of this is to facilitate a comparison of performance of individuals with an easier and more difficult version of the complex problem and explore the significance of the relationship between other variables (e.g., gesturing, working memory, planning) and performance on complex problems of varying difficulty levels (Eielts et al., 2020; Pouw et al., 2016; Schiff & Vakil, 2015). In using both problems consideration is required for the potential priming effects of the three-disk TOH for performance on the four-disk TOH where the participant may subsequently perform better on the four-disk TOH having first engaged with the three-disk TOH. For the three-disk TOH to serve as a prime for the four-disk TOH would require participants to be able to correctly identify the goals and sub-goals of the problem, the process of solving the problem and subsequently transfer this experience to the more difficult four-disk TOH. This would be representative of complex problem-solving capability. Where this transfer does not take place may be indicative that complex problem-solving capability has not been developed. As a component of this research is investigate whether complex problem-solving capacity is developed through engineering education the potential of the three-disk TOH to prime for the four-disk TOH would not have significant implications in the context of this research.

Figure 10 demonstrates the physical three-disk TOH problem. An additional smaller disk was added to this for the four-disk problem. Consistent with typical performance measures for the TOH, performance was evaluated in terms of the number of moves made and time taken to solve the problem. The three-disk TOH can be completed in a minimum of 7 moves and the four-disk TOH can be completed in a minimum of 15 moves. Participants are afforded unlimited time to solve the problem.

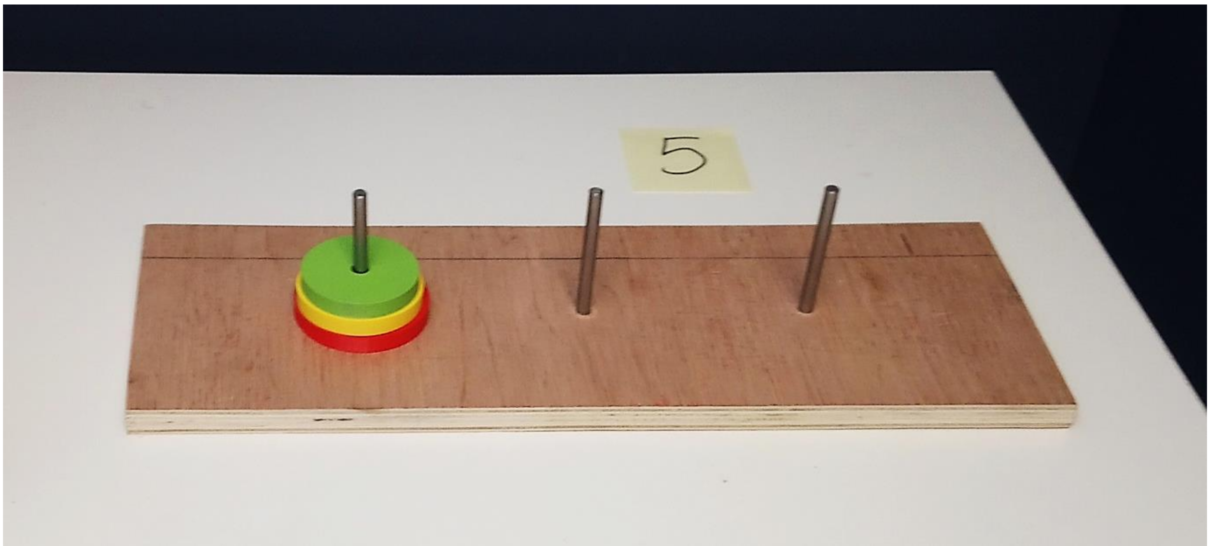


*Figure 10. A physical model of the three-disk TOH problem.*

### *3.4.2. Spatial ability*

Spatial ability is described as the ability of an individual to use simulated mental imagery to solve problems by perceiving, discrimination and manipulating images in the “mind’s eye” (Schneider & McGrew, 2018). The psychometric tradition of measuring cognitive ability will be used in this research as the predictive validity of these instruments has previously been demonstrated in research examining the relationship between spatial ability and engineering. Psychometric testing is used to measure psychological attributes of an individual such as intelligence and personality. Valid psychometric tests can be used to measure specific cognitive factors at a given point in time. Ekstrom et al. (1976) document 72 cognitive tests to measure 23 different cognitive factors in the Manual of Kit of Factor-Referenced Cognitive Tests. Psychometric tests are applied in educational settings to evaluate individual's performance and determine cognitive factors that relate to success (Kell & Lubinski, 2013; Wai et al., 2009).

Common accepted practice in the psychometric measurement of spatial ability is using tests which load on the narrow cognitive factor of visualisation (Schneider & McGrew, 2018). In almost all studies showing the predictive capacity of spatial ability for performance outcomes, visualisation is used as a proxy measure for the broad factor spatial ability (Schneider & McGrew, 2018). This is due to visualisation being the closest factor to spatial ability and strongest loading factor in analytic models of multiple spatial factors (Buckley, 2020; Carroll, 1993). To measure a narrow factor of cognitive ability, such as visualisation, requires the use of multiple tests specific to that factor as in isolation they are an imperfect measure of the factor (Schneider & McGrew, 2018). In the application of psychometric testing to measure spatial ability it is important to note that the number of tests used can result in fatigue or loss of motivation of the individual being tested (Buckley et al., 2018a). Therefore, if many tests are required or the tests being used are long, where possible the order of tests should be randomised to support a comprehensive analysis that considers test fatigue.

Psychometric measurement of spatial ability can be administered through paper-based or computer-based formats (S. K. T. Bailey et al., 2018; Friedman et al., 2020; Jamil et al., 2012; Leeson, 2006). As technology has evolved, computer-based psychometric measurements of various cognitive abilities have become increasingly common and this approach has been considered as a suitable replacement for more traditional paper-based approaches (S. K. T. Bailey et al., 2018). The benefits of computer-based measurement include automated scoring, and rapid administration (S. K. T. Bailey et al., 2018; Jamil et al., 2012; Leeson, 2006). However, there is not as much research evidencing the reliability of these computer-based approaches. Research examining the reliability of computer-based testing approaches has indicated that equivalent test taker performances are not obtained between them and paper and pencil approaches (S. K. T. Bailey et al., 2018; Leeson, 2006). Possible reasons for this include confounding variables such as computerised testing implicitly affording different strategies, and computer anxiety and experience (McDonald, 2002). In contrast, paper-based approaches for measuring spatial ability have been extensively used in research measuring spatial ability and there is a significant body of evidence of the validity and reliability of these approaches to accurately measure factors of spatial ability (Schneider & McGrew, 2018).

In the preliminary study in this thesis three paper-and-pencil psychometric measures of visualisation, the PSVT:R, SDT and PFT were used. The PSVT:R is a 30-item test where items gradually increase in difficulty through the requirement for more mental rotations in subsequent items, and individuals are given twenty minutes complete the test (Guay, 1976). It is commonly used in the measurement of spatial ability in engineering education research and is evidenced as a valid and reliable measure of spatial visualisation (Maeda et al., 2013; Sorby et al., 2006). A sample question from the PSVT:R can be seen in Figure 11A. To answer each of the 30 items on the test, individuals are instructed to study the shape in the top line of the question and how it is rotated. They are then asked to picture in their mind what the shape shown in the middle line of the question would look like when rotated in the same way. Five solution options are provided for the individual in the bottom line of the question and they are instructed to select the one that looks like the object rotated in the correct position. For the example in Figure 11A option B would be the correct solution.

The SDT and PFT are both valid and reliable measures of visualisation and are provided in the Kit of Factor-referenced Cognitive Tests (Ekstrom et al., 1976). The SDT is a two-part, 12-item test, where individuals are given six minutes to solve each part. An example of a question on the SDT is demonstrated in Figure 11B. With this test, individuals are instructed to try and imagine or visualise how a piece of paper may be folded to form some kind of object. The drawing on the left is a piece of paper that can be folded on the dashed lines to form the object that is drawn on the right. Individuals are asked to imagine this folding and figure out which of the lettered edges on the object are the same as the numbered edges of the paper on the left. The paper must always be folded so that the side marked with the ‘X’ will be on the outside of the object. The solution to the sample question in Figure 11B is detailed in the table on the right-hand side of the question.

The PFT is a two-part 20-item test where individuals are given three minutes to answer the questions in each part. A sample question on the PFT is detailed in Figure 11C. To solve the questions on the PFT individuals are instructed to imagine the folding and unfolding of pieces of paper. For each question there are some figures presented on the left side of a vertical line and five figures on the right side of the line. The figures on the left side represent a square piece of paper which is being folded, and the last figure on the left of the vertical line has one or two circles on it which represent where the folded piece of paper has been punched. Each of these holes are punched through all of the thicknesses of the paper at the respective points. One of the figures that is on the right of the vertical line shows where these holes will appear when the paper is completely unfolded. Individuals are asked to decide which of these figures is correct and to draw and 'X' through that figure to indicate their answer. For the example in Figure 11C option B is the correct solution. The order of test administration was randomised among participants in the preliminary study.

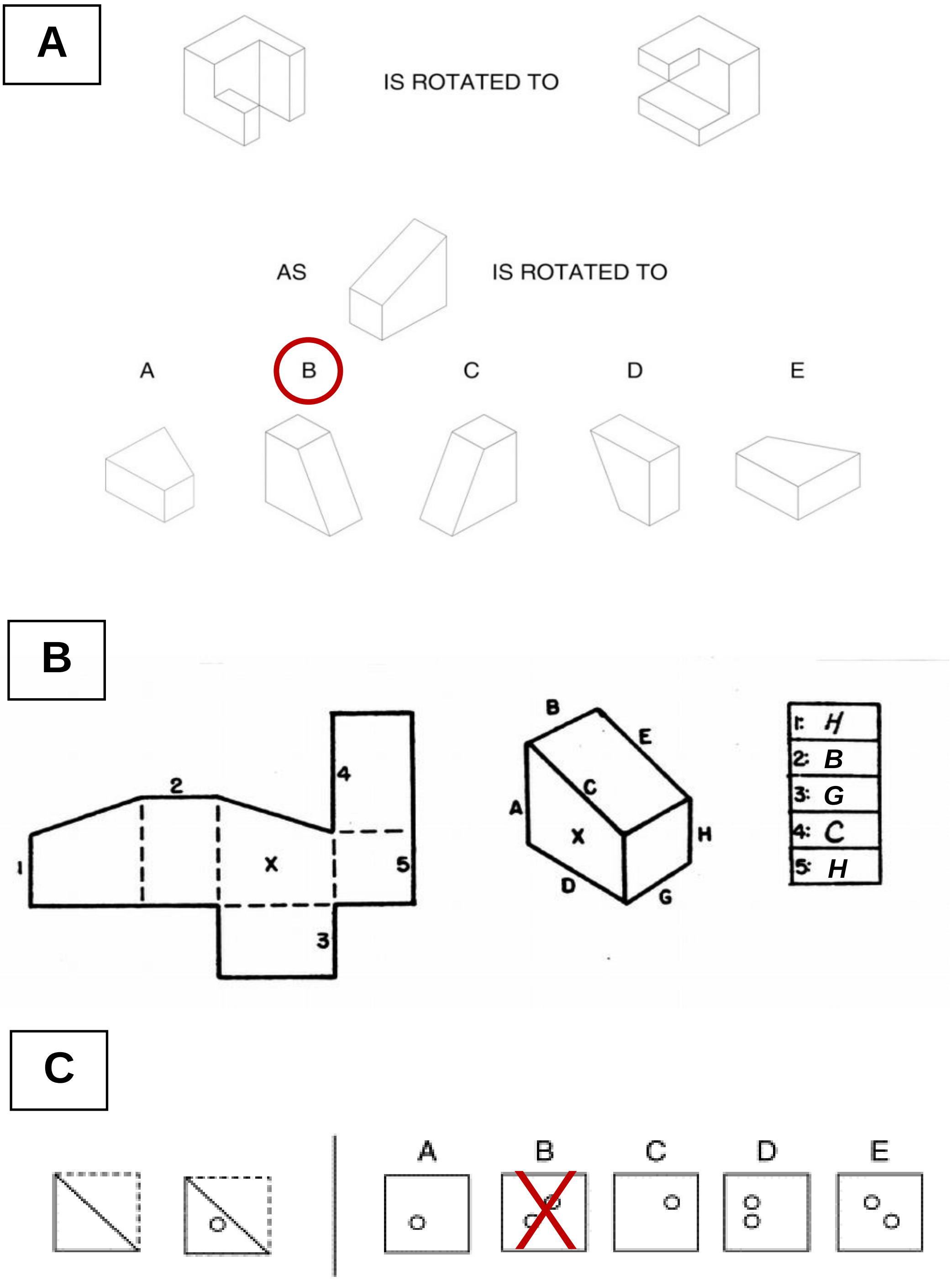


*Figure 11. Examples of questions on the spatial tests used in the research: A- PSVT:R, B- SDT and C-PFT.*

In the primary study, the MRT-A was used in place of the PSVT:R for n = 18 participants as they were taking part in another spatial ability research study. Previous research has demonstrated a significant relationship between performance on the two tests (*r* = 0.621, *p* <0.001) (Schmidt, et al., 2020). The treatment of these cases will be described in more detail in section 5.2.1. The MRT-A is a revised version of the original MRT developed by Vandenberg and Kuse (1978) (Peters et al., 1995). It is a two-part, 24-item test, where individuals are given three minutes to complete each part of the test and is demonstrated as a valid and reliable measure of visualisation. A sample question from the MRT-A is demonstrated in the figure below. For each question on the test individuals are instructed to look at a series of five figures. The figure on the left side represents and object that will be rotated into two different positions. To solve the question individuals must indicate which two of the four figures on the right side are the same object rotated into different positions. For the example included in Figure 12 figures one and three are the same object as that on the left and therefore the correct solution. The solutions should be indicated by placing a big 'X' across them. The SDT and PFT tests were also administered. As with the preliminary study, the order of test administration was randomised among participants.

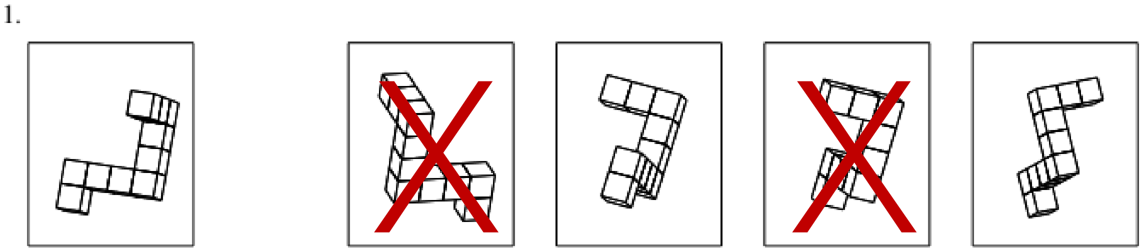


*Figure 12. Sample question on the MRT-A (Peters et al., 1995).*

### *3.4.3. Cognitive load*

Cognitive load is a cognitive reaction that is increased when demands are placed on an individuals limited working memory capacity, which is their capacity to temporarily hold and reason about information (O. Chen & Kalyuga, 2020; Sweller et al., 2019). The premise of cognitive load theory is that this processing load can affect a students' ability to process new information (Sweller et al., 2019). As earlier outlined, various subjective and objective measures can be used to investigate the cognitive load experienced by individuals when problem solving either directly or indirectly (Table 3).

#### *3.4.3.1. Subjective measurement*

Subjective measures of cognitive load can be either direct or indirect in terms of their causal relationship with cognitive load (Brünken et al., 2003). For instance, there is a direct link between cognitive load and difficulty of learning materials, whereas there is an indirect causal association between mental effort and cognitive load (ibid). Subjective measures of cognitive load commonly consist of self-reported rating scales (Paas, Renkl, et al., 2003; van Gog & Paas, 2008) and these approaches have been shown to be valid and reliable measures of the overall cognitive load experienced by an individual (Ayres, et al., 2021; Kalyuga & Plass, 2017; Leppink & van Merrienboer, 2015; Ouwehand, et al., 2021; Paas, 1992; Paas, Renkl, et al., 2003; Paas & van Merriënboer, 1994). Subjective measures of cognitive load can be taken repeatedly throughout a task or once at the end of a task providing a retrospective measure of the mental effort experienced (Leppink & van Merrienboer, 2015). However, if cognitive load is to be measured during an authentic experience and not in a controlled test environment, administering subjective scales throughout the activity would interrupt the authenticity of the experience and potentially influence performance factors being examined. Therefore, subjective measures of cognitive load should be taken at the end of an authentic problem-solving experience so as not to influence other variables under consideration. However, in doing so the cognitive load experienced throughout the problem-solving experience cannot be determined, therefore limiting the capacity to understand components which influenced cognitive load during problem solving.

#### *3.4.3.2. Objective measurements*

To overcome this limitation, objective measures of cognitive load can be used throughout the experience. Various physiological responses have been used as objective, but indirect, measures to investigate cognitive load experienced by individuals throughout activities. As earlier noted, these include pupilometry, eye-tracking, electroencephalography (EEG), heart rate (HR) measurements and electrodermal activity (EDA) measurements (Antonenko et al., 2010; Charles & Nixon, 2019; Keighrey et al., 2020; Setz et al., 2010; Sweller et al., 2011, Chapter 6). These approaches have been applied in various research settings such as in aviation (Charles & Dixon; De Rivecourt et al., 2008; Dussault et al., 2004; Lahtinen et al., 2007), medicine (Solhjoo et al., 2019), and education (Nourbakhsh et al., 2012; Thammasan et al., 2020; Villanueva et al., 2016, 2018, 2019). Physiological measures afford an objective measure of cognitive load during the completion of a task.

However, they often require the use of invasive technologies or restrictions of movement, which themselves have been criticized due to the potential negative impact that they can have on primary task performance and therefore the ecological validity of a study (Paas, Tuovinen, et al., 2003; van Merriënboer & Sweller, 2005). Recent advances in technology have provided capacity to measure physiological responses such as HR and EDA unobtrusively through, for example, wearable wristbands (Cain & Lee, 2020; Poh et al., 2010; Posada-Quintero & Chon, 2020). This increases the viability of implicitly measuring cognitive load through engagement with a problem, as significant movement restrictions are no longer a concern.

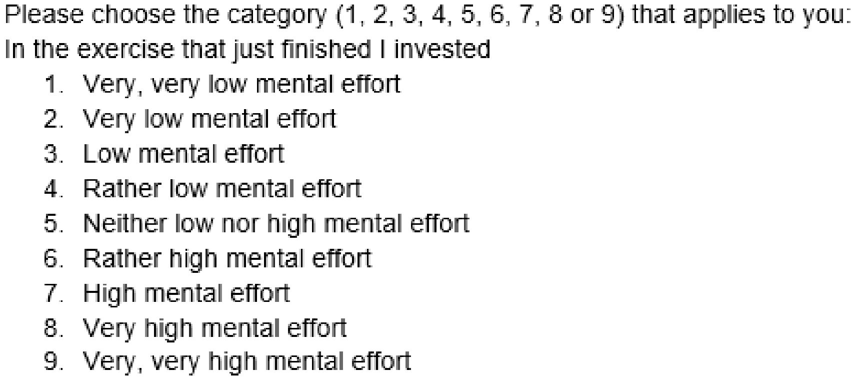
The most important factor to consider is whether these physiological measures can be validly interpreted to measure cognitive load. The validity of HR and HR variability to measure cognitive load is contested. Paas and van Merriënboer (Paas & van Merriënboer, 1994) detailed these measures as invalid and insensitive to slight fluctuations in cognitive load following a spectral-analysis technique of HR variability. Solhjoo et al. (2019), however, in conducting a correlation analysis between HR and HR variability and self-report measures of cognitive load reported a strong positive correlation between these indirect measures of cognitive load and HR variability. However, it is important to note the small sample size (n = 10) included in that study.

EDA, also referred to as galvanic skin response (GSR), relates to electrical changes that occur in the skin (Gjoreski et al., 2017; Henriques et al., 2013; Posada-Quintero & Chon, 2020; Setz et al., 2010; Son & Park, 2011; Thammasan et al., 2020). EDA measurement gauges psychophysiological activity of the sympathetic nervous system which is a part of the autonomic nervous system (Cain & Lee, 2020; Setz et al., 2010). Sweating is normally associated with thermoregulation of the body, however, in response to psychological stimuli, the body produces the physiological response of sweat through plantar and palmar sites (Posada-Quintero & Chon, 2020). This sweating causes an increase in the electrical conductance of the skin as part of the autonomic response (Boucsein et al., 2012), leading to EDA being employed for evaluating autonomic function and assessing levels of cognitive or emotional reaction to an arousing event (Gjoreski et al., 2017; Henriques et al., 2013; Posada-Quintero & Chon, 2020; Setz et al., 2010; Son & Park, 2011; Thammasan et al., 2020). EDA signals can vary between two categories, tonic change or phasic change (Boucsein et al., 2012; Braithwaite et al., 2013). The phasic component is referred to as skin conductance response (SCR) and is associated with short increases in EDA caused by arousing events such as sound, sight or smell (Braithwaite et al., 2013; Keighrey et al., 2020; Thammasan et al., 2020). The tonic component is referred to as skin conductance level (SCL) and is associated with slow change in skin conductance (Keighrey et al., 2020; Setz et al., 2010; Thammasan et al., 2020). These changes can be caused by an increase in cognitive activity (Keighrey et al., 2020; Setz et al., 2010). Measuring EDA using wrist-worn sensors provides an unobtrusive implicit indicator of these reactions experienced through engagement with a problem (Cain & Lee, 2020; Poh et al., 2010; Posada-Quintero & Chon, 2020). EDA sensors have been used in the implicit measurement of both cognitive and emotional reactions (Benedek & Kaernbach, 2010; Henriques et al., 2013; Liu & Du, 2018; Paletta et al., 2015; Poh et al., 2010; Setz et al., 2010; Villanueva et al., 2018). A limitation noted in the use of EDA sensors to measure cognitive reactions is the lack of capacity to mark specific events in the data and therefore align increases with interactions with a problem (Villanueva et al., 2018). However, this can be circumvented using additional software to mark specific timepoints of interest relative to the problem for alignment with the EDA data (Reid et al., 2020). EDA measurement affords the capacity to gain understanding of how an individual experiences an event without restricted movement where the authenticity of the activity is minimally affected.

The evaluation of behaviours during problem solving has previously been used to gain insight to the internal cognitive processes of individuals as they engage with a problem (Eielts et al., 2020; Pouw et al., 2014). Specifically, work with the TOH has evaluated the use of gesturing to free up cognitive resources during problem solving (Pouw et al., 2014). As such, simultaneously collecting behavioural observations through video recording devices while objectively measuring cognitive load would allow the capacity to gain further insight into the EDA data not previously achieved from a cognitive-behavioural perspective. It would also increase the richness of the data and support further explanation of variations in EDA and behaviours.

In summary, the selection of a method of cognitive load measurement should be based on the evaluation of its appropriateness to the subject under investigation (Flick, 2014). In the context of this work, to investigate cognitive load during problem solving, it was necessary that an individual’s movements are not limited as this would affect the authenticity of their engagement with the problem (Reid et al., 2020). From the perspective of objective measurement of cognitive load, EDA provides a suitable objective measure as it can be measured using unobtrusive wrist-worn physiological sensors and has previously been demonstrated as a measure of cognitive load (Cain & Lee, 2020; Poh et al., 2010; Posada-Quintero & Chon, 2020). Objectively observing behaviours exhibited during problem solving would also provide insight to the internal cognitive process that an individual is experiencing. Additionally, using the self-reporting subjective rating scale developed by Paas (1992) which measures the overall load experienced, would be suitable to both measure cognitive load and triangulate with EDA and behavioural data as the study seeks to examine cognitive load and not separate elements of workload which the NASA-TLX would provide (Hart & Staveland, 1988). The self-reporting scale developed by Paas does not differentiate between intrinsic and extraneous cognitive load. However, as previously noted in section 3.4.1, the instructions for the TOH problem are simplistic and can be carefully controlled for identical instruction to participants. This can normalise and limit the extraneous cognitive load that is experienced by participants and inferences can be made that variances in cognitive load experienced during problem solving are due to variances in intrinsic load experienced by participants.

In the preliminary study of this research, cognitive load was measured objectively throughout the problem-solving experience using a physiological sensor which collects EDA data. The lack of capacity in existing approaches to mark important events during problem solving was mitigated through the development of a bespoke time and movement mapping software. This software provided capacity to mark moves and the time that they occurred on the EDA data for later comparison with other data collected in relation to cognitive load. To facilitate this, it was critical that the bespoke software was run on the same device that the EDA device is remotely connected to so that the timestamps of important events can be synchronized with the EDA data using the operating system times. Cognitive load was also measured subjectively using the 9-point Likert-type item developed by Paas (1992) (see Figure 13) when participants concluded each problem. Both mechanisms for measuring cognitive load were used as EDA is a continuous measurement which allows for insight to be gained into the cognitive load experienced throughout the problem, while the Likert scale measure reflects the overall cognitive load that was experienced by the problem solver. The aim in combining these approaches was to gain a robust insight to the cognitive load experienced by individuals with different levels of spatial ability during complex problem-solving. In addition, audio and video recording equipment was used to observe behaviours exhibited during problem solving which may be indicative of cognitive resource management. This decision was reached as video observation does not interfere with the natural problem-solving process, affords the capacity to retrospectively analyse interactions, and has previously been successfully used in research examining the role of behaviours during problem solving, specifically gestures (Pouw et al., 2014).



Please choose the category (1, 2, 3, 4, 5, 6, 7, 8 or 9) that applies to you:
In the exercise that just finished I invested

1. Very, very low mental effort
2. Very low mental effort
3. Low mental effort
4. Rather low mental effort
5. Neither low nor high mental effort
6. Rather high mental effort
7. High mental effort
8. Very high mental effort
9. Very, very high mental effort

*Figure 13. Paas (1992) 9-point Likert-type item for measuring mental effort (overall cognitive load).*

In the primary study the aforementioned self-reporting 9-point Likert-type item is administered to obtain measures of cognitive load and audio and video recording equipment is used to collect behavioural observations.

### 3.5. Ethical considerations

Ethical approval was sought and granted for each of the studies included in this thesis from the research ethics committee at Athlone Institute of Technology at the time this research was being conducted, now Technological University of the Shannon: Midlands Midwest. Ethical considerations varied across studies due to the nature of information being collected from the various cohorts. Participation in the research was voluntary and participants were provided with details of the studies including aims and methods in line with institutional guidelines. Individuals that participated in the research were free to withdraw their participation at any time and information that withdrawers had provided would be removed from the study, therefore being excluded from data analysis and dissemination. The data collected was anonymised for the purposes of dissemination with participant numbers assigned. Participants in the preliminary study were a convenience sample of postgraduate students. Participants in the main study were a purposive sample of undergraduate engineering students at the initial and latter stages of their formal engineering training, a decision made as the variability sought was in assumed differences in levels of disciplinary expertise. In both studies two complex problems were administered to participants with performance monitored through audio and video recording equipment and a self-rated 9-point Likert-type item. In the initial study, physiological EDA sensors were used to collect objective indications of cognitive load throughout problem solving. Recording equipment focused on the hands of participants with the participant number visible to maintain participant anonymity. Three psychometric tests were used to measure the spatial ability of participants. Participants in both studies were fully aware of the purposes of the research and their role as participants. Participants were offered an incentive for participation and upon request were provided with an overview of their performance on an individual level. No potential participant identifiers have been or will be published as part of this research.

# 4. Preliminary study

## 4.1. Preliminary study method

### *4.1.1. Preliminary study aim*

As there are no existing methodological approaches to evaluating cognitive load continuously throughout problem solving whilst simultaneously monitoring behavioural interactions to determine potential associations, it was necessary to pilot an approach to determine its appropriateness and capacity to address the research questions. Simultaneously the empirical intention of the preliminary study was to expand understanding of the observable behaviours demonstrated by individuals during the TOH.

### *4.1.2. Preliminary study participants*

A convenience sample of postgraduate students was used for this study. A convenience sample was appropriate as the focal variable, observed behaviour exhibited during the TOH, was going to be observed until saturation, and there is no theoretical assumption that such behaviours will differ between populations. The study sample size was determined by reaching theoretical saturation. Participants were recruited through email and notice board advertisements with an incentive of being entered into a draw to win a €20 gift voucher. The investigator was known to the prospective participants.

The participants (*n* = 26) recruited were all above the age of 18 (21 - 48yrs). Participants were aware that they could withdraw their participation at any time without providing reason and that their participation was completely voluntary. Ethical approval was granted through the Athlone Institute of Technology research ethics committee. Participants were provided with an information sheet with details of the study and written informed consent obtained before the data collection began. All participant information was securely stored, and participant numbers assigned to protect participant anonymity.

### *4.1.3. Preliminary study implementation*

Individuals interested in participating were scheduled to attend two different sessions. Session one, the problem-solving session, and session two, a spatial ability testing session as demonstrated in Figure 8.

*Session one*: Participants completed two versions of the TOH problem, the three-disk and more difficult four-disk problem. Each session was attended by one participant only to minimize distractions which may influence variations in the participants EDA activity. When entering the session participants were provided with the information sheet and opportunity to ask any questions they might have. Participants then completed a written consent form.

The medical grade Empatica E4 physiological sensor was then fitted to participants in line with device specifications to obtain measures of EDA during the problem-solving process. As per the device recommendations, participants were fitted with the Empatica E4 on the non-dominant hand. The sensors for EDA were positioned in-line with the joint of the second and third finger at the wrist (Figure 14) to allow for an accurate and precise reading of the physiological response. Participants were instructed to relax for a period of five minutes to obtain baseline measurements of EDA. This baseline measurement is critical for identifying situations during the test that caused increases from an individual's "normal" EDA levels. During these five minutes the facilitator left the room as their presence may affect the individual's ability to relax fully.

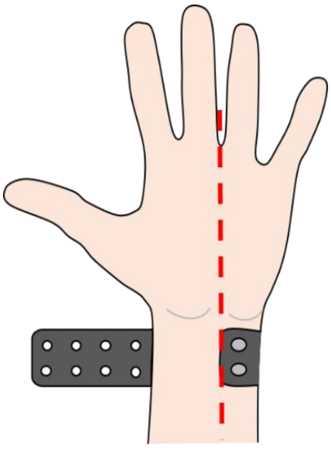


*Figure 14. The Empatica E4 physiological sensor placement on participants.*

After the five minutes had elapsed the investigator re-entered the session. To provide the opportunity for retrospective analysis and observation of the problem-solving session, the session was audio and video recorded. Participants were informed of this through their participant information sheet and the camera was orientated to only capture the participants hands as they solved the problem and participant number, which was taped to the table, as illustrated in Figure 15.

*Figure 15. Experiment setup with participant number hidden to preserve anonymity.*

Following the brief camera setup, the participant was presented with their first problem, the three-disk TOH and asked to indicate whether they had seen the problem before and if they subsequently knew how to solve it. These indications did not act as an element of selection criteria. The instructions were delivered to participants, and it was emphasized that questions could be asked at any point during the problem if clarification was required. The instructions were as follows:

‘The goal of the task is to get the arrangement of 3 discs on the left-most peg to the right most peg. The discs must be arranged in the same order i.e., largest on the bottom to smallest on the top. There are two conditions:

1. Only one disc can be moved at a time from one peg to another
2. A larger disc cannot be placed on a peg that already contains a smaller disc’

Participants were provided with the opportunity to ask questions and were then instructed to begin the problem.

In parallel to this, a bespoke software was utilized to objectively measure the number of moves made. A start point for the activity was incorporated into the software so that the time could be observed between participants being told to begin the problem and the first move being made. The software solution automated the capture of problem completion (time), interaction time, and number of moves taken to complete the problem. The objective of this was to create an accurate measure of user performance. Through incorporating this feature, the time taken between moves could be mapped onto the sensor data to monitor the physiological responses of participants between moves. Each move was marked by the investigator electronically with the final move concluding the time tracking for this problem. A move was defined as when the disc was placed on a peg and released by the participant. When the problem was completed, participants were presented with the 9-point Likert-type item and asked to rate the amount of mental effort they experienced when solving the problem as an indicator of cognitive load. Once completed, the second problem, the four-disk TOH, was administered to participants. The same process was repeated for the four-disk problem including the self-reported cognitive load experienced during the problem. Following the completion of the final problem and self-report scale the video recording equipment and physiological sensor were switched off.

*Session Two*: Spatial ability was measured in a separate session to reduce the effects of test fatigue that may be experienced through its inclusion in the problem-solving session. The session facilitator administered three spatial visualisation tests to participants, the PSVT:R, SDT, and PFT. The order of test administration was randomised between participants to account for order bias in the administration of the tests. Each test was administered in line with test specifications.

### *4.1.4. Preliminary study data analysis procedure*

#### *4.1.4.1. Quantitative analysis procedure*

The raw quantitative data for spatial ability, problem-solving performance, and self-reported cognitive load was initially compiled in Microsoft Excel before being cleaned and analysed using R Studio (R version 4.0.3.), and IBM SPSS (Statistics 27). Descriptive statistics were determined for all quantitative variables and a factor analysis was conducted on the spatial test scores. The objective cognitive load data was then prepared. Noise in the EDA data was cleaned with a 4Hz filter. A baseline EDA level was determined for each participant. The EDA data was then graphed, and the performance data was mapped onto these EDA graphs for each participant.

#### *4.1.4.2. Qualitative analysis procedure*

The observational video data collected during the problem-solving session was thematically analysed with consideration of the phases described by Braun and Clarke (2006). As an analysis of this nature had not previously been conducted on physical behaviours exhibited during problem solving, there were no preconceptions of codes or themes that would emerge from the problem-solving performance process. To that end, inductive thematic saturation was the intention in coding this data (Saunders et al., 2018). The process of coding is outlined in the following steps:

1. Initial viewing of footage for each participant to become familiar with the data
2. Further review of footage, making interpretative memos on sense of what was happening
3. Meeting to discuss and review samples of footage, interpretations, and sense of the data with no conception of codes or themes discussed
4. Initial inductive code generation begins. Footage reviewed, stopped, and replayed wherever it was interpreted that a behaviour which could potentially be a code was being exhibited. A note of the timestamp on the video footage at this point was recorded as an example of the potential code for later discussion with members of the supervisory team. Four videos reviewed at a time until inductive thematic saturation was reached (after reviewing twelve videos)
5. Behavioural codes were reviewed, discussed, and refined with the project supervisors. Similar codes were merged where suitable, and the names of codes were refined.
6. Following consensus on behavioural codes they were compared relative to problem solving and one another. Similar behaviour codes in the context of problem-solving performance and to one another were grouped. These groupings formed themes which were refined with consideration of the video data and interpretative memos noted at the beginning of the coding process. Factors from the memos not relating to the themes were noted for later discussion. The themes were also compared to ensure that each theme was internally homogeneous and externally heterogeneous (Braun & Clarke, 2006). The themes were then defined and named.

### *4.1.5. Merged analysis*

The video footage for each participant was reviewed again following the finalization of themes and the associated codes to identify the exact time points that the codes and themes occurred for each participant. This was carried out so that the occurrence of the coded behaviours and themes could be graphically mapped onto the physiological data to gain a visual overview of the EDA experienced at the time of these codes. As the EDA data records a timepoint in seconds with the inclusion of milliseconds, a free video software, VideoPad Video Editor, was used that displayed time in this manner to increase the time accuracy of the occurrence of the code on the EDA data. The steps involved in this process were:

1. Set the point where the participant was instructed to begin the problem as the start point of the video as this would also be the starting point of the graphs.
2. When a behaviour began to be exhibited the video footage would be reviewed frame-by-frame to determine the exact moment, in milliseconds, it began and exact moment it concluded.
3. This process was repeated for each code occurrence for all participants.

RStudio (R version 4.0.3.) was then used to visually map the codes onto the EDA data.

## 4.2. Preliminary study findings

Four participants were removed from the study due to issues with technical equipment. Six of the remaining participants (*n* = 22) indicated that they had seen the problem before. No participant indicated that they knew how to solve the problem. Consistent with a convergent mixed method design, the quantitative and qualitative data was analysed independently before being compared (Bailey & Gammage, 2020; Cohen et al., 2018; Creswell & Creswell, 2018; Creswell & Plano Clark, 2011; Hatta et al., 2020).

### *4.2.1. Quantitative analysis*

#### *4.2.1.1. Descriptive statistics*

Descriptive statistics were computed for performance on the TOH, self-reported mental effort, and performance on the three psychometric tests of spatial ability (Table 4). Performance on the TOH, an indicator of problem-solving capacity, was measured by the number of moves required to solve the problem and the time taken to reach a solution. The minimum number of moves required to solve the three-disk TOH is 7, while the minimum number of moves required to solve the four-disk TOH is 15. The mean number of moves on the three-disk TOH was found to be *M* = 8.23 which is close to the minimum moves to solve the problem, while on the four-disk TOH the mean value (*M* = 29.55) is almost double that of the minimum number of moves. The length of time to complete the problem also significantly increased on the four-disk

TOH. These findings indicate as the difficulty of the problem increases, the length of time and number of moves to solve it also increases, as expected.

*Table 4. Descriptive statistics for the preliminary study.*

| | M | SD | Med | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| | | | | **Spatial tests** | | | |
| PSVT:R | 59.24 | 23.92 | 61.67 | 20 | 93.33 | -0.21 | -1.38 |
| SDT | 60.23 | 24.38 | 62.50 | 11.67 | 96.67 | -0.38 | -0.92 |
| PFT | 48.86 | 19.57 | 47.50 | 15 | 90 | 0.33 | -0.16 |
| | | | | **Three-disk TOH** | | | |
| Moves | 8.23 | 2.29 | 7 | 7 | 14 | 1.77 | 1.71 |
| Time | 29.71 | 18.95 | 23.50 | 13.56 | 88.16 | 2.30 | 5.41 |
| Mental Effort | 3.36 | 1.65 | 4 | 1 | 6 | -0.16 | -1.03 |
| | | | | **Four-disk TOH** | | | |
| Moves | 29.55 | 12.11 | 26 | 16 | 63 | 1.06 | 1.08 |
| Time | 110.99 | 67.05 | 90.44 | 34.83 | 322.93 | 1.45 | 3.26 |
| Mental Effort | 5.41 | 1.47 | 6 | 2 | 7 | -0.80 | -0.08 |

M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

Participants completed the 9-point Likert-type item to indicate the overall cognitive load that they experienced (mental effort) as they solved the problem. As can be seen in Table 4, there was an increase in the mean measures between the three- and four-disk problems which would be expected in line with the increased difficulty of the problem.

Three spatial visualisation tests were administered to participants as a measure of spatial ability. The maximum score possible on the PSVT:R is 30, SDT is 60, and PFT is 20 (Ekstrom et al., 1976; Guay, 1976). Descriptive statistics of spatial test score percentages for the sample are included in Table 4. Based on the mean percentage values in the table it is demonstrated that participants performed to a marginally higher level on the SDT (*M* = 60.23) than on the PSVT:R (*M* = 59.24). Participants in the sample scored lowest on the PFT (*M* = 48.86).

#### *4.2.1.2. EDA data presentation*

Prior to the merged analysis, the physiological EDA data collected through the Empatica E4 sensor was cleaned with a 4Hz filter to remove any signal noise from the recordings. A baseline EDA level was determined and graphed with the EDA data for each participant's problem-solving experience. This was to provide a comparison of the individuals EDA at rest and during problem solving. In Figure 16, the horizontal dashed line across the graph represents the individuals baseline EDA. The data from the bespoke time and movement tracking software was then mapped onto this data, as illustrated in Figure 16. The first point on the graph represents when participants were instructed to begin the problem with each subsequent point indicating the next move, time it occurred, and EDA at that moment for the participant.

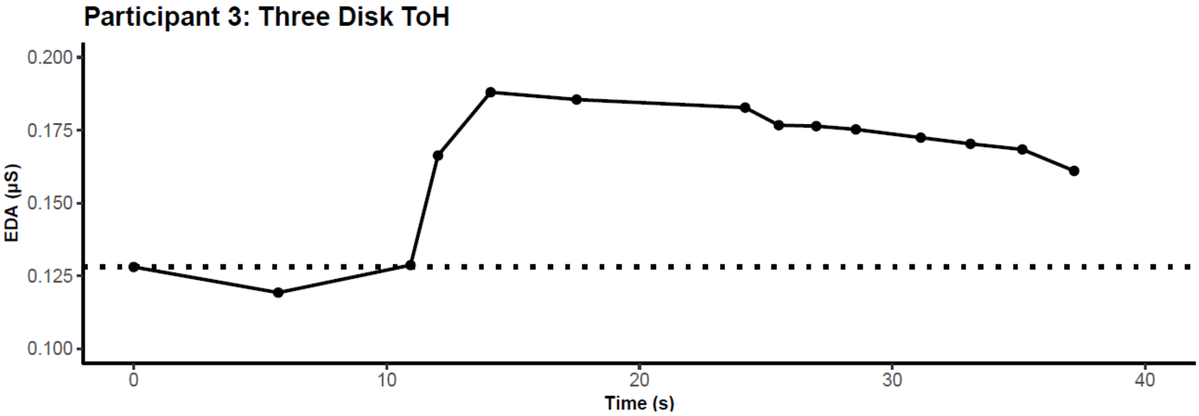


*Figure 16. Graphical representation of EDA, moves, and time where baseline EDA is represented by a horizontal dashed line and moves and the time that they occurred are represented by solid black dots.*

#### *4.2.1.1. Computing a single spatial ability variable*

In line with best practice recommendations for the measurement of single cognitive factors, three measures of the spatial ability, specifically for the visualisation factor were used (Buckley, 2020; Schneider & McGrew, 2018). To determine whether these tests did predominantly measure a single factor and could be transformed into a single variable as a more robust indicator that the three tests taken individually, a factor analysis was conducted using IBM SPSS. The results, presented in Table 5, indicated a single underlying latent variable which explained 82.92% of the variance. As this value is above 60% this is deemed satisfactory to interpret the tests as loading on a single factor (Hair Jr. et al., 2014).

*Table 5. Preliminary study spatial tests factor analysis for a single underlying latent variable.*

| Total Variance Explained | | | |
|---|---|---|---|
| | | Initial Eigenvalues | |
| Factor | Total | % of Variance | Cumulative % |
| 1 | 2.488 | 82.920 | 82.920 |
| 2 | .352 | 11.729 | 94.649 |
| 3 | .161 | 5.351 | 100.000 |

Extraction Method: Maximum Likelihood.

To compute a single spatial ability variable, all three variables (one for each individual spatial test) were first transformed into z-scores using the formula detailed in equation 1 below to standardise their results in terms of distance in standard deviations from the mean. A composite score was taken as the average of the three z-scores. This is similar to the procedure implemented by Hambrick et al. (2012) to investigate spatial ability.

$$z = \frac{x - \bar{x}}{\sigma} \tag{1}$$

The descriptive statistics for this single spatial variable are demonstrated in Table 6. As the sample values have been converted to z-scores the mean value will be 0 (Abdi, 2007).

*Table 6. Descriptive statistics for the preliminary study single spatial variable.*

| | M | SD | Med | Min | Max | Skew | Kurt |
|---|---|---|---|---|---|---|---|
| Spatial ability | 0.00 | 0.91 | 0.04 | -1.74 | 1.54 | -0.14 | -0.79 |

M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

### *4.2.2. Qualitative analysis*

Following initial inductive coding of the video footage, inductive thematic saturation was reached after reviewing three groups (n = 4) of footage. Figure 17 indicates the number of codes generated per group, with 12 codes being generated in total. The number of codes was later reduced to nine following review and discussion of the codes. Codes that did not relate to problem-solving performance itself were removed at this time and noted for later discussion.

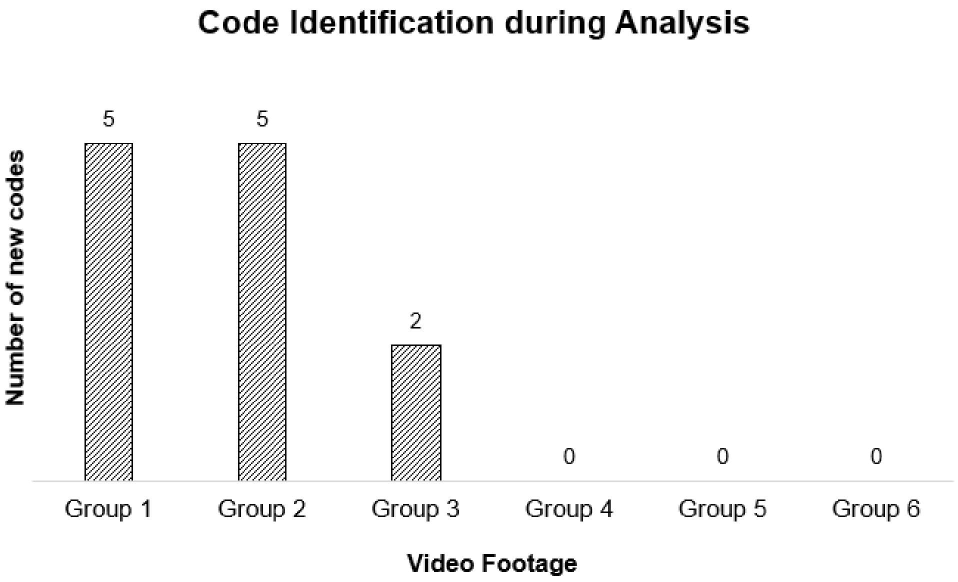


*Figure 17. The emergence of behavioural codes through inductive thematic coding where saturation was reached after reviewing three groups of footage (n = 4).*

An inductive thematic analysis of these coded behaviours resulted in three main themes in the context of problem-solving performance on the TOH; "stop", "indecision", and "progress". The themes and codes within these themes are outlined in Table 7.

*Table 7. Inductive thematically coded behaviours and associated themes.*

| **Theme/Code** | **Description** |
|---|---|
| "Stop" | |
| Pause | Stops, no movements being made to progress. |
| Finger movements | Tapping, drumming, or wiggling fingers while not making a move. |
| Disk play | Disk bounced up and down, spun, or moved around on the peg. |
| "Indecision" | |
| Hesitation | Hesitation to pick up or to remove a disk from a peg to make the next move. |
| Reach and stop | Reaches to make a move, brings hand back and stops. |
| Gesturing | Using finger to point at the position of a disk and gesture to where it might be moved. |
| Rule confirmation | Facilitator asked a question in relation to the rules/instructions for the problem. |
| "Progress" | |
| Hovering | Disk held or slowly lowered over new position, before being placed/released. |
| Trialling | Disk in hand but unsure about which peg to place it on, trialling positions by moving above the pegs. |

Examples of some codes are demonstrated in Figure 18 and video examples of all codes, with the exception of rule confirmation to protect participant anonymity, can be found at: https://osf.io/pf3kw/?view_only=d0bafdd65aa248b182e1071186782da7

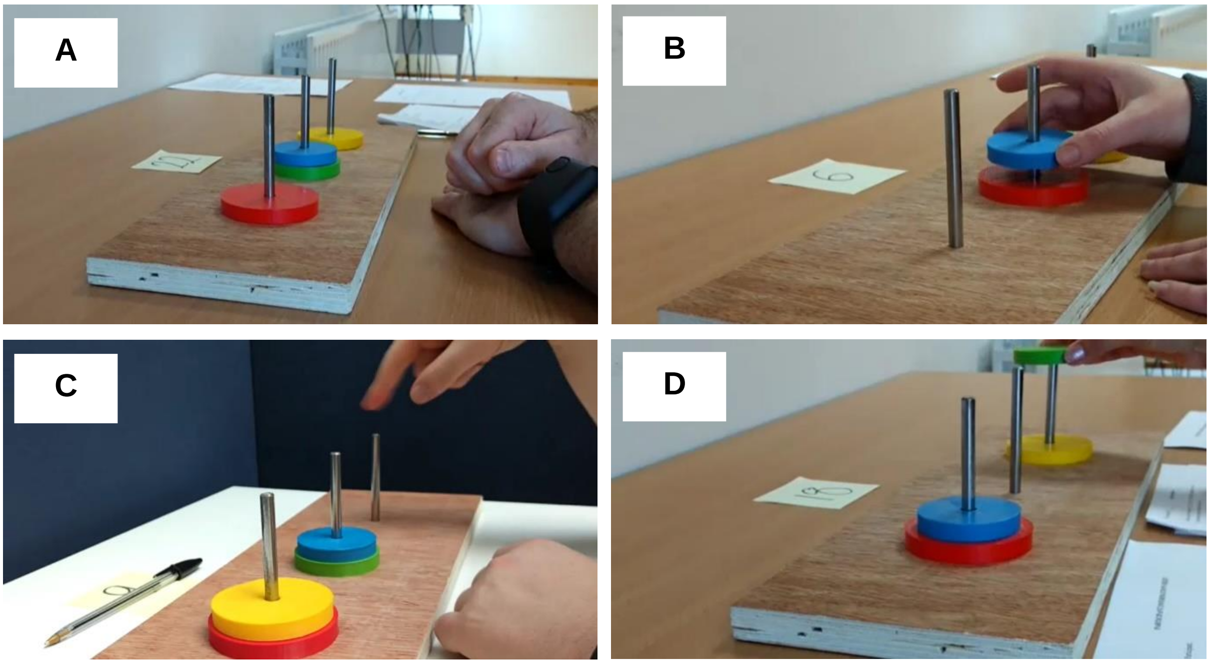


*Figure 18. Examples of coded behaviours in the "stop" (A- pause), "indecision" (B- hesitation and C- gesturing), and "progress" (D- hovering) themes.*

#### *4.2.2.1. "Stop" theme*

When codes categorised in the "stop" theme were exhibited the individual was not making any physical interaction with the problem. In generating the interpretive memos at the early stages of data analysis, it was noted that in these stopping times the individuals appeared to be reflecting on their approach and how to proceed. It was also noted that they appeared to sit back and observe what was in front of them, demonstrated in Figure 18A.

#### *4.2.2.2. "Indecision" theme*

Codes categorised in the "Indecision" theme were those where an action was being made to solve the problem, but the individual was not actively holding or moving any of the disks. Figure 18B and C demonstrate examples of behaviour codes which were categorised within the "indecision" theme. The rule confirmation code was also a behaviour in this theme with examples of rule confirmations including:

> PT 06: "You can't move two at once can't you not?"
>
> PT 14: "Can I go backwards with them as well?"

The interpretive memos from the investigator noted that at these stages the individual seemed uncertain of what they should do next.

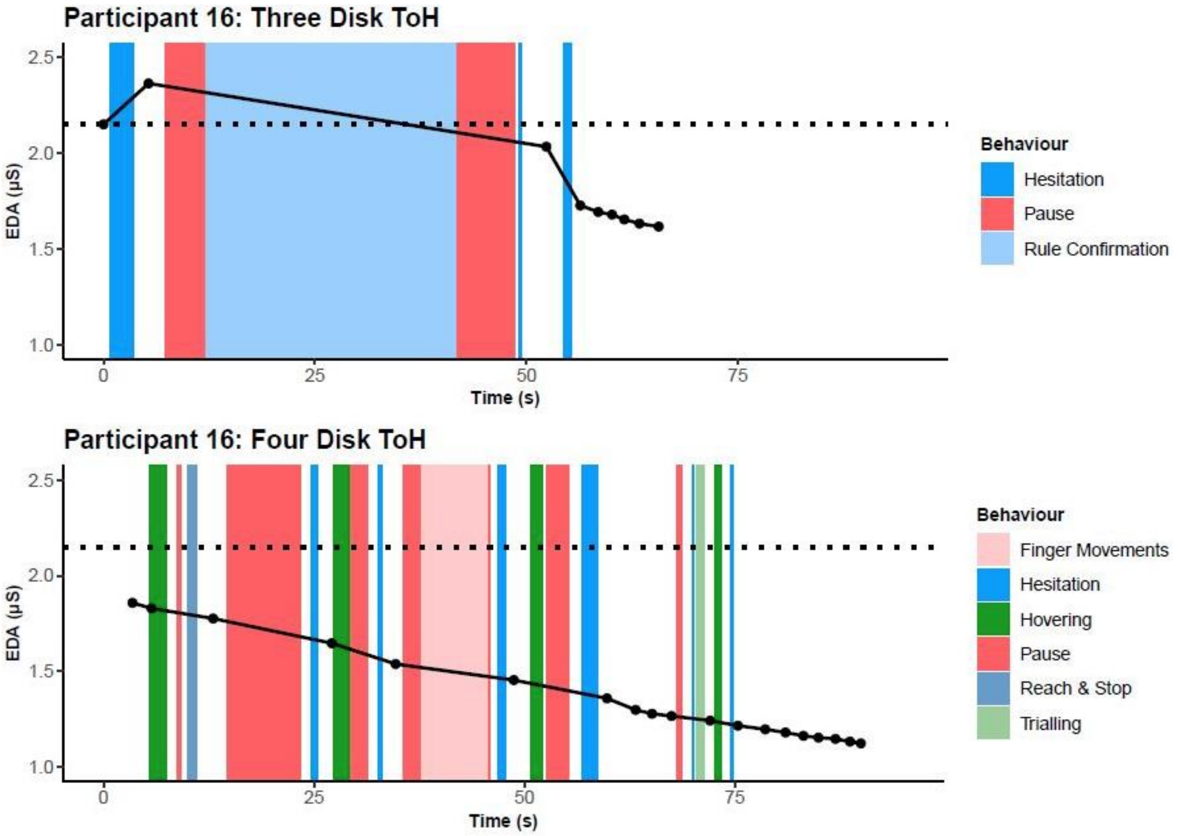
#### *4.2.2.3. “Progress” theme*

Behaviours categorised under the “progress” theme were those where the individual was exhibiting an action towards making a move, but the move had not been completed. In these instances, individuals had a disk in their hand and off a peg attempting to move forward with the problem, as demonstrated in Figure 18D which depicts the hovering code. When exhibiting this code, individuals had removed a disk from a peg and brought it to a new position where the disk was subsequently held over this position for a period or slowly lowered down the peg before being released to complete the move. From the interpretive memos it was noted that when individuals exhibited behaviours in the “progress” theme they displayed a degree of uncertainty around which position the disk should be placed in.

#### *4.2.2.4. Additional observations*

As previously noted, there were observations made in the interpretive memos throughout the analysis of video observations which did not directly relate to the participants physical behaviours towards the problem and therefore were not included as codes. Several participants made audible exclamations throughout problem solving regardless of whether they were exhibiting codes in the “stop”, “indecision”, or “progress” themes. Some acknowledged awareness of how they should go about solving the problem in saying phrases such as; “*something like this”*. Others made statements acknowledging that they had made a mistake in their view; “*I messed up*” or “*I’ve made this a lot harder on myself*”. Other participants queried whether it was possible to solve the problem with one participant asking; “*can it be done?*” and following this, gasped when they appeared to realize how to solve it. These audible exclamations occurred most frequently during the more difficult four-disk TOH and were particularly interesting as none of the participants indicated that they knew how to solve the problem. The audible occurrences indicated that despite not initially knowing how to solve the problem participants reached a point of awareness of the solution. At this point, it could be seen as transitioning from being a problem to a task for the problem solver (Buckley et al., 2018b; Schoenfeld, 1983). With the audible occurrences indicating a point in the four-disk TOH where participants realized how to reach the solution, it is necessary to examine the EDA graphs with mapped behavioural codes to determine if this indication may be reflected there.

To determine the capacity of the method to inform understanding of a cognitive load-behaviour interaction during problem solving, and to examine a potential point where a solution path was identified, it was necessary to create a joint display of the qualitative and quantitative data to facilitate the inferences that can be made (Fetters et al., 2013; Fetters & Freshwater, 2015; Guetterman et al., 2015). The exhibited problem-solving behaviours were mapped onto each participants EDA, move and time graph for each of the two problems. Figure 19 illustrates how the joint display appeared in the format of graphs for each participant. The horizontal dashed line represents the individuals baseline EDA for comparison to EDA measures during problem solving. The first point on the graph represents when participants were instructed to begin the problem with each subsequent point indicating the next move, time it occurred, and EDA at that moment for the participant. The vertical lines represent the behaviours the individual exhibited with behaviours in shades of blue representing “indecision”, green representing “progress”, and red representing “stop” themes of coded behaviours. Joint display graphs for all participants are displayed in Appendix I.


Participant 16: Three Disk ToH

Participant 16: Four Disk ToH

*Figure 19. Participant 16 joint display of behaviours, EDA, and problem-solving performance.*

The joint display graphs were qualitatively assessed to explore whether any association existed between behaviours or themes and increases or decreases in EDA signal. Through comparisons of the quantitative EDA data and qualitative observational data in the joint display graphs, no obvious influence was observed of behaviours or themes on variations in EDA measurement. This insight does not definitively determine whether an association exists between cognitive load and behaviours exhibited during problem solving. Rather, in the current study design, increases or decreases of EDA cannot be attributed to particular themes. Determining a statistical association between behaviours and self-reported cognitive load was not possible given the sample size included in this study. Such an association may be tested in a study with a larger sample size where statistical analysis can be conducted.

A possible association was identified between the cessation of behaviours and the number of moves remaining to solve the four-disk TOH, approximately seven. This indicated that at this point, the participants could see a clear path to the solution. The appearance of the problem with seven remaining moves was explored. Figure 20 illustrates the arrangement of disks at this stage of the problem. The bottom right corner represents the solution, with all disks arranged in order from smallest to largest on the final peg. The top and left most point represent alternative paths to reach the solution with approximately seven moves remaining. The diagram demonstrates that the largest disk in all instances is on the final peg. At this point, the problem solver has three more disks to arrange on the final peg. Having previously completed the three-disk TOH before being presented with the four-disk, the problem is now in a familiar state to the problem solver. The cessation of behaviours at this point could be interpreted to indicate that the path to reach the solution is identifiable, potentially converting it from a problem to a task (Buckley et al., 2018b; Schoenfeld, 1983). However, ascertaining this would require additional insight from problem solvers and observation of variations of cognitive load during this period to inform understanding of these occurrences. Theoretically, as a problem transitions to a task the cognitive load experienced would be expected to reduce as the approach to solving the problem is identified. However, upon reflection on the joint displays of participants, no notable trends of increases or decreases in EDA were identified during the final moves for participants where behaviours ceased.

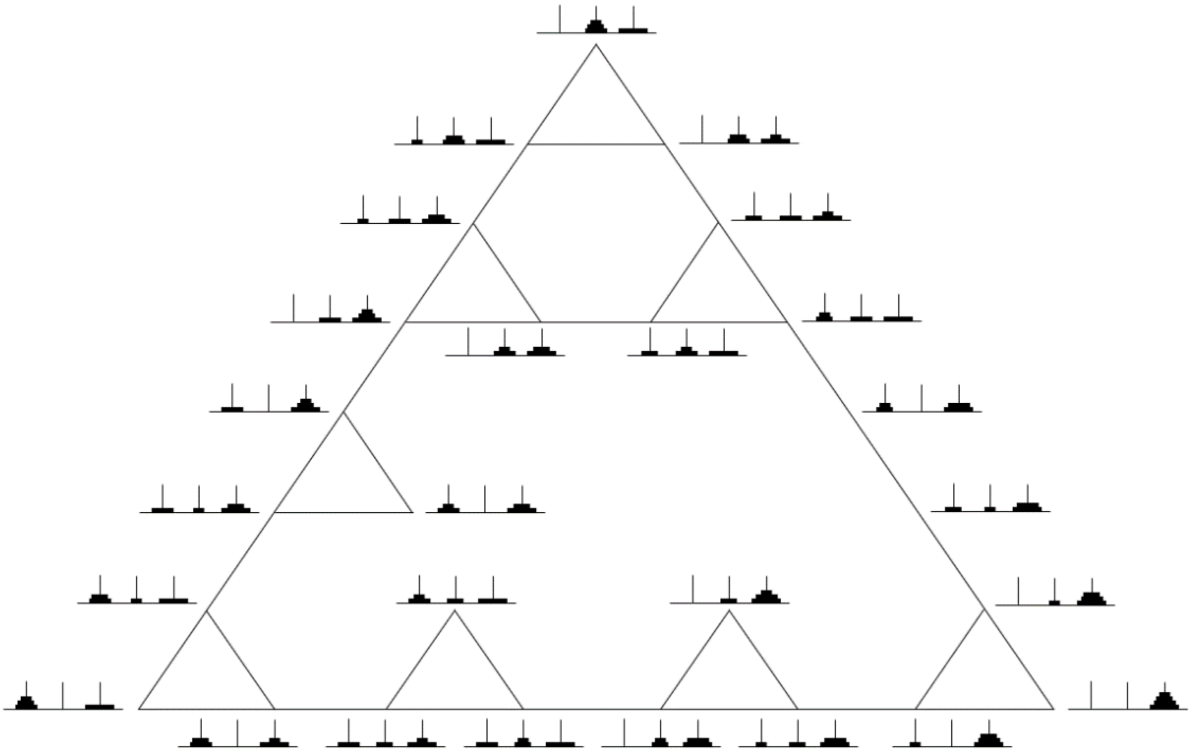


*Figure 20. The configuration of the four-disk TOH with seven moves remaining to reach the solution.*

### 4.3. Implications for the primary study

The insights gained through the preliminary study which will contribute to the primary study include:

- The unobtrusive observation of problem-solving interactions using video recording equipment provided the capacity to move beyond considering problem solving as a broader process towards reflecting on the performance of individuals on a move-by-move basis which goes beyond the current literature base (Patsenko & Altmann, 2010). This affords the opportunity to conduct an in-depth investigation of problem solvers interactions with the problem-solving experience and evaluate the potential role of pertinent factors such as spatial ability in this process.
- The inductive thematic coding of behaviours identified 9 behaviours exhibited during problem solving, including the previously known gesturing, advancing understanding of problem-solving interactions.
- Through evaluating the effectiveness of the methodological design, the psychometric tests used to measure spatial ability were found to significantly

load on a single factor, demonstrated in Table 5. This indicates that combined, the use of these test provides an appropriately precise measure of spatial ability, supporting their use in the main research study.

- Additionally, through piloting of the method it was determined that the EDA sensor did not yield any additional significant insight which could not be gained through the Likert-type item to contribute towards addressing the research objectives. Therefore, the EDA sensor should not be used in the primary study as a measure of cognitive load due to redundancy and efficiency in the context of this research.

The following section sets out the method for the main research study to investigate the research aim in light of the insights gained through this preliminary study.

# 5. Primary study

## 5.1. Primary study method

### *5.1.1. Primary study aim*

The primary study in this research sought to address the overarching research aim to investigate the role of spatial ability in problem solving to advance understanding of the causal theory between spatial ability and success in engineering education.

### *5.1.2. Primary study participants*

Individuals invited to participate in this study were undergraduate students in one Irish third-level institution studying on mechanical and software engineering programmes accredited by Engineers Ireland/QQI in line with the ABET professional standards and Washington Accord. Students in the first year of their engineering course and students in the third mandatory year of their course were invited to participate in this study as an indicator of expertise. As such, a comparison could be made between the complex problem-solving capacity of students at the beginning of their engagement with engineering education and the end of their training.

The necessary sample size required was calculated using G*Power 3.1.9.7 software. The calculation of sample size was based on an effect size of d = 0.5, alpha = 0.05, and power = 0.8 as is common in social science research. A required sample size of *n* = 64 per group was calculated. The investigator sought to recruit all interested participants through email, visits to lectures, and notice board advertisements with an incentive of a place in a draw to win a Samsung Galaxy Tab A7 Lite. The investigator was known to some of the prospective participants. The participants (*n* = 114 1st year students, and *n* = 79 3rd year students) were aware that they could withdraw their participation at any time without providing reason and that their participation was completely voluntary. Further, the recruited sample size, based on a sensitivity analysis, permitted an effect size as small as d = 0.41 to be reliably observed with an alpha level of 0.05 and 80% statistical power. Participant ages ranged from 17 - 61 yrs. Ethical approval was granted through the Athlone Institute of Technology research ethics committee. Participants were provided with an information sheet with details of the study and written informed consent obtained before the data collection began. All participant information was securely stored, and participant numbers assigned to protect participant anonymity.

### *5.1.3. Primary study implementation*

Individuals interested in participating were scheduled to attend two different sessions. Session one, the problem-solving session, and session two, a spatial ability session as demonstrated in Figure 9.

*Session one*: Participants completed two versions of the TOH problem, the three-disk and more difficult four-disk problem. Although through the preliminary study the mean number of moves for the three-disk TOH was close to the minimum number of moves to solve the problem (see Table 4), indicating a potential ceiling effect, the three-disk TOH was included in the primary study. As the participants included in the primary study were at earlier stages of their formal education their complex problem-solving capability may differ to that of the postgraduate students that participated in the preliminary study. Therefore, a ceiling effect may not be seen in this population. Additionally, through proceeding using both problems a comparison could be facilitated between the behaviours exhibited and potential significance of these behaviours when solving an easier and more difficult complex problem. When entering the session participants were provided with the information sheet and opportunity to ask any questions they might have. Participants then completed a written consent form.

To provide the opportunity for retrospective analysis and observation of the problem-solving session, the session was audio and video recorded. Participants were informed of this through their participant information sheet and the camera was orientated to only capture the participants hands as they solved the problem and participant number, which was taped to the table, as in the preliminary study.

Following the brief camera setup, the participant was presented with their first problem, the three-disk TOH and asked to indicate whether they had seen the problem before and if they subsequently knew how to solve it. These indications did not act as an element of selection criteria. The instructions were delivered to participants, and it was emphasized that questions could be asked at any point during the problem if clarification was required. The instructions were as follows:

‘The goal of the task is to get the arrangement of 3 discs on the left-most peg to the right most peg. The discs must be arranged in the same order i.e., largest on the bottom to smallest on the top. There are two conditions:

1. Only one disc can be moved at a time from one peg to another
2. A larger disc cannot be placed on a peg that already contains a smaller disc’

Participants were provided with the opportunity to ask questions and were then instructed to begin the problem.

Performance (i.e., number of moves to solve the problem and time taken) on the problem was retrospectively noted through the video observations. A move was defined as when the disc was placed on a peg and released by the participant. When the problem was completed, participants were presented with the 9-point Likert-type item and asked to rate the amount of mental effort they experienced when solving the problem. Once completed, the second problem, the four-disk TOH, was administered to participants. The same process was repeated for the four-disk problem including the self-reported cognitive load experienced during the problem. Following the completion of the final problem and self-report scale the video recording equipment was switched off.

*Session Two*: Spatial ability was measured in a separate session to reduce the effects of test fatigue that may be experienced through its inclusion in the problem-solving session. The investigator administered three spatial visualisation tests to participants, the PSVT:R/MRT, SDT, and PFT. The order of test administration was randomised between participants to account for order bias. Each test was administered in line with test specifications.

### *5.1.4. Primary study data analysis procedure*

#### *5.1.4.1. Qualitative analysis procedure*

To determine the behaviours exhibited during problem solving, participants interactions with the problem were qualitatively thematically coded using the codebook produced in the preliminary study of nine inductive codes categorised under three themes; “indecision”, “progress”, and “stop”. The video observations collected during the problem-solving session were analysed using the VideoPad Video Editor software. For each participant, the total time to complete the problem and number of moves to solve the problem were documented before qualitative coding began. The footage was reviewed frame-by-frame to determine the exact amount of time spent (in seconds) exhibiting the observable behaviours. This information was documented in a password protected Microsoft Excel file.

#### *5.1.4.2. Quantitative analysis procedures*

The raw quantitative data presented in this thesis was initially compiled in Microsoft Excel before being cleaned and analysed using R Studio (R version 4.1.0.), and IBM SPSS (Statistics 27). Initially the descriptive statistics were determined for the quantitative variables. Following this, a factor analysis was conducted on the spatial test scores. Univariate outliers were then removed for all relevant variables. Comparative analysis across groups were conducted followed by correlation analysis and regression analysis of pertinent variables.

## 5.2. Primary study findings

### *5.2.1. Data cleaning and pre-processing*

Of participants that completed the psychometric spatial tests, 18 completed the MRT in place of the PSVT:R. These 18 participants MRT scores were merged into one variable with the PSVT:R scores of the other 160 participants as previous research has demonstrated a significant correlation with very large effect size between the performance on these tests ($r = 0.621$, $p < 0.001$) (Schmidt, et al., 2020). It is common practice in social science for Cohen's benchmarks for interpretation of the strength of an effect to be used (Cohen, 1962,1988). Cohen recommended that an effect size of $d = 0.5$ to be medium (corresponding to $r = 0.24$), an effect size of $d = 0.2$ as small (corresponding to $r = 0.1$), and a large effect size to be $d = 0.8$ (corresponding to $r = 0.37$). However, there is an issue with these qualifications of strengths as they may not transfer across fields or individual studies (Schäfer & Schwarz, 2019). New effect size interpretations have been proposed based on comparisons of effect sizes with well-understood benchmarks and considering them in terms of consequences such as long-term impact (Fund & Ozer, 2019). Fund and Ozer (2019) propose effect sizes of $r = 0.05$ to be very small (corresponding to $d = 0.1$), $r = 0.1$ to be small (corresponding to $d = 0.2$), medium to be $r = 0.2$ (corresponding to $d = 0.4$), a large effect size to be $r = 0.3$ (corresponding to $d = 0.6$), and a very large effect size to be $r = 0.4$ (corresponding to $d = 0.87$). Based on this, in this thesis, the strength of effect sizes will be described in relative descriptive terms based on the work of Funder and Ozer (2019) as these are more likely to reflect the nature of this research than Cohen's descriptions.

To correct for the MRT and PSVT:R having different numbers of items the scores were first converted to percentages as a method of standardisation before being merged. PFT and SDT scores were subsequently also converted to percentages for consistency. Table 8 presents the descriptive statistics for the spatial measures.

*Table 8. Descriptive statistics for the four individual and two merged spatial tests.*

| | **N** | **M** | **SD** | **Med** | **Min** | **Max** | **Skew** | **Kurt** |
|---|---|---|---|---|---|---|---|---|
| PSVT:R | 160 | 66 | 18.78 | 66.67 | 10.00 | 100 | -.518 | -.124 |
| SDT | 178 | 60.64 | 24.70 | 63.34 | 1.67 | 100 | -.432 | -.768 |
| PFT | 177 | 58.62 | 18.35 | 60 | 10.00 | 100 | -.133 | -.364 |
| MRT | 18 | 52.08 | 18.15 | 56.25 | 12.50 | 75.00 | -.829 | -.174 |
| PSVT:R/ MRT | 178 | 64.60 | 19.14 | 66.67 | 10.00 | 100 | -.507 | -.114 |

N = sample, M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

A factor analysis of the spatial test scores was conducted using IBM SPSS and the results, presented in Table 9, indicated a single underlying latent variable with 70.47% of variance explained by one variable. As this value is above 60% this is deemed satisfactory to interpret the tests as loading on a single factor (Hair Jr. et al., 2014). To compute that variable, the data from the three spatial tests individually were transformed into z-scores. A composite score was then taken as the average across these three variables. This is similar to the procedure implemented by Hambrick et al., (2012) to investigate spatial ability.

*Table 9. Primary study spatial tests factor analysis for a single underlying latent variable*

| **Total Variance Explained** | | | |
|---|---|---|---|
| | | **Initial Eigenvalues** | |
| **Factor** | **Total** | **% of Variance** | **Cumulative %** |
| 1 | 2.114 | 70.470 | 70.470 |
| 2 | .540 | 17.987 | 88.457 |
| 3 | .346 | 11.543 | 100.000 |

Extraction Method: Maximum Likelihood.

Following the z-score calculation RStudio was used to identify outliers in the data. Univariate outliers are extreme values that fall outside the expected values for a variable and are therefore distanced from the majority of cases which are found in the centre of the normal distribution for that variable (Mowbray et al., 2019). These outliers can effect the robustness of a statistical test result. Univariate outliers were identified by distance from the median (Leys et al., 2013), with upper and lower limits denoted by:

$$\mathrm{Q1} - (1.5 \times \mathrm{IQR}) \text{ and } \mathrm{Q3} + (1.5 \times \mathrm{IQR}) \qquad 2$$

In equation 2, Q1 represents quartile one - the cut-off point at the 25th percentile, Q3 represents the third quartile - the cut-off point at the 75th percentile, and IQR is the Interquartile Range which represents the midspread or middle 50% of the data i.e., the difference between the 25th and 75th percentiles. Univariate outliers were identified in numerous variables as detailed in Table 10.

*Table 10. Identification of univariate outliers in the data.*

| | **Univariate outliers in variables** | |
|---|---|---|
| **Variable** | **Univariate outliers** | **% of the data** |
| Spatial ability | 1 | 0.56 |
| Three-disk moves | 5 | 2.82 |
| Three-disk time | 14 | 10.21 |
| Four-disk moves | 10 | 5.95 |
| Four-disk time | 12 | 8.76 |
| Three-disk “indecision” | 13 | 9.49 |
| Three-disk “progress” | 10 | 7.30 |
| Three-disk “stop” | 17 | 12.41 |
| Three-disk behaviour time | 17 | 12.41 |
| Four-disk “indecision” | 10 | 7.30 |
| Four-disk “progress” | 7 | 5.11 |
| Four-disk “stop” | 11 | 8.03 |
| Four-disk behaviour time | 13 | 9.49 |

Following the identification of outliers, they were transformed to the upper and lower limits of respective variables.

The data analysis will be presented in the following order:

a) Comparisons of performance, behaviours, and spatial ability across different levels of engineering expertise (based on progression through an engineering training programme)
b) Associations between spatial ability, performance, cognitive load, and behaviour variables
c) Cause-and-effect analysis of pertinent correlative relationships.

### *5.2.2. Comparing levels of expertise relative to pertinent factors in the research*

The research sought to examine to what, if any, degree spatial ability and engineering expertise predict performance on the TOH. To address this question, the role of expertise in performance was initially explored. This was followed by investigating whether spatial ability is influenced by levels of expertise and finally, a correlation analysis was conducted to determine whether there was an association between spatial ability and performance.

#### *5.2.2.1. Performance and expertise*

Initially, the descriptive statistics were examined to gain insights to the distribution of the data, set out in Table 11. Given skewness and kurtosis values of greater and less than +/- 2 normality of distribution was not indicated (Gravetter & Wallnau, 2014; Trochim & Donnelly, 2006). Visual inspection of the distribution of moves through Figure 21, also indicated that the assumption of normal distribution may not hold.

*Table 11. Descriptive statistics for the number of moves made at different levels of expertise.*

| | **Y** | **N** | **M** | **SD** | **Med** | **Min** | **Max** | **Skew** | **Kurt** |
|---|---|---|---|---|---|---|---|---|---|
| Three-disk moves | 1 | 107 | 9.59 | 3.88 | 8 | 7 | 26 | 2.16 | 5.42 |
| | 3 | 70 | 10.50 | 4.05 | 9.5 | 7 | 26 | 1.67 | 3.78 |
| Four-disk moves | 1 | 99 | 29.67 | 17.78 | 25 | 15 | 101 | 2.01 | 4.05 |
| | 3 | 69 | 32.49 | 18.13 | 27 | 15 | 96 | 1.51 | 2.11 |

Y = Year of study, N = sample size, M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

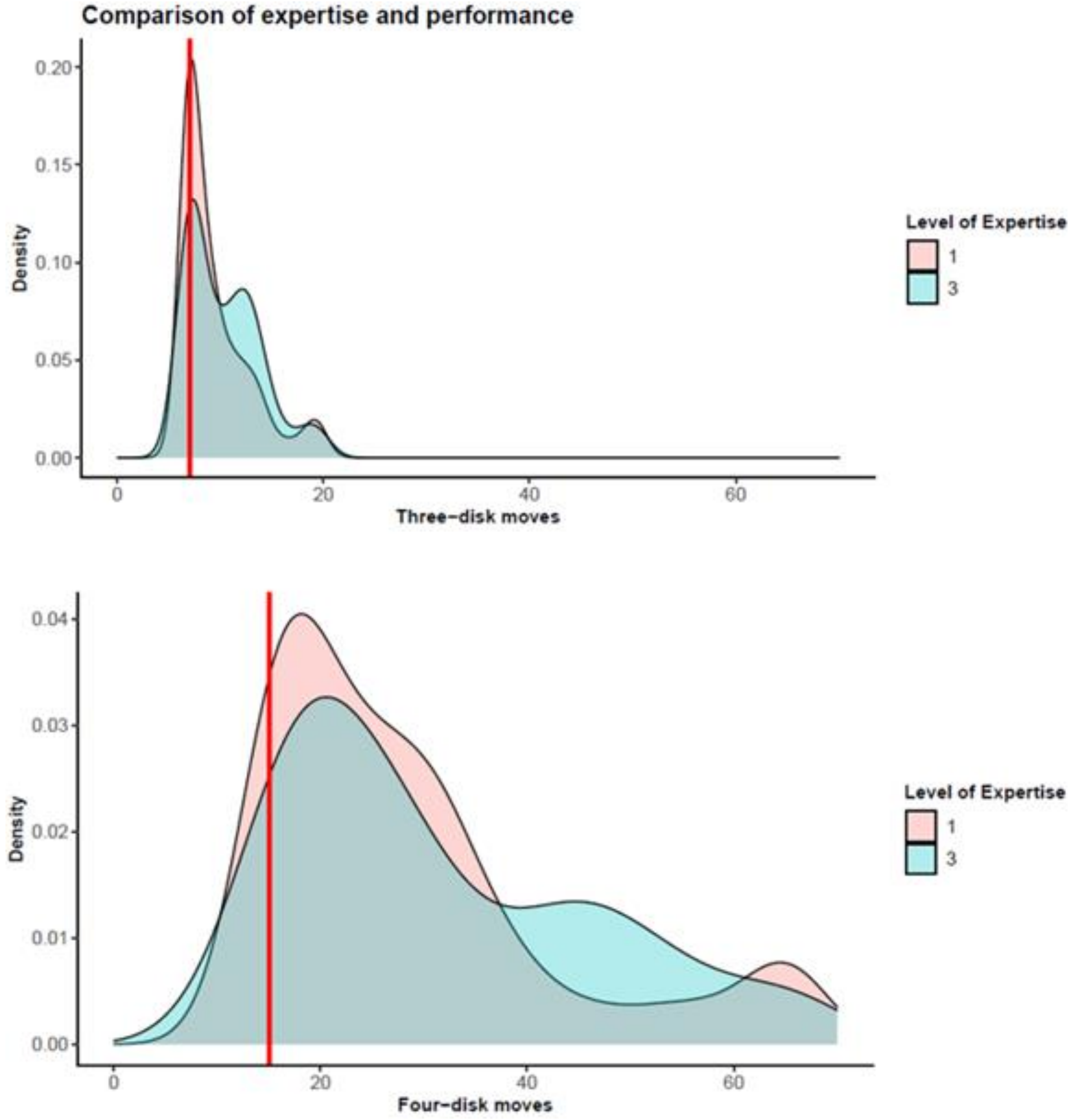


*Figure 21. Density plot comparing moves and expertise where the vertical red line on each graph represents the minimum number of moves possible to solve the problem.*

To statistically confirm these indications, a Shapiro Wilk test was used to determine whether the residuals of the dependent variable, the number of moves to solve the problem, were normally distributed using equation 3, where *n* = sample size, $a_i$ = slope of the observed data, $x_{(i)}$= expected normal values, $x_i$ = variance/spread of the values in the data set, and $\bar{x}$ = mean. The result was statistically significant for both three-disk moves, $W = .836$, $p < 0.05$, ($p = 0.000000000000777$), and four-disk moves, $W = .861$, $p < 0.05$, ($p = 0.0000000000257$). Therefore, the null hypothesis that the residuals are normally distributed was rejected and it is inferred that the data differ significantly from normal distribution.

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{3}$$

To determine homogeneity of variances, a Levene’s test was used. This is outlined in equation 4 below where, $k$ = the number of different groups which the sampled cases belong, $N_i$ = the number of cases in the $i$th group, $N$ is the total number of cases in all groups, $Z_{ij}$ = mean of the $i$th group, $Z_{i.}$ = mean of the $Z_{ij}$ for group $i$, $Z_{..}$ = mean of all $Z_{ij}$. The result was not statistically significant between groups for performance on the three-disk problem, $F(1, 175) = 1.389$, $p = 0.240$, or on the four-disk problem $F(1, 166) = 0.827$, $p = 0.364$. Therefore, the null hypothesis that the variance between groups is identical is accepted.

$$W = \frac{N-k}{k-1} \cdot \frac{\sum_{i=1}^{k} N_i (Z_{i.} - Z_{..})^2}{\sum_{i=1}^{k}\sum_{i=1}^{N_i}\left(Z_{ij} - Z_{i.}\right)^2} \tag{4}$$

Given that the assumption of normality is violated, the Mann-Whitney U test was used to compare the group differences of levels of expertise and performance, the equation for which is outlined through equation 5. $R_1$ = the sum of ranks in sample 1, $n_1$ = the sample size for sample 1.

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} \tag{5}$$

Comparing expertise and moves made on the three-disk TOH with the Mann-Whitney U test, the $p$ value was not statistically significant, $U$= 3128, $p$ > .05 (p= 0.05213), r = 0.15. The $p$ value was also found not to be statistically significant when comparing expertise and moves on the four-disk TOH $U$= 2977.5, $p$ > .05 (p= 0.1576), r = 0.09. As neither of the $p$ values were statistically significant, no evidence was observed to indicate a difference in problem solving capacity, as measured by performance on the TOH, between 1st and 3rd year engineering students where these groups were assumed to indicate a difference in disciplinary expertise.

#### *5.2.2.2. Distribution of moves*

On the three-disk TOH the mean number of moves taken (*M* = 9.45, *SD* = 3.39), is close to the minimum number of moves possible (7 moves) indicating a ceiling effect. To examine this further, boxplots for the total number of moves made in both the three- and four-disk TOH problems (Figure 22) were generated. From these and the distribution graphs (Figure 21), it is evident that there is a ceiling effect with the three-disk TOH where performance is clustered around the minimum number of moves needed to solve the problem.

This ceiling effect in the three-disk TOH indicates that the problem was too easy for the participants which impedes its capacity to differentiate between problem solving performance for the different levels of expertise. Inspecting the four-disk TOH we see a greater spread of moves without a ceiling effect. In this problem, the minimum number of moves needed to solve the problem is 15, and the mean number of moves taken was 28.57 with a standard deviation of 14.39. This indicates that the four-disk TOH is a more useful indicator of problem-solving capacity to proceed with due to having increased potential to discriminate performance.

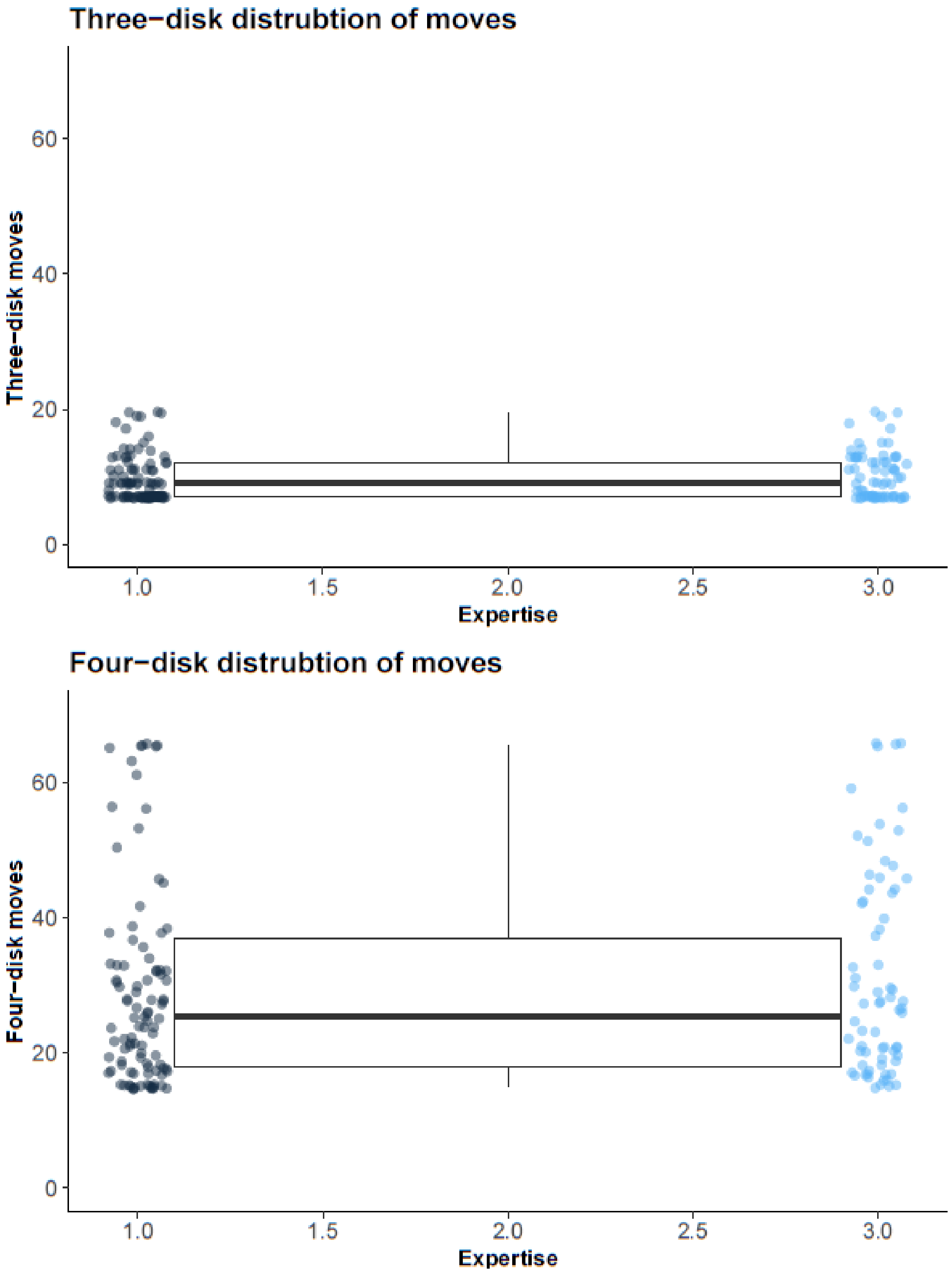


*Figure 22. Boxplot comparison of TOH move distribution at different levels of expertise.*

#### *5.2.2.3. Behaviour themes and expertise*

To investigate the potential for expertise to influence the behaviours displayed during problem solving a comparison of means test was conducted. Initially, the descriptive statistics were examined to gain insights to the distribution of the data (see Table 12). The number of participants included in the analysis of behaviours was reduced due to issues with recording equipment. The assumptions of independent comparison of mean tests were then tested for the overall behaviour time and each of the individual themes of behaviours.

*Table 12. Descriptive statistics for the time spent exhibiting the inductively themed observed behaviours at different levels of expertise.*

| | **Y** | **N** | **M** | **SD** | **Med** | **Min** | **Max** | **Skew** | **Kurt** |
|---|---|---|---|---|---|---|---|---|---|
| Behaviour time | 1 | 76 | 36.38 | 33.78 | 21.79 | 2.03 | 107.8 | 1.07 | -0.14 |
| | 3 | 61 | 41.13 | 33.26 | 31.88 | 0.36 | 107.8 | 0.76 | -0.74 |
| “Indecision” | 1 | 76 | 17.39 | 17.57 | 11.38 | 0 | 58.86 | 1.27 | 0.50 |
| | 3 | 61 | 19.50 | 16.80 | 14.16 | 0 | 58.86 | 0.83 | -0.42 |
| “Progress” | 1 | 76 | 7.45 | 6.77 | 5.13 | 0 | 22.8 | 0.92 | -0.31 |
| | 3 | 61 | 7.26 | 6.70 | 5.26 | 0 | 22.8 | 1.07 | -0.01 |
| “Stop” | 1 | 76 | 11.28 | 14.26 | 4.51 | 0 | 44.59 | 1.26 | 0.36 |
| | 3 | 61 | 13.37 | 14.72 | 6.80 | 0 | 44.59 | 1.07 | -0.24 |

Y = Year of study, N = sample size, M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

A Shapiro Wilk test was used to determine whether the residuals of the dependent variable, behaviour time, were normally distributed. The result was statistically significant, $W = .860$, $p = 0.000000000467$. Therefore, the null hypothesis that the residuals are normally distributed was rejected and it is inferred that the data differ significantly from normal distribution. To determine homogeneity (equality) of variances, the Levene’s test was used. The result was not statistically significant between groups for behaviour time, $F(1, 135) = 0.0373$, $p = 0.8472$. Therefore, the null hypothesis that the variance between groups is identical is accepted. Given that the assumption of normality was violated, the Mann-Whitney U test was used to compare the means of different levels of expertise and performance.

Comparing expertise and behaviour time with the Mann-Whitney U test, the *p* value was not statistically significant, $U= 2065$, $p > .05$ (p= 0.2739), r = 0.09. As the *p* value was not statistically significant, no evidence was observed to indicate that expertise influences the amount of time spent exhibiting the themed behaviours differed significantly between levels of expertise.

#### *5.2.2.4. Spatial ability and expertise*

To determine whether spatial ability differed across different levels of expertise a comparison of means test was conducted. Initially, the descriptive statistics were examined to gain insight to the distribution of the data (Table 13).

*Table 13. Descriptive statistics of spatial ability across levels of expertise.*

| | **Y** | **N** | **M** | **SD** | **Med** | **Min** | **Max** | **Skew** | **Kurt** |
|---|---|---|---|---|---|---|---|---|---|
| Spatial ability | 1 | 102 | -0.03 | 0.86 | 0 | -2.32 | 1.55 | -0.21 | -0.61 |
| | 3 | 75 | 0.04 | 0.80 | 0.15 | -1.95 | 1.9 | -0.27 | -0.47 |

Y = Year of study, N = sample size, M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

Again, a Shapiro Wilk test was used to determine whether the residuals of the dependent variable, spatial ability, were normally distributed. The result was not statistically significant, $W$ = .989, $p$ = 0.198, therefore, the null hypothesis that the residuals are normally distributed can be accepted. Using the Levene's test to determine the homogeneity of variances found $F(1, 175) = 0.7424$, $p > .05$ ($p$ = 0.3901), which is not statistically significant, indicating that the null hypothesis (variances are equal) should be accepted.

An independent samples t-test was conducted on the sample of 177 engineering students to determine if there was a statistically significant mean difference in spatial ability between 1st (n = 102) and 3rd (n = 75) year students. The equation for an independent samples t-test is outlined in equation 7 where $\bar{x}_1$ = mean value of group 1, $\bar{x}_2$ = mean value of group 2, $n_1$ = size of group 1, $n_2$ = size of group 2, $s_1$ = standard deviation for group 1, and $s_2$ = standard deviation for group 2. The results indicate no statistical difference between the two groups t(175) = -0.56, p = 0.57, d= 0.43.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad 6$$

### *5.2.3. Investigating associations between pertinent factors in the research*

In addition to determining whether an association exists between spatial ability and performance, to address the remaining research questions required the examination of potential associations between these variables and each of the other variables i.e., cognitive load and behaviours. As the performance, cognitive load, and behavioural data are not normally distributed, and the cognitive load data is Likert data, Spearman's rank correlation with Holm correction was conducted. The equation for Spearman's rank correlation is outlined in equation 7 where $d_i$ = difference between the two ranks of each observation, and $n$ = number of observations.

$$\rho = 1 - \frac{6\Sigma\, d_i^2}{n(n^2 - 1)} \tag{7}$$

#### *5.2.3.1. Associations between spatial ability and problem-solving performance*

A Spearman's rank correlation was conducted to evaluate the role of spatial ability in problem-solving performance. As demonstrated in Table 14, significant negative correlations were identified between spatial ability and each of the performance measures of (a) the number of moves to solve the problem, r(155) = -.23, p = .033, and (b) the time taken to solve the problem, r(135) = -.28, p = .022, with medium effect sizes. This indicates that while spatial ability increases the number of moves to solve the problem and time taken to solve it decreases.

*Table 14. Spearman correlation for spatial ability and problem-solving performance.*

| **Variable** | **M** | **SD** | **1** | **2** |
|---|---|---|---|---|
| 1. Spatial ability | 0.01 | 0.83 | | |
| 2. Moves | 29.55 | 14.81 | -.23**<br>[-.38, -.08] | |
| 3. Time | 110.57 | 72.31 | -.28**<br>[-.43, -.11] | 0.78***<br>[.70, .84] |

p-value adjustment method: Holm (1979).
Observations: 137-155.
Note: **. Correlation is significant at the 0.01 level (two-tailed).
*. Correlation is significant at the 0.05 level (two-tailed).

#### *5.2.3.2. Associations between spatial ability and behaviour themes*

A Spearman's rank correlation was conducted to evaluate the relationship between spatial ability and the inductively behaviour themes of "indecision", "stop", and "progress". Total time spent exhibiting all themes was also included in this correlation. As demonstrated in Table 15, significant negative correlations were identified between spatial ability and time spent (a) exhibiting behaviours in the “indecision” theme, r(135) = -.27, p = .005, (b) exhibiting behaviours in the “progress” theme, r(135) = -.20, p = .04, and (c) overall behaviour time r(135) = -.22, p = .03, with medium effect sizes. This indicates that while spatial ability increases the time spent exhibiting these themes of behaviours and overall time spent exhibiting behaviours decreases.

*Table 15. Spearman correlation for spatial ability and time spent exhibiting the themed behaviours.*

| **Variable** | **M** | **SD** | **1** | **2** | **3** | **4** |
|---|---|---|---|---|---|---|
| 1. Spatial ability | 0.01 | 0.83 | | | | |
| 2. “Indecision” | 18.30 | 17.11 | -.27**<br>[-.42, -.10] | | | |
| 3. “Stop” | 12.11 | 14.41 | -.11<br>[-.28, .06] | .63***<br>[.51, .72] | | |
| 4. “Progress” | 7.31 | 6.71 | -.20*<br>[-.36, -.03] | .55***<br>[.40, .65] | .47***<br>[.33, .60] | |
| 5. Behaviour time | 38.28 | 33.32 | -.22*<br>[-.37, -.05] | .90***<br>[.86, .93] | .84***<br>[.78, .88] | .70***<br>[.60, .78] |

p-value adjustment method: Holm (1979).
Observations: 137.
Note: **. Correlation is significant at the 0.01 level (two-tailed).
*. Correlation is significant at the 0.05 level (two-tailed).

### *5.2.3.3. Associations between spatial ability and mental effort*

A Spearman’s rank correlation was conducted to investigate the relationship between spatial ability and mental effort (overall cognitive load) experienced during problem solving. A significant negative correlation was identified between spatial ability and mental effort, (r(155) = -.18, *p* = .03), of small effect size. This indicates that while spatial ability increases the mental effort experienced during problem solving decreases.

### *5.2.3.4. Associations between mental effort and behaviour themes*

A Spearman’s rank correlation was conducted to investigate an association between the mental effort experienced during problem solving, behaviour themes, and total time spent exhibiting the behaviour themes. As demonstrated in Table 16, significant positive correlations were determined between mental effort and “indecision” (r(135) = .38, *p* < .001), “stop” (r(135) = .37, *p* < .001), “progress” (r(135) = .37, *p* < .001), with large effect sizes, and with total behaviour time (r(135) = .44, *p* < .001) having a very large effect size. This indicates that as mental effort increases during problem solving, the time spent exhibiting the observable coded behaviours also increases.

*Table 16. Spearman correlation for mental effort and time spent exhibiting observed behaviours.*

| Variable | M | SD | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|
| 1. Mental effort | 4.72 | 1.67 | | | | |
| 2. “Indecision” | 18.30 | 17.11 | .38***<br>[.22, .51] | | | |
| 3. “Stop” | 12.11 | 14.41 | .37***<br>[.19, .49] | .63***<br>[.51, .72] | | |
| 4. “Progress” | 7.31 | 6.71 | .37***<br>[.19, .49] | .55***<br>[.40, .65] | .47***<br>[.33, .60] | |
| 5. Behaviour time | 38.28 | 33.32 | .44***<br>[.28, .56] | .90***<br>[.86, .93] | .84***<br>[.78, .88] | .70***<br>[.60, .78] |

p-value adjustment method: Holm (1979).
Observations: 137.
Note: **. Correlation is significant at the 0.01 level (two-tailed).
*. Correlation is significant at the 0.05 level (two-tailed).

#### *5.2.3.5. Associations between mental effort and problem-solving performance*

A Spearman’s rank correlation was conducted to evaluate the role of mental effort in problem-solving performance. As demonstrated in Table 17, significant positive correlations were identified between mental effort and each of the performance measures of (a) the number of moves to solve the problem, (r(166) = .31, *p* < .001), where a large effect size was found, and (b) the time taken to solve the problem, (r(135) = .49, *p* < .001), where a very large effect size was found. This indicates that while mental effort increases the number of moves to solve the problem and time taken to solve it also increases.

*Table 17. Spearman correlation for mental effort and performance.*

| Variable | M | SD | 1 | 2 |
|---|---|---|---|---|
| 1. Mental effort | 4.72 | 1.67 | | |
| 2. Moves | 29.55 | 14.81 | .31***<br>[.17, .45] | |
| 3. Time | 110.57 | 72.31 | .49***<br>[.35, .61] | .78***<br>[.71, .84] |

p-value adjustment method: Holm (1979).
Observations: 137-168.
Note: **. Correlation is significant at the 0.01 level (two-tailed).

*. Correlation is significant at the 0.05 level (two-tailed).

#### *5.2.3.6. Associations between performance and behaviour themes*

Finally, a Spearman’s rank correlation was conducted to investigate the relationship between problem-solving performance, behaviour themes, and total time spent exhibiting the behaviour themes. As demonstrated in Table 18, significant positive correlations were identified between number of moves made to solve the problem and time r(155) = .78, p = .033, “indecision” r(155) = .61, p = .033, “stop” r(155) = .40, p = .033, “progress” r(155) = .56, p = .033, and behaviour time r(155) = .60, p = .033, with very large effect sizes. Significant positive correlations were also found between time to solve the problem and “indecision” r(155) = .86, p = .033, “stop” r(155) = .75, p = .033, “progress” r(155) = .71, p = .033, and behaviour time r(155) = .93, p = .033, also with very large effect sizes.

*Table 18. Spearman correlation for performance and time spent exhibiting observed behaviours.*

| **Variable** | **M** | **SD** | **1** | **2** | **3** | **4** | **5** |
|---|---|---|---|---|---|---|---|
| 1. Moves | 4.72 | 1.67 | | | | | |
| 2. Time | 110.57 | 72.31 | .78*** | | | | |
| | | | [.70, .84] | | | | |
| 3. “Indecision” | 18.30 | 17.11 | .61*** | .86*** | | | |
| | | | [.49, .71] | [.81, .90] | | | |
| 4. “Stop” | 12.11 | 14.41 | .40*** | .75*** | .63*** | | |
| | | | [.24, .53] | [.65, .81] | [.51, .72] | | |
| 5. “Progress” | 7.31 | 6.71 | .56*** | .71*** | .55*** | .47*** | |
| | | | [.41, .66] | [.60, .78] | [.40, .65] | [.33, .60] | |
| 6. Behaviour time | 38.28 | 33.32 | .60*** | .93*** | .90*** | .84*** | .70*** |
| | | | [.47, .70] | [.90, .95] | [.86, .93] | [.78, .88] | [.60, .78] |

p-value adjustment method: Holm (1979).
Observations: 137.
Note: **. Correlation is significant at the 0.01 level (two-tailed).
*. Correlation is significant at the 0.05 level (two-tailed).

### *5.2.4. Investigating predictive associations between the pertinent variables in the research*

#### *5.2.4.1. Spatial ability and performance*

Having identified a significant negative correlation between spatial ability and the number of moves to solve the problem, r(155) = -.23, p = .033, and time, r(135) = -.28, p = .022 (Table 14), the association was explored further. To evaluate whether spatial ability predicts performance on the TOH regression analysis was explored. As the dependent variable (spatial ability) is continuous, the assumptions of multiple regression analysis, a linear model, were tested. The descriptive statistics for each of the variables were initially examined to gain insight to the distribution of the data (Table 19). A Shapiro Wilk test was used to test the distribution of the variable residuals. The result was statistically significant $W = .944$, $p < 0.05$, ($p = 0.000024$), indicating that the null hypothesis that the residuals are normally distributed was rejected and it is inferred that the data differ significantly from normal distribution.

*Table 19. Descriptive statistics for spatial ability, time, and moves.*

| | **N** | **M** | **SD** | **Med** | **Min** | **Max** | **Skew** | **Kurt** |
|---|---|---|---|---|---|---|---|---|
| Spatial ability | 177 | 0 | 0.84 | 0.04 | -2.32 | 1.9 | -0.25 | -0.52 |
| Moves | 168 | 29.71 | 14.66 | 25.5 | 15 | 65.5 | 1.11 | 0.22 |
| Time | 137 | 110.72 | 72.65 | 88.06 | 25.57 | 271.1 | 1.07 | 0 |

N = sample, M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

The assumption of linearity was tested through visual inspection. Through this inspection (Figure 23), no evidenced was found to indicate a linear relationship between the variable. Multicollinearity was also determined between the independent variables (Table 14). As multiple assumptions were breached, non-parametric ordinal logistic regression was also conducted.

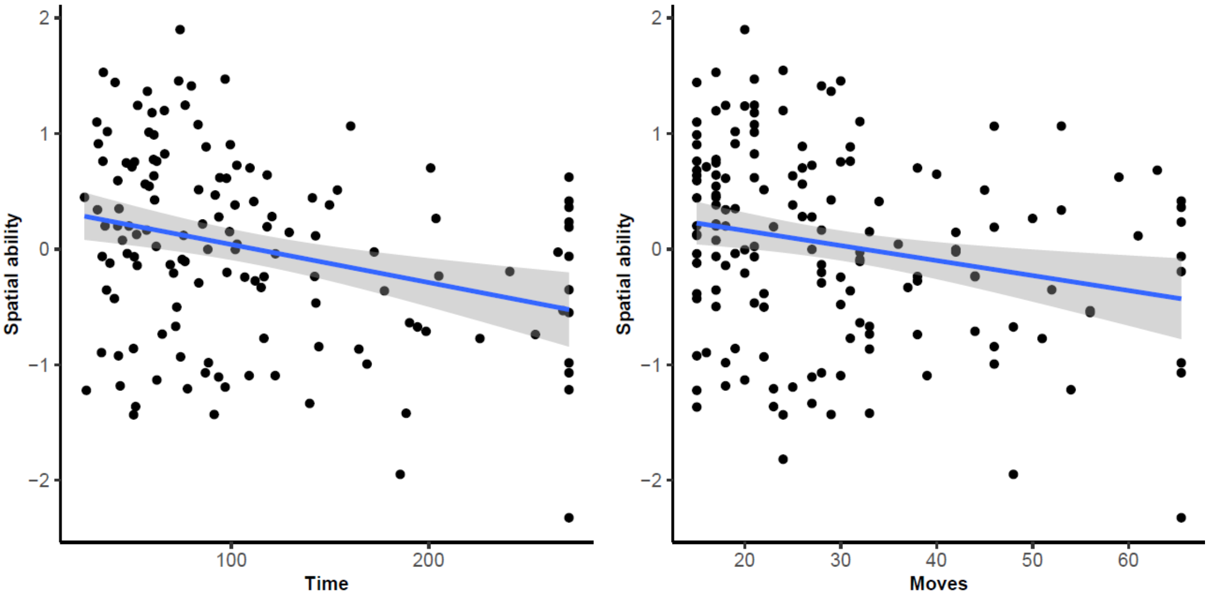


*Figure 23. Testing linearity between the variables of spatial ability, time, and moves.*

To conduct an ordinal logistic regression, spatial ability which has thus far been treated as a continuous variable had to be considered in terms of groups. In this study, quartiles were used to define participants relatively within the sample into four groups of different levels of spatial ability, as has been done in various other studies (Buckley et al., 2019; Veurink & Hamlin, 2011; Wai et al., 2009; Webb et al., 2007). In categorising the spatial variable in quartiles, an ordinal logistic regression could be conducted to determine the log odds of a dependent variable being influenced by independent variables. There are four assumptions of ordinal logistic regression:

1. Dependent variable is ordinal
2. Independent variables are continuous, categorical, or ordinal
3. No multicollinearity
4. Proportional odds- there is an identical effect of the independent variable at each cumulative split in the dependent variable

The spatial ability variable was categorised into quartiles where 1 = upper quartile and 4 = lowest quartile. The independent variables are continuous. However, multicollinearity was determined between the independent variables (Table 14). Finally, the proportional odds assumption was tested and assessed graphically.

When testing the proportional odds assumption, the independent variables, in this instance moves and time, are categorised into quartiles ranging from the lowest to the largest measures that were collected. This is demonstrated on the left side of Figure 24 below which presents fictitious data to graphically explain the difference between the proportional odds assumption holding and not holding. For the proportional odds assumption to hold, the coefficients which describe the relationship between for example the lowest level of the dependent variable and each of the quartiles of moves and time should be the same. This should apply for each of the levels of the dependent variable. The levels of the dependent variable are represented by symbols such as a triangle, diamond, etc. as demonstrated in Figure 24 where the fictitious dependent variable has two categories which are represented by a triangle and diamond. In instances where the proportional odds assumption holds, the symbol for the level of the dependent variable should be in approximately the same position for each of the levels of the independent variable, this would indicate that the coefficients are the same. Viewing the fictitious data in Figure 24, the proportional odds assumption would hold for the time variable as the diamond symbol (representing a level of the dependent variable) is in approximately the same position for the quartile splits of the independent variable. However, the proportional odds assumption would not hold for the moves variable as there is significant variation in the alignment of the diamond symbol. The triangle symbol is aligned for both independent variables as this category of the dependent variable is set to zero as a reference point to compare the alignment of other categories of the dependent variable at each quartile level.

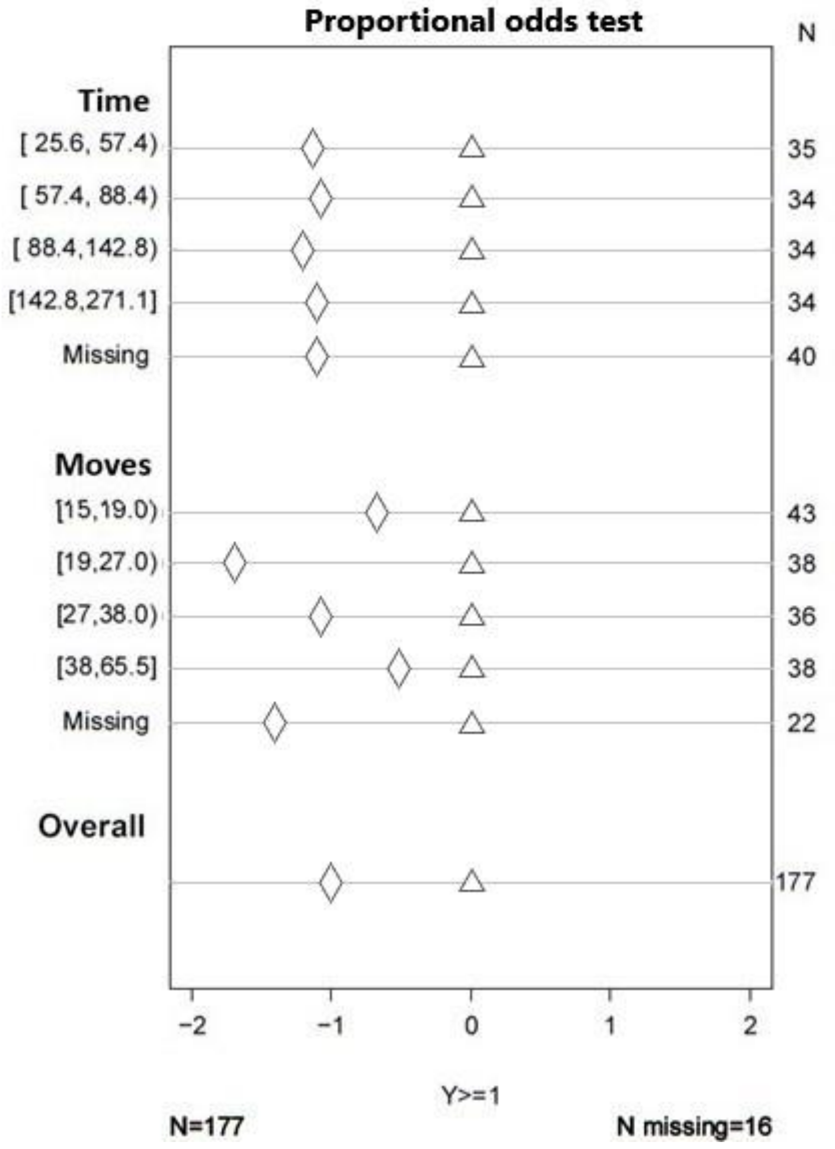


*Figure 24. Testing the proportional odds assumption differentiating between instances where the assumption holds and does not hold.*

Figure 25 details the test of the proportional odds assumption for spatial ability, time, and moves. Again, one level of the dependent variable spatial ability is set at zero (triangle symbol) as a reference point to compare the alignment of the other symbols. Through visual inspection of the graph, it is evident that the alignment of each of the symbols is not similar at each quartile split in the time and move variables. This indicates that the proportional odds assumption for the ordinal logistic regression was violated.

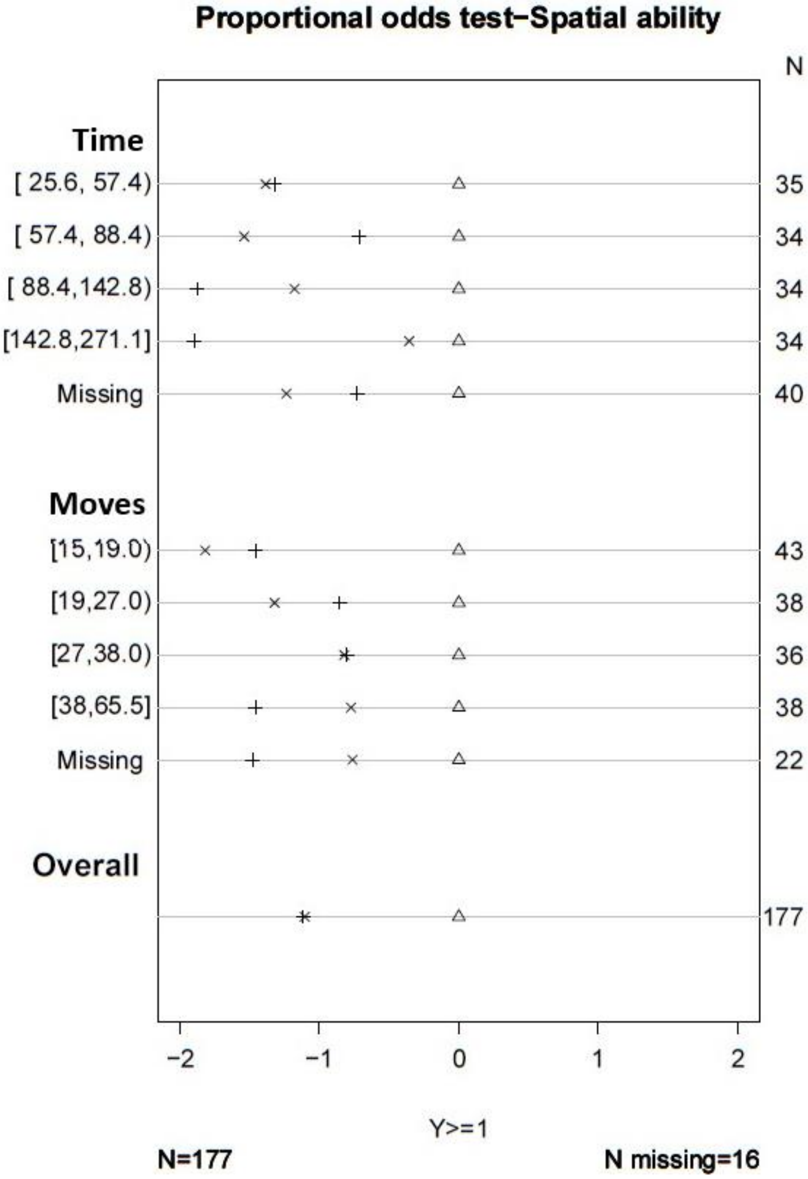


*Figure 25. Test of the proportional odds assumption for spatial ability, time, and moves.*

As assumptions of both regression models were violated, the results obtained using either model may not be valid. However, both models were implemented to determine what tentative insight they may provide.

A multiple linear regression was used to determine whether spatial ability was predictive of the time spent solving the problem or number of moves. Results indicated that there was a collective significant effect between time, moves, and spatial ability, $F(6.7, 2) = 134$, $p = 0.002$, $R^2 = 0.09$, $R^2_{\text{adjusted}} = 0.07$. On an individual level neither time ($t = -1.140$, $p = 0.257$) or moves ($t = -0.814$, $p = 0.417$) were found to be significant predictors in the model.

An ordinal logistic regression was conducted, with spatial ability categorised in quartile. The results from this analysis also found no significant effect between time, moves, and spatial ability (Table 20). The odds ratio for time (*OR* = 1.00), indicates that for every one-unit increase or decrease in time, the odds of being more likely to have higher spatial ability versus lower spatial ability is multiplied 1 time, holding constant all other variables. The odds ratio calculated for moves (*OR* = 1.01), indicates that for every one-unit increase or decrease in time spent displaying move behaviours, the odds of being more likely to have higher or lower spatial ability is multiplied by 1.01 (i.e., increases 1%), holding constant all other variables. Caution should be taken when interpreting these results as the assumptions of both models were violated.

*Table 20. Odds ratio data for spatial ability, time, and moves.*

| | | | | | **95% CI** | |
|---|---|---|---|---|---|---|
| **Variable** | **Estimate** | **SE** | ***p*** | **OR** | **Lower** | **Upper** |
| Time | 0.004 | 0.004 | 0.26 | 1.00 | 1.00 | 1.01 |
| Moves | 0.014 | 0.019 | 0.47 | 1.01 | 0.98 | 1.05 |

SE = Standard error, *p* = p-value, OR = Odds ratio, CI = confidence interval

#### *5.2.4.2. Spatial ability and behaviour themes*

Having identified significant negative correlations between spatial ability and the time exhibiting behaviours in the “indecision”, r(135) = -.27, p = .005, and “progress” themes, r(135) = -.20, p = .04, and overall behaviour time r(135) = -.22, p = .03 (Table 15), further analysis was conducted to explore this association. To determine whether prediction existed between the variable’s a multiple regression analysis was conducted. Given that the dependent variable (spatial ability) is continuous, the assumptions of multiple regression analysis, a linear model, were tested. The descriptive statistics for each of the variables were initially examined to gain insight to the distribution of the data (Table 21). A Shapiro Wilk test was used to test the distribution of the variable residuals. The result was statistically significant $W = .939$, $p < 0.05$, ($p = 0.0000102$), indicating that the null hypothesis that the residuals are normally distributed was rejected and it is inferred that the data differ significantly from normal distribution.

*Table 21. Descriptive statistics for spatial ability and time spent exhibiting inductively coded behaviours.*

| | **N** | **M** | **SD** | **Med** | **Min** | **Max** | **Skew** | **Kurt** |
|---|---|---|---|---|---|---|---|---|
| Spatial ability | 177 | 0 | 0.84 | 0.04 | -2.32 | 1.9 | -0.25 | -0.52 |
| “Indecision” | 137 | 18.33 | 17.2 | 12.23 | 0 | 58.86 | 1.08 | 0.11 |
| “Progress” | 137 | 7.37 | 6.71 | 5.2 | 0 | 22.8 | 1 | -0.41 |
| Behaviour time | 137 | 38.5 | 33.51 | 26.6 | 0.36 | 107.8 | 0.93 | -0.41 |

N = sample, M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

The assumption of linearity was tested through visual inspection (Figure 26) yielding no evidence to suggest a linear relationship between the variables. Multicollinearity was found between the independent variables (Table 15). As multiple assumptions were breached, non-parametric ordinal logistic regression was also conducted.

For the ordinal logistic regression spatial ability was considered in terms of quartiles. In doing this the dependent variable would be ordinal and the independent variables continuous. However, the assumption of multicollinearity was breached. Finally, the proportional odds assumption was tested. Through visual inspection of the graph in Figure 27, the alignment of the symbols for each set of categories of the dependent variable, spatial ability, are not similar. This indicates that the proportional odds assumption for the ordinal logistic regression was violated.

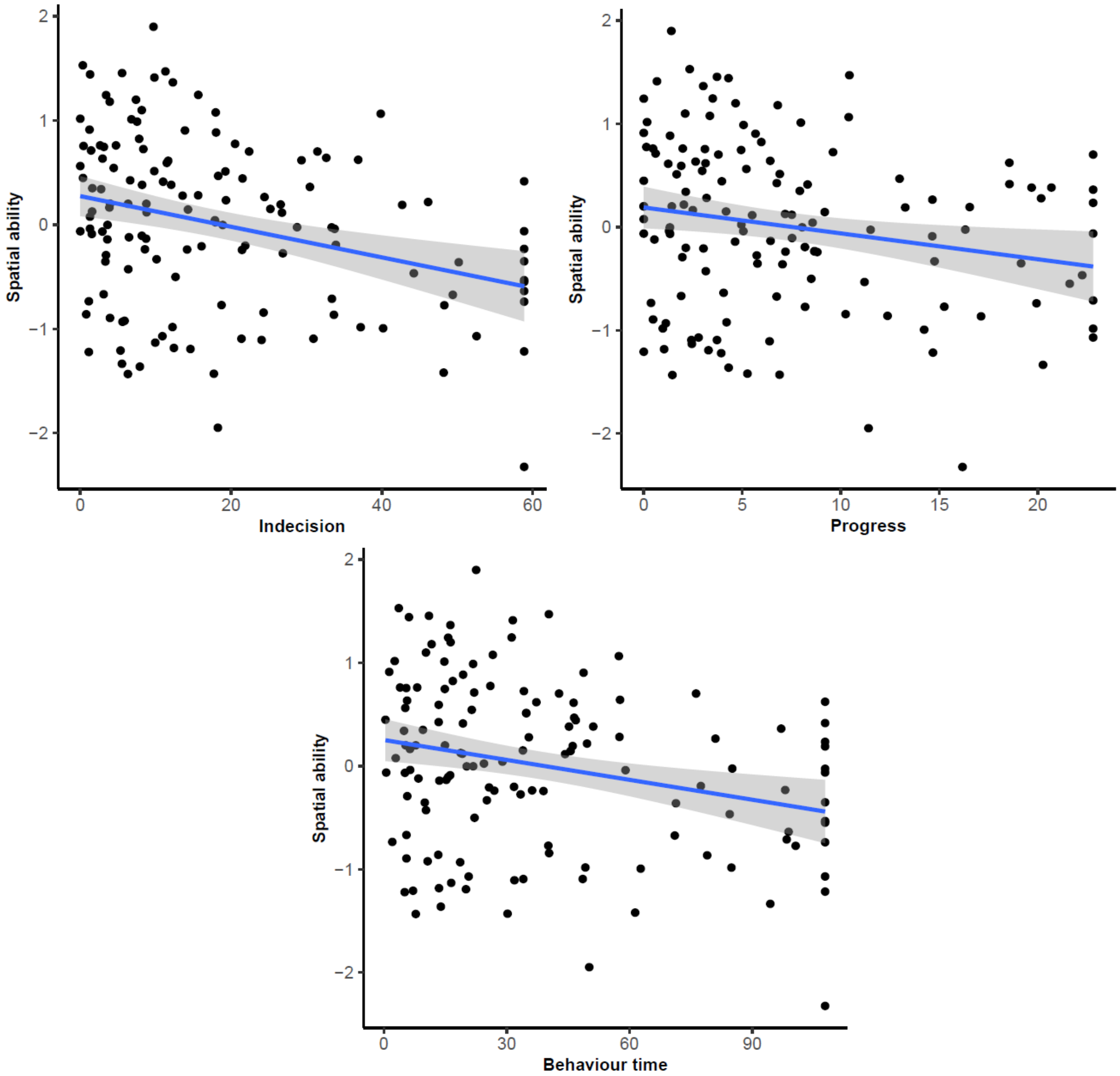


*Figure 26. Testing linearity between spatial ability and time spent exhibiting “indecision” and “progress” themes and total behaviour time.*

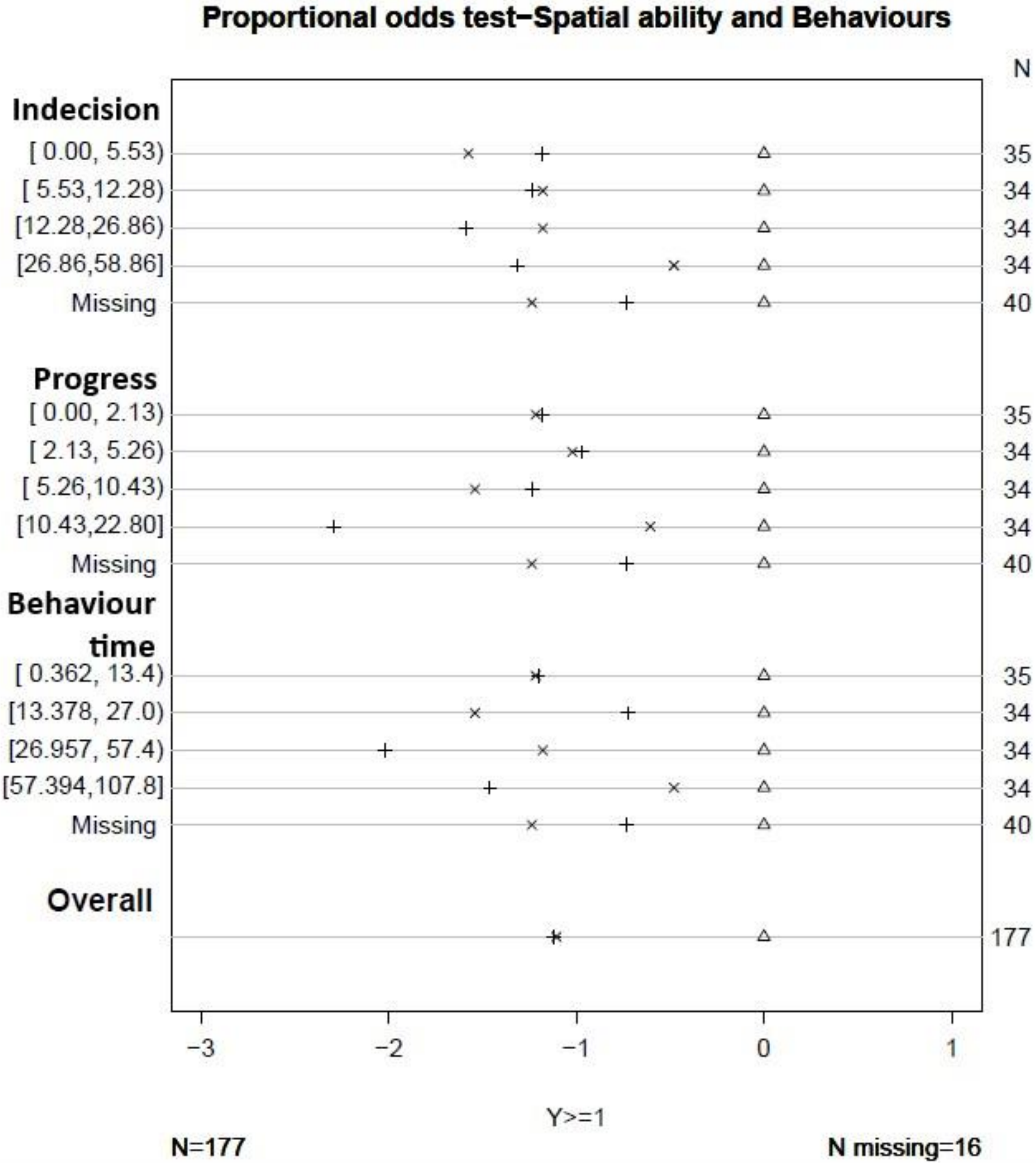


*Figure 27. Test of proportional odds assumption for spatial ability and behaviour themes and time.*

As the assumptions of both regression models were violated, the results obtained using either model may not be valid. However, both models were implemented to determine what tentative insight they may provide. A multiple linear regression was used to determine whether spatial ability was predictive of the time spent exhibiting the observed behavioural themes of “indecision” and “progress” and overall behaviour time. Results indicated that there was a collective significant effect between “indecision”, “progress”, behaviour time, and spatial ability, $F(5.1, 3) = 133$, $p = 0.002$, $R^2 = 0.10$, $R^2_{\text{adjusted}} = 0.08$. On an individual level neither “progress” ($t = -0.924$, $p =$

0.357) or behaviour time (*t* = 0.783, *p* = 0.435) were found to be significant predictors in the model. The time spent exhibiting the behaviour themed as "indecision" was found to have a significant effect (*t* = -2.194, *p* = 0.03). As multiple assumptions of the model were violated, this cannot be reliably interpreted as a significant finding.

An ordinal logistic regression was conducted, with spatial ability categorised in quartiles. The odds ratios calculated for "indecision" (*OR* = 1.04), "progress" (OR = 1.03), and behaviour time (*OR* = 0.99) (Table 22), indicate that for every one-unit increase or decrease in any of the variables, the odds of being more likely to have higher spatial ability versus lower spatial ability is negligible.

*Table 22. Odds ratio data for spatial ability and behaviour themes and time.*

| | | | | | 95% CI | |
|---|---|---|---|---|---|---|
| **Variable** | **Estimate** | **SE** | ***p*** | **OR** | **Lower** | **Upper** |
| "Indecision" | 0.041 | 0.02 | 0.04 | 1.04 | 1.01 | 1.08 |
| "Progress" | 0.028 | 0.04 | 0.43 | 1.03 | 0.97 | 1.09 |
| Behaviour time | -0.009 | 0.01 | 0.46 | 0.99 | 0.97 | 1.01 |

SE = Standard error, *p* = p-value, OR = Odds ratio, CI = confidence interval

#### *5.2.4.3. Spatial ability and mental effort*

From the significant positive correlation determined between spatial ability and the mental effort experienced during problem solving, (r(155) = .18, *p* = .03), in section 5.2.3.3, regression analysis was conducted to gain further insight to this association. As the data collected from the self-reported cognitive load measure was ordinal, an ordinal logistic regression was deemed to be the most suitable approach to consider a potential predictive association. The independent variable mental effort was ordinal, the dependent variable of spatial ability was divided into quartiles and the assumption of multicollinearity could not be violated as there were no other variables in the model. Finally, the proportional odds assumption was tested.

Through visual inspection of the graph in Figure 28, the alignment of the symbols for each set of categories of the dependent variable, spatial ability, are not similar. This indicates that the proportional odds assumption for the ordinal logistic regression was violated.

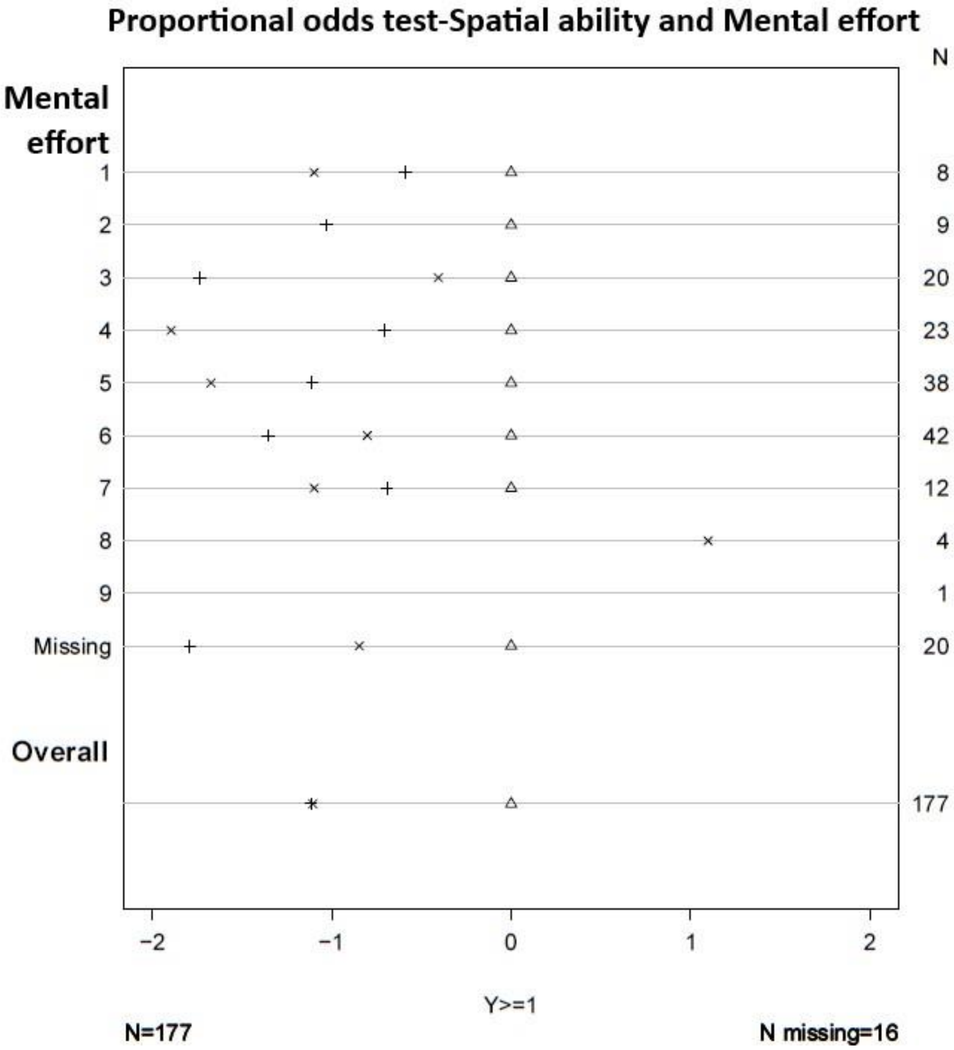


*Figure 28. Test of proportional odds assumption for spatial ability and mental effort.*

As the assumptions of the regression were violated, the results obtained using the model may not be valid. However, the analysis was conducted to determine what tentative insight it could provide. The results from this analysis indicated no significant effect between mental effort and spatial ability. The odds ratio for mental effort (*OR* = 1.23), indicates that for every one-unit increase or decrease in mental effort, the odds of being more likely to have higher spatial ability versus lower spatial ability is multiplied by 1.23 (i.e., increases 23%), holding constant all other variables. This indicates a non-significant result.

#### *5.2.4.4. Mental effort and behaviours*

Strong positive correlations were determined between mental effort and “indecision” (r(135) = .38, *p* < .001), “stop” (r(135) = .37, *p* < .001), “progress” (r(135) = .37, *p* < .001), and behaviour time (r(135) = .44, *p* < .001). Having identified the associations necessitating further investigation, the four assumptions of ordinal logistic regression were tested. The dependent variable, mental effort, is ordinal. Each of the independent variables are continuous. Multicollinearity was identified between the independent variables as demonstrated by the significant correlations between them in Table 16. The proportional odds assumption was then tested where it was indicated through visual inspection of the graph in Figure 29 that the assumption was violated. As the proportional odds and multicollinearity assumptions are violated, the results determined from the model may not be valid, therefore interpretation of the significance of results should be done with caution.

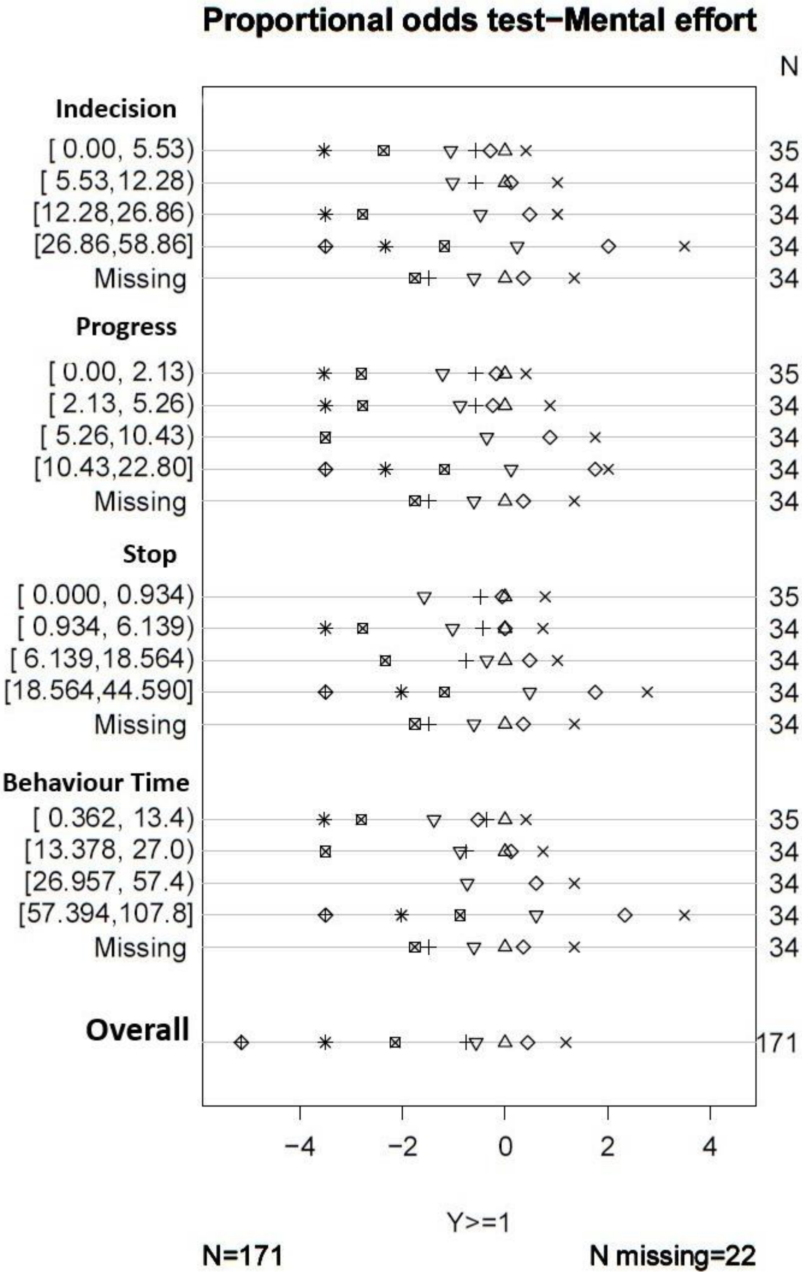


*Figure 29. Test of proportional odds assumption for mental effort and time spent exhibiting behaviours.*

The results from the ordinal logistic regression (Table 23) found no significant effect between mental effort and the behaviour theme variables. The odds ratio for "indecision" (*OR* = 1.03), "progress" (OR = 1.07), "stop" (*OR* = 1.04) and behaviour time (*OR* = 0.99), indicate that for every one-unit increase or decrease in any of the variables, the odds of being more likely to have higher mental effort versus lower mental effort is negligible. These findings tentatively indicate that mental effort is not predictive of behaviour indicators during complex problem-solving performance.

*Table 23. Odds ratio data for mental effort and time spent exhibiting behaviours.*

| | | | | | **95% CI** | |
|---|---|---|---|---|---|---|
| **Variable** | **Estimate** | **SE** | ***p*** | **OR** | **Lower** | **Upper** |
| "Indecision" | 0.032 | 0.03 | 0.29 | 1.03 | 0.98 | 1.09 |
| "Progress" | 0.070 | 0.05 | 0.14 | 1.07 | 0.99 | 1.16 |
| "Stop" | 0.035 | 0.03 | 0.28 | 1.04 | 0.98 | 1.10 |
| Behaviour time | -0.007 | 0.03 | 0.80 | 0.99 | 0.94 | 1.04 |

SE = Standard error, *p* = p-value, OR = Odds ratio, CI = confidence interval

#### *5.2.4.5. Mental effort and performance*

Strong positive correlations were determined between mental effort, moves (r(166) = .31, *p* < .001), and time (r(135) = .49, *p* < .001). Having identified the associations necessitating further investigation, a non-parametric regression analysis was conducted. The four assumptions of ordinal logistic regression were tested. The dependent variable, mental effort, is ordinal. Each of the independent variables are continuous. Multicollinearity was identified between the independent variables as demonstrated by the significant correlations between them Table 17. The proportional odds assumption was then tested where it was indicated through visual inspection of the graph in Figure 30 that the assumption was violated. As the proportional odds and multicollinearity assumptions are violated, the results determined from the model may not be valid, therefore interpretation of the significance of results should be done with caution.

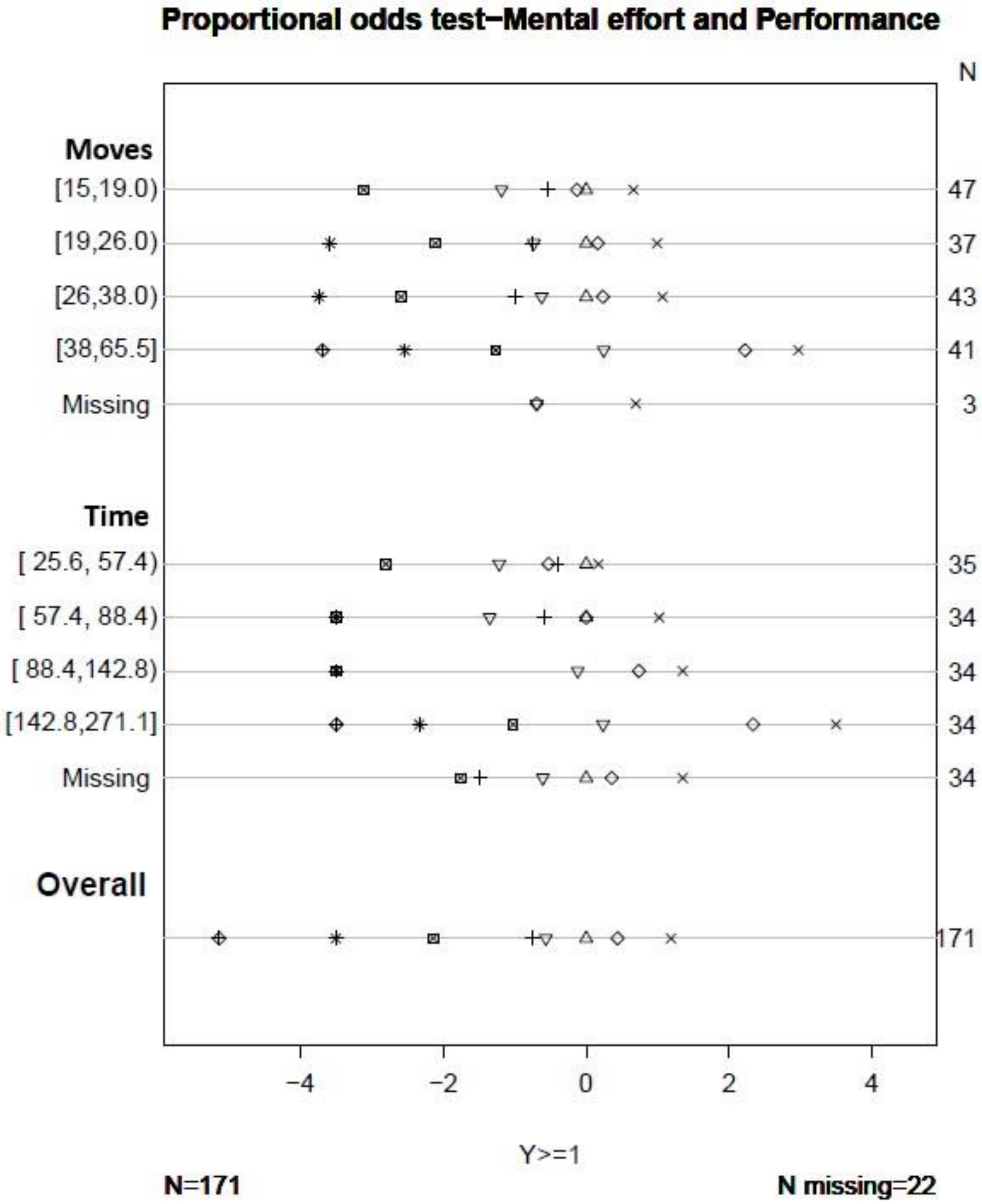


*Figure 30. Test of proportional odds assumption for mental effort and performance.*

The results from the ordinal logistic regression (Table 24) found no significant effect between mental effort and the behaviour variables. The odds ratio for moves (*OR* = 1.01) and time (OR = 1.01), indicate that for every one-unit increase or decrease in any of the variables, the odds of being more likely to have higher mental effort versus lower mental effort is negligible. These findings tentatively indicate that mental effort is not predictive of behaviour indicators during complex problem-solving performance.

*Table 24. Odds ratio data for mental effort and performance.*

| | | | | | 95% CI | |
|---|---|---|---|---|---|---|
| **Variable** | **Estimate** | **SE** | ***p*** | **OR** | **Lower** | **Upper** |
| Moves | 0.014 | 0.02 | 0.51 | 1.01 | 0.98 | 1.05 |
| Time | 0.013 | 0.00 | 0.003 | 1.01 | 1.01 | 1.02 |

SE = Standard error, *p* = p-value, OR = Odds ratio, CI = confidence interval

#### *5.2.4.6. Performance and behaviours*

Positive correlations were also found between the number of moves made to solve the problem and time r(155) = .78, p = .033, “indecision” r(155) = .61, p = .033, “stop” r(155) = .40, p = .033, “progress” r(155) = .56, p = .033, and behaviour time r(155) = .60, p = .033. Significant positive correlations were also found between time to solve the problem and “indecision” r(155) = 86, p = .033, “stop” r(155) = .75, p = .033, “progress” r(155) = .71, p = .033, and behaviour time r(155) = .93, p = .033. To further investigate this relationship regression analysis was explored. Given that the dependent variable (moves) is continuous, the assumptions of multiple regression analysis, a linear model, were tested. Initially the descriptive statistics for each of the variables were explored, detailed in Table 25. The Shapiro Wilk test was used to test the distribution of the variable residuals. The result was statistically significant *W* = .963, *p* > 0.05, (*p* = 0.059), indicating that the null hypothesis of normal distribution of residuals should be accepted.

*Table 25. Descriptive statistics for performance measures and time spent exhibiting behaviours.*

| | **N** | **M** | **SD** | **Med** | **Min** | **Max** | **Skew** | **Kurt** |
|---|---|---|---|---|---|---|---|---|
| Moves | 168 | 29.71 | 14.66 | 25.5 | 15 | 65.5 | 1.11 | 0.22 |
| Time | 137 | 110.72 | 72.65 | 88.06 | 25.57 | 271.1 | 1.07 | 0 |
| “Indecision” | 137 | 18.33 | 17.2 | 12.23 | 0 | 58.86 | 1.08 | 0.11 |
| “Stop” | 137 | 12.21 | 14.45 | 5.8 | 0 | 44.59 | 1.18 | 0.11 |
| “Progress” | 137 | 7.37 | 6.71 | 5.2 | 0 | 22.8 | 1 | -0.41 |
| Behaviour time | 137 | 38.5 | 33.51 | 26.6 | 0.36 | 107.8 | 0.93 | -0.41 |

N = sample size, M = Mean, SD = Standard deviation, Med = Median, Min = Minimum value, Max = Maximum value, Skew = Skewness, Kurt = Kurtosis

The assumption of linearity was tested through visual inspection. As detailed in Figure 31, there is no evidence to indicate a linear relationship between the variables. Further to this, there was multicollinearity between the independent variables as demonstrated through the significant positive correlations in Table 18. Given that multiple assumptions were breached, a logistic regression analysis was also conducted.

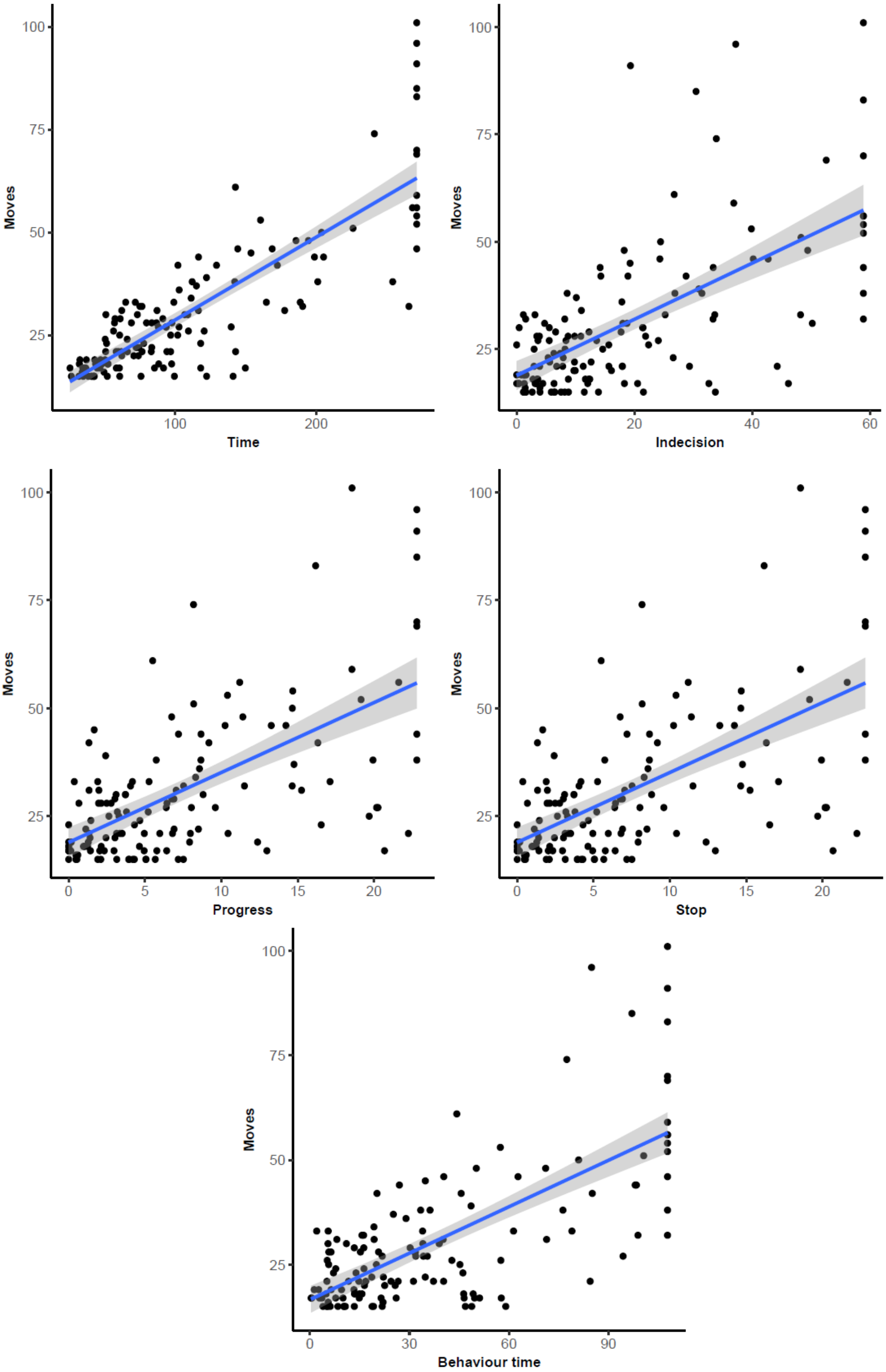


*Figure 31. Testing linearity between moves, time and time spent exhibiting behaviours individually and collectively.*

As with the spatial ability variable in section 5.2.4.1, the moves variable could be categorised through quartiles to facilitate an ordinal logistic regression. In doing this the dependent variable would be ordinal and the independent variables continuous. However, the assumption of multicollinearity was breached. Finally, the proportional odds assumption was tested. Through visual inspection of the graph in Figure 32, the alignment of the symbols for each set of categories of the dependent variable, moves, are not similar. This indicates that the proportional odds assumption for the ordinal logistic regression was violated.

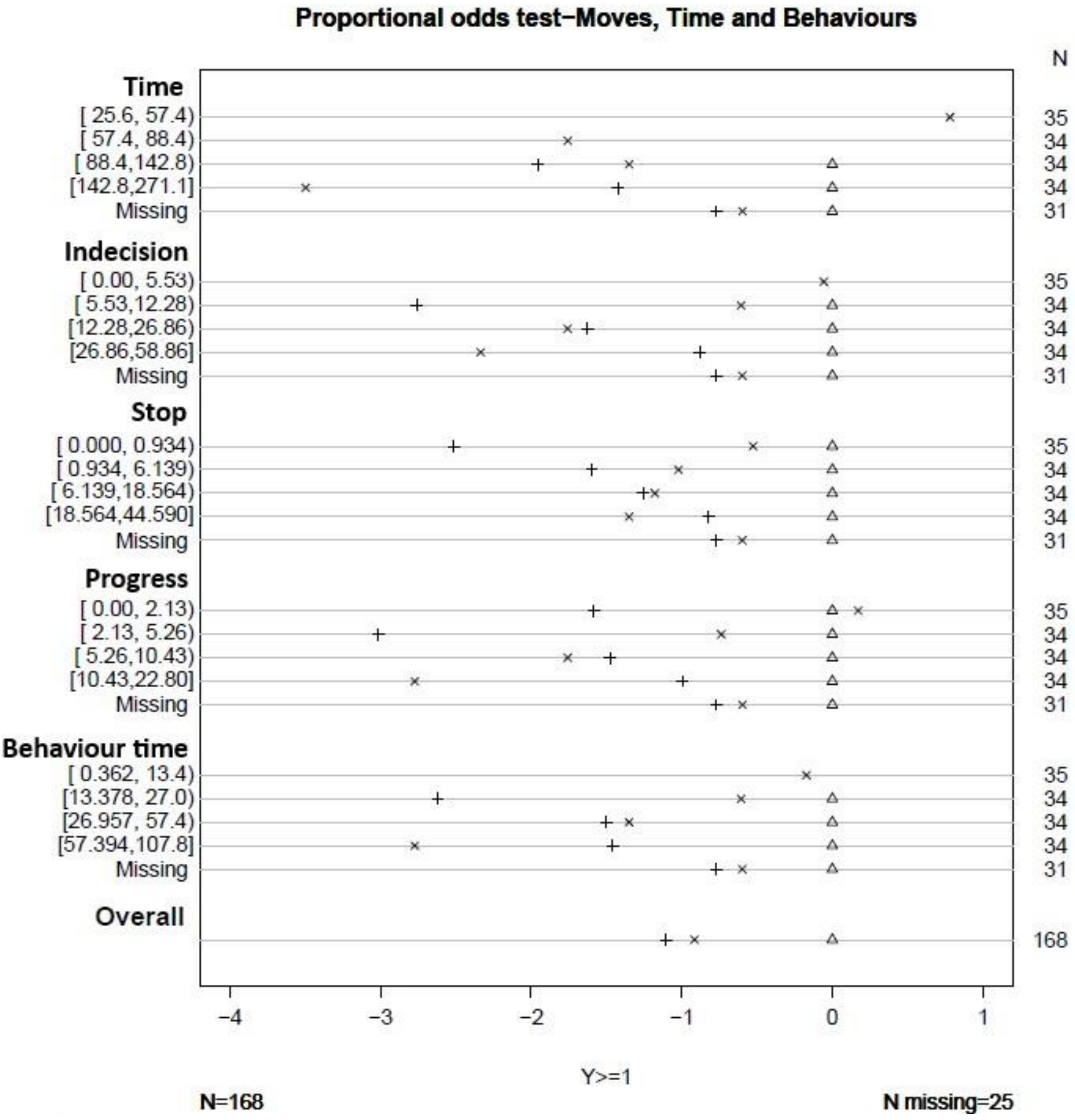


*Figure 32. Test of proportional odds assumption for performance and time spent exhibiting behaviours.*

Given that the assumptions of both regression models were violated, the results obtained using either model may not be valid. However, both models were implemented to determine what tentative insight they may provide. A multiple linear regression was used to determine whether moves were predictive of the time spent solving the problem, exhibiting "indecision", "stop" and "progress" themes, and overall spent exhibiting behaviours. Results indicated that there was a collective significant effect between "time", "indecision", "stop", "progress" and behaviour time, *F*(92.05, 5) = 131, $p < 0.001$, $R^2 = 0.78$, $R^2_{\text{adjusted}} = 0.77$. On an individual level neither "indecision" ($t = 0.228$, $p = 0.820$), "progress" ($t = 1.461$, $p = 0.146$) or "stop" ($t = 0.650$, $p = 0.517$) were found to be significant predictors in the model. The time spent solving the problem ($t = 42.884$, $p < 0.001$) and behaviour time ($t = -3.873$, $p < 0.001$) were found to be significant predictors in the model. As multiple assumptions of the model were violated, these outcomes cannot be reliably interpreted as a significant finding.

An ordinal logistic regression was conducted, with moves categorised in quartiles as previously described for evaluating the proportional odds assumption. The odds ratio for time (*OR* = 0.91), "indecision" (*OR* = 1.07), "progress" (OR = 1.08), "stop" (*OR* = 1.00) and behaviour time (*OR* = 1.07), outlined in Table 26, indicate that for every one-unit increase or decrease in any of the variables, the odds of being more likely to have a higher versus lower number of moves is negligible. These findings tentatively indicate that the number of moves to solve the problem is not predictive of time or behaviour indicators during complex problem-solving performance.

*Table 26. Odds ratio data for moves, time, and time spent exhibiting behaviours.*

| | | | | | **95% CI** | |
|---|---|---|---|---|---|---|
| **Variable** | **Estimate** | **SE** | ***p*** | **OR** | **Lower** | **Upper** |
| Time | -0.107 | 0.01 | 0.00 | 0.91 | 0.87 | 0.92 |
| "Indecision" | 0.072 | 0.08 | 0.39 | 1.07 | 0.92 | 1.22 |
| "Progress" | 0.002 | 0.09 | 0.98 | 1.08 | 0.92 | 1.24 |
| "Stop" | 0.075 | 0.09 | 0.40 | 1.00 | 0.85 | 1.16 |
| Behaviour time | 0.065 | 0.08 | 0.41 | 1.07 | 0.95 | 1.24 |

SE = Standard error, *p* = p-value, OR = Odds ratio, CI = confidence interval

In summary, the key empirical findings of the study are:

*Key Finding 1*

Domain general complex problem-solving performance was not found to differ across levels of engineering discipline expertise (as determined by year of study in an undergraduate degree programme) for students engaging in a representative engineering education programme.

*Key Finding 2*

Increased time spent exhibiting externalised behaviours throughout problem solving positively correlates with poorer performance in completing the problem i.e., increased number of moves and time.

*Key Finding 3*

Spatial ability relates to performance on a domain general complex problem, where increased spatial ability negatively correlates with both performance measures (i.e., overall time taken and number of moves).

*Key Finding 4*

Spatial ability relates to the overall time spent exhibiting behaviours and time spent displaying inductively themed behaviours of “indecision” and “progress”[1]. As spatial ability increased the time spent displaying these behaviours decreased.

*Key Finding 5*

Spatial ability relates to cognitive load experienced where, as spatial ability increased the levels of self-reported mental effort decreased.

*Key Finding 6*

Cognitive load relates to performance, where increased cognitive load correlated with poorer performance.

*Key Finding 7*

Cognitive load relates to behaviours, where increased cognitive load positively correlated with increased time exhibiting the inductively coded behaviour themes.

---

[1] Indecision behaviours: hesitation, reach and stop, gesturing, rule confirmation
Progress behaviours: hovering, trialling
Stop behaviours: pause, finger movements, disk play

These key empirical findings are graphically represented through Figure 33 below.

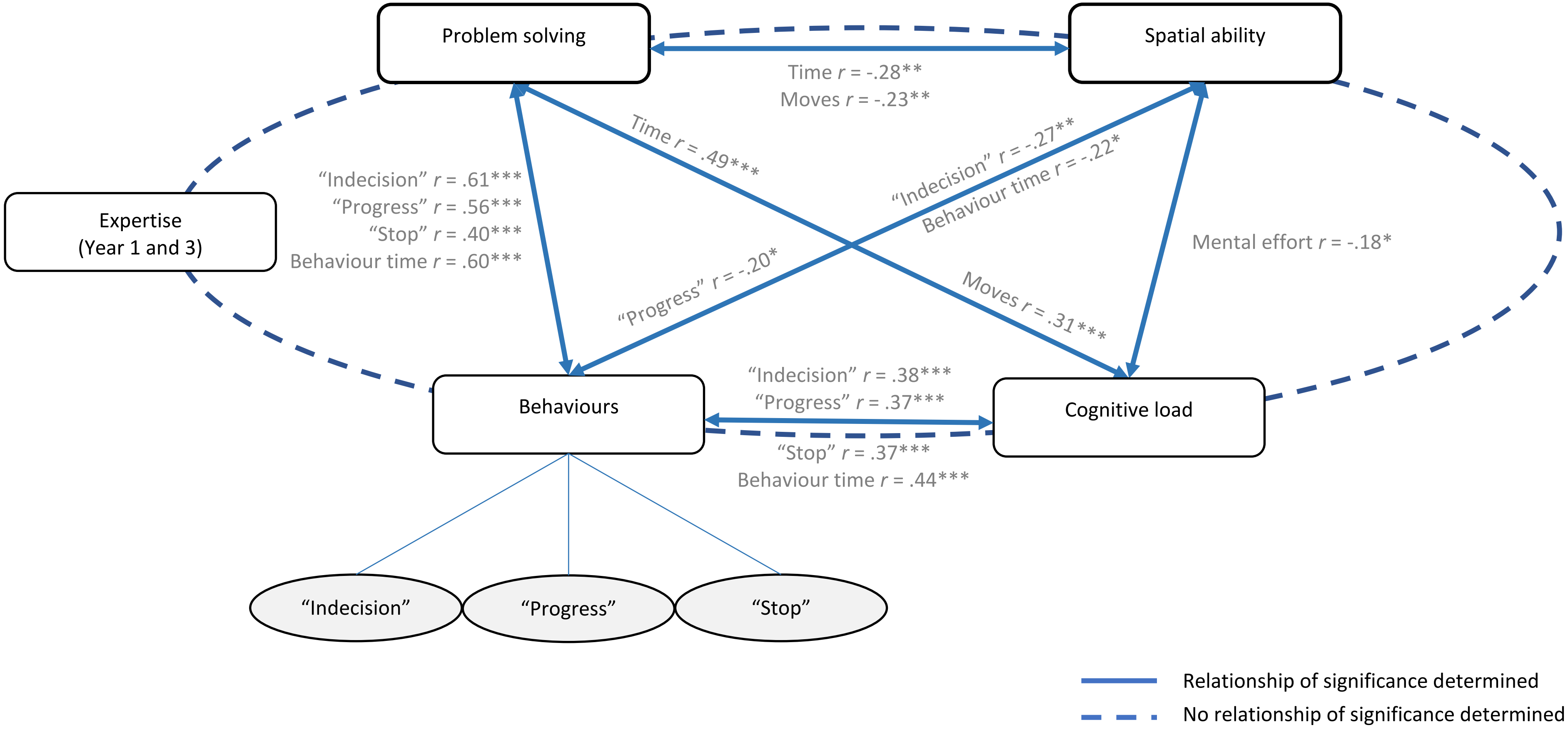


*Figure 33. Key empirical findings of relationships identified between core variables through the research.*

# 6. Discussion

## 6.1. Synopsis of study and overview of discussion

The overarching aim of this research study was to investigate the role of spatial ability in problem solving to contribute towards the understanding of a causal association between spatial ability and success in engineering education.

In attending to this aim the following research question was investigated:

> How does spatial ability's role in problem solving contribute to a broader causal theory between spatial ability and success in engineering education?

The key findings that resulted from investigating the outlined research questions will be discussed throughout this chapter under the headings of:

- Spatial ability as a critical cognitive factor for problem solving
- Behaviours and cognitive load in problem-solving
- Implications of the research for engineering education practice.

## 6.2. Spatial ability as a critical cognitive factor for problem solving

Within the CHC theory of human intelligence, spatial ability is outlined as facilitating the use of simulated mental imagery to solve problems by supporting perception, discrimination, and manipulation of that imagery in the "mind's eye" (Schneider & McGrew, 2018). Individuals with greater levels of spatial ability have a greater capacity to manipulate mental models and consider how components of a problem may relate to one another than individuals with lower levels of spatial ability.

This study found that levels of spatial ability had a significant relationship of medium effect size with domain general problem-solving performance indicators, where higher levels of spatial ability related to improved performance. This improved performance is of particular significance when we consider the requirement of graduate engineers to be adaptive to real-world engineering problems. This adaptive capability requires graduates to have the duality of both disciplined competences alongside domain general problem-solving skills to meet the evolving needs of current and future industrial activities. It is in this context that the correlative relationship that spatial skills, as a key cognitive factor, has with general problem-solving performance comes into clear focus. This specific finding of the study also underscores a testable and causal theory as to why spatial skills more broadly relate to success in engineering education due to the key role that problem solving plays within the overall goals of engineering education.

It was theorised in this research that spatial ability may support engineering students in managing the cognitive load experienced during problem solving. Through this study levels of spatial ability were also found to have a significant relationship of small effect size with the cognitive load experienced during problem solving where higher levels of spatial ability related to decreased ratings of cognitive load. There is minimal research that has previously explored the specific relationship between spatial ability and cognitive load levels during general problem-solving performance. A finding of this research suggests a means of supporting students in managing cognitive load to support successful problem solving and contributes towards further understanding for future investigations of a causal theory between spatial ability and educational success.

Previous research has demonstrated the use of alternative problem-solving strategies across levels of spatial ability. Higher visualisers have been found to employ holistic problem-solving approaches whereas lower visualisers may employ analytic or ‘piecemeal’ approaches to solve a problem (Khooshabeh et al., 2011; Lin, 2016; Tzuriel & Egozi, 2010). The limited spatial capacity of lower visualisers is viewed as impeding their ability to holistically manipulate mental imagery and therefore leads to them having to employ less optimal approaches. These limited approaches may amplify their inability to deal with more abstracted or out-of-context problems that are representative of real-world engineering activities.

Additionally, research investigating the role of behaviours in problem solving performance has demonstrated a relationship between visual working memory and the use of behaviours during problem solving where, for example, low visualisers are more likely to use gesturing to support problem solving (Cooperrider et al., 2015; Eielts et al., 2020; Handley et al., 2002; Pouw et al., 2016; Salthouse et al., 2003). This study found that higher levels of spatial ability had a significant negative relationship of medium effect size with the overall time spent exhibiting behaviours. At a more nuanced level, spatial ability also correlated with the time spent exhibiting behaviours categorised in the “indecision” and “progress” themes. The findings indicate that individuals with higher levels of spatial ability spend less time exhibiting these categories of observable behaviours which are interpreted as being indicative of piecemeal, less optimal approaches. Therefore, this outcome aligns with existing research on alternative problem-solving strategies and behaviours being exhibited by higher and lower visualisers, where higher visualisers may be able to implement more effective spatial strategies for problem solving. This emphasises that consideration should be given to the behaviours demonstrated during problem solving as a medium that either circumvents challenges or optimises the activities of problem-solving. This will help to advance understanding of the cognitive relationship between spatial skills and externalised behaviours as a function of engagement and performance in general problem-solving activities explicitly addressing the shortcomings identified by industry.

### 6.3. Behaviours and cognitive load in problem solving

Previous research evaluating the use of hand gestures in problem solving has suggested a capacity for these behaviours to free up cognitive resources during problem solving (Pouw et al., 2014). This work suggests that by using gesturing the load being placed on memory during information processing can be temporarily off-loaded by physically maintaining some information in the environment through a gesture (Eielts et al., 2020; Paas & van Merriënboer, 2020; Pouw et al., 2014).

Through this research significant correlations of large effect size were found between cognitive load and each of the behavioural themes of “indecision”, “stop”, and “progress” where higher levels of cognitive load experienced during problem solving were related to increased time spent displaying these themed behaviours. In essence, this finding indicates that; (1) the more mental effort an individual experiences when trying to solve a problem, the more time they will spend exhibiting behaviours which may be an approach to offload information to free up cognitive resources or (2) that externalisation of behaviours is not the most cognitively efficient strategy to use and therefore the levels of cognitive load are increased. This particular finding contributes to the understanding of the potential role of behaviours in cognitive resource management and demonstrates that observable behaviours may be indicative of an increase in the cognitive demands being experienced by an individual.

When an individual has not acquired the necessary schema to solve a problem, they can use alternative strategies such as heuristics to support them in reaching a solution (Gigerenzer, 2008). Demonstrating these behaviours or heuristics may be indicative of the individual not understanding the underlying parameters of the problem. This is significant in considering the capacity to acquire domain specific knowledge and skills and has implications for task and problem design in engineering education. Through this study significant correlations of very large effect size were found between the performance measures of moves and time and the time spent displaying the observed behaviours. This suggests that where poorer problem-solving performance is demonstrated, more time is spent displaying the behavioural themes of “indecision”, “stop”, and “progress”.

The exhibition of the observable behaviours during problem solving may be a result of individuals being uncertain as to how they should proceed with the problem. Based on the theoretical framework in this thesis (specifically the five principles of natural information processing systems (Sweller et al., 2011)) this finding infers that one of two of the following approaches were employed:

1. Randomness as genesis – where the problem solver has no existing schema to relate to or have not developed them at an earlier point in the problem, therefore, they randomly generate and test strategies to reach a solution
2. Environmental organising and linking – where the problem solver has developed strategies through interacting with the problem and when an obstacle is met they draw upon the previous strategies they employed to overcome this obstacle.

However, it is not possible to determine which approach was taken by individuals to move forward with the problem, but this is a significant insight into approaches for domain general learning. As these behaviours were more commonly displayed by learners with poorer performance in solving the problem, it highlights a useful indicator for educators to recognise when a learner requires additional instruction or support.

### *6.3.1. Cognitive load and performance*

The premise of cognitive load theory is to explain how the load imposed on memory during learning tasks can affect learners’ ability to process information effectively and transfer from the experience to long-term memory (Sweller et al., 2019). When considering cognitive load as a factor of a learner's educational performance during problem solving, the aim is to optimise intrinsic load and decrease extraneous load experienced to facilitate the learner in achieving the intended learning goals (Sweller et al., 2011). The instrument used in this study measured the overall cognitive load experienced and did not differentiate between intrinsic and extraneous cognitive load. However, the instructions for the TOH are simplistic and were carefully controlled so that they were identical for each participant. This was a key factor in the selection of this general, complex problem and the design of the method of how the participants engaged with the problem. This therefore both limited and normalised the amount of extraneous cognitive load experienced by the participants due to the instructional design of the problem. It was therefore inferred that the variance in cognitive load experienced across the participant group was due to variance in intrinsic load experienced by the participants. A significant correlation of large effect size was determined between the cognitive load experienced during problem solving and poorer performance on the domain general problem. It could be the case that the individuals had not acquired the necessary problem-solving schemas or strategies to solve the problem and therefore they experienced excessive intrinsic load which resulted in poorer performance. This is another key concern from the research that can be drawn upon to inform the design and overseeing of educational tasks in practice; where the research suggests that more robust problem-solving schema and strategies should be experienced while developing the necessary cognitive factors for success.

## 6.4. Implications of the research for engineering education practice

### *6.4.1. Perspectives on practice*

Existing research indicates that complex problem-solving skills are underdeveloped relative to industry needs in graduate engineers (Nair et al., 2009; Valentine et al., 2017; Ward & Thiriet, 2010). The World Economic Forum (2020) describe complex problem solving as reviewing information that is related to a problem to develop and evaluate options available to solve the problem. The capacity to successfully solve problems requires individuals to correctly identify the goals of the problem, be persistent, adopt efficient search strategies, regulate actions, and be able to trace back to previous points in the solution process (Litzinger et al., 2010; Y. Wang & Chiew, 2010). The problem used as a measure of complex problem-solving performance in this research, the TOH, required individuals to identify and manage sub-goals of the problem by recursively thinking ahead, considering the implications of their actions and regulate their actions to reach the solution (Schiff & Vakil, 2015). The findings of the research indicated no differences in complex problem-solving performance between individuals at the beginning and latter stages of their formal engineering education.

Evidence was found to suggest that spatial ability level influences complex problem-solving performance, behaviours exhibited during problem solving, and the cognitive load experienced during problem solving. However, no evidence was found to suggest differences in spatial ability level and years of experience in an engineering degree programme. Where spatial skills are a function of the threshold for effective engagement, lower visualisers will demonstrate inefficient problem-solving heuristics, while experiencing higher levels of cognitive load and reduced performance. Furthermore, the impact of having poorer spatial skills limits the capacity of individuals to acquire more discipline specific knowledge and skills but also the capacity to engage with conceptually broader problems.

This research has identified a malleable cognitive factor that relates to success and efficiency in complex problem solving. Therefore, spatial skills should be developed to support students in solving complex problems and contribute towards addressing the goal of developing domain general problem-solving capabilities. If the intention is to continue to develop effective problem-orientated, student-centred approaches in engineering education, key considerations must be made around integrating spatial skills development into associated pedagogical approaches and curriculum design.

### *6.4.2. Spatial skills in engineering curricula*

Although a significant gap remains in our understanding of the causal role of spatial ability in STEM achievement, the outcomes of this thesis combined with previous research indicates there is merit in incorporating spatial skills training into engineering education curricula. Spatial skills have been demonstrated to be malleable and can be developed through both direct and indirect approaches such as spatialising the curricula and semester-long training interventions (Julià & Antolí, 2016; Julià & Antolì, 2017; Mohler & Miller, 2008; Sorby, 2005; Sorby & Baartmans, 1996, 2000). As spatial ability is a critical cognitive factor that relates to successful and efficient problem solving, a need is highlighted for curriculum designers to consider the purposive and strategic integration of spatial skills training into engineering education programmes. This should not be done through superficial curricular objectives, but rather by meaningfully spatialising the curriculum with direct instructional approaches to cater for the development of spatial skills while also considering the specific problem design and tasks relevant to engaging students in more adaptive, context-rich and context-weak problems that rely on spatial skills for optimal engagement.

The findings from this research highlight the importance of an approach to (a) encourage students to employ spatial skills whilst problem solving and (b) to recognise that externalised behaviours can be used by individuals with lower spatial ability to possibly supplement them in building the spatial connections between the information provided. If educators were to be informed of the implications of observing student interactions during problem solving, they could (a) encourage students to use externalised behaviours to support them in overcoming deficiencies in their spatial skills, whilst (b) recognising that an over reliance on such behaviours may be indicative that higher levels of cognitive load are being experienced which can hinder their capacity to acquire the necessary discipline specific information and build robust problem-solving schema. This pedagogical awareness may take two forms (1) ensure engagement in knowledge transfer exercises that require the move from concrete (discipline specific) to abstract (domain general) through varying contexts (2) to offer specific instruction for the development of the student's spatial skills. This research indicates that the latter pedagogical intervention can lead to an increase in the capabilities of the student to more holistically and effectively approach problem solving.

As detailed through the findings and discussion in this thesis, this research highlights how spatial ability is a critical cognitive factor to be considered in supporting engineering students to solve complex problems. The following chapter will outline the key conclusions, contributions, limitations, and future work in relation to this study.

# 7. Conclusion

## 7.1. Conclusions

The following conclusions are presented in relation to each of the research objectives as defined in Chapter 1.

Objective 1: To examine performance on a domain general complex problem across levels of engineering student disciplined expertise.

> Conclusion 1: Through the analysis conducted in this thesis, domain general complex problem-solving performance was not found to differ across levels of engineering student disciplined expertise (as determined through progression).

Objective 2: To investigate if levels of spatial ability vary between engineering students with different levels of disciplined expertise.

> Conclusion 2: The research found no indication to suggest that spatial ability levels differed between engineering students in the first and final mandatory year of an engineering programme.

Objective 3: To examine if spatial ability influences engineering students' performance in general complex problem solving.

> Conclusion 3: Through this research, evidence was found that spatial ability relates to the performance of an individual on a complex general problem where higher spatial ability reduces the number of moves made and total time taken to solve the problem. Spatial ability was also found to have an association with behaviours during problem solving where higher levels of spatial ability were found to decrease the total time spent exhibiting behaviours; specifically behaviours in the "indecision" and "progress" themes.

Objective 4: To investigate whether levels of spatial ability affect the cognitive load experienced during problem solving.

> Conclusion 4: The work carried out in this thesis indicates that spatial ability significantly relates to the levels of cognitive load experienced during problem solving where, as spatial ability level increases the cognitive load experienced decreases.

## 7.2. Contributions

This research aimed to investigate the role of spatial ability in problem solving in engineering education. Understanding of the causal theory of spatial ability's relation to success in engineering education has been advanced. This research has indicated the following:

- Aligned with the critique from industry there is no evidence to suggest the development of complex problem-solving capability across the programme of study.
- Individuals with higher levels of spatial ability perform to a higher level when solving complex domain general problems.
- Higher levels of spatial ability also reduced the time spent exhibiting externalised behaviours and reduced the amount of cognitive load experienced during problem solving.
- The combination of these findings outline how higher spatial ability can positively contribute to students' performance and educational experiences during problem solving.
- As problem solving is a fundamental component of engineering education practice, it is posited that the role of spatial ability in problem solving contributes to the causal relationship between spatial ability and success in engineering education.

## 7.3. Limitations

The limitations of the work are acknowledged as follows:

- This is a single site study, that although showed significant causal evidence, it cannot be generalised. Further studies are required that control variables such as task, curriculum, pedagogy, and environment across multiple contexts.
- The study may have benefitted from having two separate groups within the participant cohort where variables of interest were manipulated across the groups e.g., time, problem difficulty, experience with the problem, spatial ability. This could have supported the formation of more specific inferences being made around the threshold of spatial ability for problem-solving performance.

- Access to the population post analysis to engage with confirmatory interviews could have added a rich qualitative dimension to the research. However, this was not possible due to the current global pandemic.
- Although differences are reported in spatial ability between males and females (Doyle et al., 2012; Feng et al., 2007; Metz & Sorby, 2016; S. A. Sorby, 1999; S. A. Sorby et al., 2013; Vandenberg & Kuse, 1978; Xu et al., 2016; Yilmaz, 2009), an analysis of problem-solving performance and spatial ability across genders was not conducted. This was due to the lack of a representative sample to conduct such an analysis. Given the underrepresentation of females in engineering this limitation was anticipated but unfortunately unavoidable.

### 7.4. Recommendations and future work

Through the course of this research tentative implications for spatial skills development in engineering education practice have been examined and several areas requiring further research have emerged.

Two recommendations for engineering education practice are:

1. Integrating spatial skills training into engineering education programmes to facilitate the development of a core cognitive factor for problem solving in engineering.
2. There is a necessity for engineering educators to be aware of the significance of observable behaviours as indicators of internal cognitive processes during problem solving so that they can optimise the pedagogical support that they offer individual students.

Areas requiring additional research to inform practice include:

- Evaluating whether problem-solving capacities are being developed through engagement in formal engineering education at the site of the primary study would require further research through a longitudinal study to confirm the findings in this research.
- If the findings that complex problem-solving capacities are not being developed are confirmed, it would be necessary to conduct further research examining whether these capacities are being developed through alternative pedagogical

and curricular approaches that use both curricular learning outcomes and direct intervention for the development of problem-solving skills.

- Examining whether spatial skills training can improve complex problem-solving performance. This research indicated that spatial ability relates to problem-solving performance, therefore, it is necessary to explore whether engagement in a spatial skills training intervention may support improved performance in complex problem solving and the students that most benefit from such an intervention.
- Investigating the association between spatial ability and cognitive load in problems of increasing complexity and limited time. These investigations would inform understanding of the role of spatial ability in cognitive resource management during problem solving.
- Exploring the role of behaviours in representing internal cognitive processes in the external world. This research advanced on the awareness of behaviours exhibited when engaging with the TOH problem. It is necessary that further research is conducted to understand these behaviours from the perspective of the problem solver to determine the inferences that can be made around these.

# Bibliography

Abdi, H. (2007). Z-scores. *Encyclopedia of measurement and statistics*, 3, 1055-1058.

Akera, A. (2017). Setting the standards for engineering education: A history. *Proceedings of the IEEE*, *105*(9), 1834–1843. https://doi.org/10.1109/JPROC.2017.2729406

Amodio, P., del Piccolo, F., Marchetti, P., Angeli, P., Iemmolo, R., Caregaro, L., Merkel, C., Gerunda, G., & Gatta, A. (1999). Clinical Features and Survivial of Cirrhotic Patients With Subclinical Cognitive Alterations Detected by the Number Connection Test and Computerized Psychometric Tests. *Hepatology*, *29*(6), 1662–1667.

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. In *Educational Psychology Review, 22*(4), 425–438. https://doi.org/10.1007/s10648-010-9130-y

Anwar, A. A., & Richards, D. J. (2018). Comparison of EC and ABET Accreditation Criteria. *Journal of Professional Issues in Engineering Education and Practice*. https://doi.org/10.1061/(ASCE)

Ardies, J., de Maeyer, S., Gijbels, D., & van Keulen, H. (2015). Students attitudes towards technology. *International Journal of Technology and Design Education*, *25*(1), 43–65. https://doi.org/10.1007/s10798-014-9268-x

Ayres, P., Lee, J. Y., Paas, F., & van Merriënboer, J. J. (2021). The Validity of Physiological Measures to Identify Differences in Intrinsic Cognitive Load. *Frontiers in Psychology*, 12.

Bailey, K. A., & Gammage, K. L. (2020). Applying Action Research in a Mixed Methods Positive Body Image Program Assessment With Older Adults and People With Physical Disability and Chronic Illness. *Journal of Mixed Methods Research*, *14*(2), 248–267. https://doi.org/10.1177/1558689819871814

Bailey, S. K. T., Neigel, A. R., Dhanani, L. Y., & Sims, V. K. (2018). Establishing Measurement Equivalence Across Computer- and Paper-Based Tests of Spatial

Cognition. *Human Factors The Journal of the Human Factors and Ergonomics Society*, *60*(3), 340–350. https://doi.org/10.1177/0018720817747731

Balaji, K. V. A., & Somashekar, S. (2009). A Comparative Study of Soft Skills among Engineers. *The IUP Journal of Soft Skills*, *3 & 4*(3), 50–57.

Ball, C., Huang, K. T., Rikard, R. V, & Cotten, S. R. (2019). The emotional costs of computers: an expectancy-value theory analysis of predominantly low-socioeconomic status minority students' STEM attitudes. *Information Communication and Society*, *22*(1), 105–128. https://doi.org/10.1080/1369118X.2017.1355403

Barak, M. (2013). Teaching engineering and technology: Cognitive, knowledge and problem-solving taxonomies. In *Journal of Engineering, Design and Technology, 11*(3), 316–333. https://doi.org/10.1108/JEDT-04-2012-0020

Barkaoui, K. (2011). Think-aloud protocols in research on essay rating: An empirical study of their veridicality and reactivity. *Language Testing*, *28*(1), 51–75. https://doi.org/10.1177/0265532210376379

Barlex, D., & Steeg, T. (2017). *Re-building Design & Technology In the secondary school curriculum version 2: A working paper*.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., Gijselaers, W., & Westendorp, J. (2008). Cognitive load measurements and stimulated recall interviews for studying the effects of information and communications technology. *Educational Technology Research and Development*, *56*(3), 309–328. https://doi.org/10.1007/s11423-006-9020-7

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science*, *21*(11), 1605–1610. https://doi.org/10.1177/0956797610385353

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Björklund, T. A. (2013). Initial mental representations of design problems: Differences between experts and novices. *Design Studies*, *34*(2), 135–160. https://doi.org/10.1016/j.destud.2012.08.005

Blasko, D. G., & Holliday-Darr, K. A. (2010). Longitudinal Analysis of Spatial Skills Training in Engineering Graphics. *Proceedings of the 65th Midyear Meeting of the Engineering Design Graphics Division*, 138–151. https://www.researchgate.net/publication/266170923

Blums, A., Belsky, J., Grimm, K., & Chen, Z. (2017). Building Links Between Early Socioeconomic Status, Cognitive Ability, and Math and Science Achievement. *Journal of Cognition and Development*, *18*(1), 16–40. https://doi.org/10.1080/15248372.2016.1228652

Bodner, G. M., & Guay, R. B. (1997). The Purdue Visualization of Rotations Test. *The Chemical Educator*, *2*(4), 1–17. https://doi.org/10.1007/s00897970138a

Boucsein, W., Fowles, D. C., Grimnes, S., Ben-Shakhar, G., Roth, W. T., Dawson, M. E., & Filion, D. L. (2012). Publication recommendations for electrodermal measurements. *Psychophysiology*, *49*(8), 1017–1034. https://doi.org/10.1111/j.1469-8986.2012.01384.x

Braithwaite, J., Watson, D., Robert, J., & Mickey, R. (2013). A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments. *Psychophysiology*, *49*(1), 1017–1034. https://doi.org/10.1017.S0142716405050034

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist*, *38*(1), 53–61. https://doi.org/10.1207/S15326985EP3801_7

Buckley, J. (2020). The need to consider the predictive capacity of intelligence and its malleability within design and technology education research. *International*

*Journal of Technology and Design Education*, 1–15. https://doi.org/10.1007/s10798-020-09588-9

Buckley, J., Seery, N., & Canty, D. (2018a). A Heuristic Framework of Spatial Ability: a Review and Synthesis of Spatial Factor Literature to Support its Translation into STEM Education. *Educational Psychology Review*, *30*(3), 947–972. https://doi.org/10.1007/s10648-018-9432-z

Buckley, J., Seery, N., & Canty, D. (2018b). Heuristics and CAD modelling: an examination of student behaviour during problem solving episodes within CAD modelling activities. *International Journal of Technology and Design Education*, *28*(4), 939–956. https://doi.org/10.1007/s10798-017-9423-2

Buckley, J., Seery, N., & Canty, D. (2019). Investigating the use of spatial reasoning strategies in geometric problem solving. *International Journal of Technology and Design Education*, *29*(2), 341–362. https://doi.org/10.1007/s10798-018-9446-3

Burgess, R. A., Davis, M., Dyrud, M. A., Herkert, J. R., Hollander, R. D., Newton, L., Pritchard, M. S., & Vesilind, P. A. (2013). *Engineering Ethics: Looking Back, Looking Forward*.

Burt, C. (1949). The structure of the mind. *Educational Psychology*, *19*(3), 176–199.

Caeiro-Rodriguez, M., Manso-Vazquez, M., Mikic-Fonte, F. A., Llamas-Nistal, M., Fernandez-Iglesias, M. J., Tsalapatas, H., Heidmann, O., de Carvalho, C. V., Jesmin, T., Terasmaa, J., & Sorensen, L. T. (2021). Teaching Soft Skills in Engineering Education: An European Perspective. *IEEE Access*, *9*, 29222–29242. https://doi.org/10.1109/ACCESS.2021.3059516

Caicedo, J. M., Berge, N., Madarshahian, R., Timmerman, B., Pierce, C. E., Flora, J. R. v, Berge, N. D., & Madarshahian, R. (2014). Integrating Professional and Technical Engineering Skills with the EFFECTs Pedagogical Framework. *International Journal of Engineering Education*, *30*(6(B)), 1579–1589. https://www.researchgate.net/publication/278964690

Cain, R., & Lee, V. R. (2020). Measuring Electrodermal Activity in an Afterschool Maker Program to Detect Youth Engagement. In R. Z. Zheng (Ed.), *Cognitive and*

*Affective Perspectives on Immersive Technology in Education,* 128–150. https://doi.org/10.4018/978-1-7998-3250-8.ch007

Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics*, *58*(1), 45–75. https://doi.org/10.1007/s10649-005-0808-x

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies* (1st ed.). Cambridge University Press. https://doi.org/10.1080/15332276.1998.11672894

Carroll, J. B. (1997). The three-stratum theory of cognitive abilities. In *In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues.* New York, NY, US: The Guilford Press.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*(3), 153–193. https://doi.org/10.1111/j.1939-0025.1939.tb05653.x

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1–22. https://doi.org/10.1037/h0046743

Cattell, R. B., & Horn, J. L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, *15*(3), 139–164.

CDIO. (2010). *THE CDIO STANDARDS v 2.0 (with customized rubrics)*. http://www.cdio.org

Chang, Y. (2014). 3D-CAD effects on creative design performance of different spatial abilities students. *Journal of Computer Assisted Learning*, *30*(5), 397–407. https://doi.org/10.1111/jcal.12051

Chang, Y. S., Chien, Y. H., Lin, H. C., Chen, M. Y., & Hsieh, H. H. (2016). Effects of 3D CAD applications on the design creativity of students with different representational abilities. *Computers in Human Behavior*, *65*, 107–113. https://doi.org/10.1016/j.chb.2016.08.024

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

Chen, J., Kolmos, A., & Du, X. (2021). Forms of implementation and challenges of PBL in engineering education: a review of literature. *European Journal of Engineering Education*, *46*(1), 90–115. https://doi.org/10.1080/03043797.2020.1718615

Chen, O., & Kalyuga, S. (2020). Cognitive Load Theory, Spacing Effect, and Working Memory Resources Depletion: Implications for Instructional Design. In S. Hai-Jew (Ed.), *Form, function, and style in instructional design: Emerging research and opportunities,* 1–26.

Chi, M. T. H., Glaser, R., & Rees, E. (1981). *Expertise in problem solving*. Pittsburgh Univ PA Learning Research and Development Center.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. The Journal of Abnormal and Social Psychology, 65(3), 145–153. https://doi.org/10.1037/h0045186

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (8th ed.). Routledge. https://doi.org/10.1111/j.1467-8527.2007.00388_4.x

Colman, B., & Willmot, P. (2016). How Soft Are ‘Soft Skills’ in the Engineering Profession? *44th Annual SEFI Conference: Engineering Education on Top of the World: Industry University Cooperation.*

Cooperrider, K., Wakefield, E., & Goldin-Meadow, S. (2015). More than Meets the Eye: Gesture Changes Thought, even without Visual F eedback. *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, 441–446.

Crawley, E. F., Malmqvist, J., Östlund, S., Brodeur, D. R., & Edström, K. (2014). *Rethinking Engineering Education: The CDIO Approach, Second Edition,* (2nd ed.). https://doi.org/10.1007/978-3-319-05561-9

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). SAGE Publications Inc. https://doi.org/10.1017/CBO9781107415324.004

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed). Thousand Oaks, CA: Sage.

Cunningham, C. M., & Kelly, G. Y. J. (2017). Epistemic Practices of Engineering for Education. *Science Education*, *101*(3), 486–505. https://doi.org/10.1002/sce.21271

de Vries, M. J. (2016). Teaching about Technology. *Teaching about Technology*. https://doi.org/10.1007/978-3-319-32945-1

Dixit, R. J., & Pathak, M. (2012). An Overview of International Engineering Accords with special reference to Washington Accord. *Journal of Engineering, Science & Management Education*, *5*(2), 467–471.

Dixon, R. A., & Johnson, S. D. (2011). Experts vs. Novices: Differences in How Mental Representations Are Used in Engineering Design. *Journal of Technology Education, 23*(1), 47–65.

Dörner, D., & Funke, J. (2017). Complex problem solving: What it is and what it is not. *Frontiers in Psychology*, *8*, 1–11. https://doi.org/10.3389/fpsyg.2017.01153

Doyle, R. A., Voyer, D., & Cherney, I. D. (2012). The relation between childhood spatial activities and spatial abilities in adulthood. *Journal of Applied Developmental Psychology*, *33*(2), 112–120. https://doi.org/10.1016/j.appdev.2012.01.002

Dyrenfurth, M., & Murphy, M. (2006). Understanding the European Bologna Process. *American Society for Engineering Education*.

Edström, K. (2016). Aims of Engineering Education Research - The role of the CDIO Initiative. *Proceedings of the 12th International CDIO Conference, Turku University of Applied Sciences*.

Edström, K., & Kolmos, A. (2014). PBL and CDIO: complementary models for engineering education development. *European Journal of Engineering Education*, *39*(5), 539–555. https://doi.org/10.1080/03043797.2014.895703

Eielts, C., Pouw, W., Ouwehand, K., van Gog, T., Zwaan, R. A., & Paas, F. (2020). Co-thought gesturing supports more complex problem solving in subjects with lower visual working-memory capacity. *Psychological Research*, *84*(2), 502–513. https://doi.org/10.1007/s00426-018-1065-9

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for Kit of Factor-Referenced Cognitive Tests. *Educational Testing Service*.

Engineers Ireland. (2021). *The State of Ireland 2020: Engineering a green and digital recovery*. www.engineersireland.ie

Ericsson, K. A. (2006). The Influence of Experience and Deliberate Practice on the Development of Superior Expert Performance. In *The Cambridge handbook of expertise and expert performance,* 38, 685–705.

Ericsson, K. A. (2008). Deliberate Practice and Acquisition of Expert Performance: A General Overview. *Academic Emergency Medicine*, *15*(11), 988–994. https://doi.org/10.1111/J.1553-2712.2008.00227.X

Faber, C., & Benson, L. C. (2017). Engineering Students' Epistemic Cognition in the Context of Problem Solving. *Journal of Engineering Education*, *106*(4), 677–709. https://doi.org/10.1002/jee.20183

Feilzer, M. Y. (2010). Doing Mixed Methods Research Pragmatically: Implications for the Rediscovery of Pragmatism as a Research Paradigm. *Journal of Mixed Methods Research*, *4*(1), 6–16. https://doi.org/10.1177/1558689809349691

Felder, R. M. (2012). Engineering Education: A Tale of Two Paradigms. In B. McCabe, M. Pantazidou, & D. Phillips (Eds.), *Shaking the Foundations of Geo-Engineering Education,* 9–14. Leiden: CRC Press.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*(10), 850–855. https://doi.org/10.1111/j.1467-9280.2007.01990.x

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs - Principles and practices. *Health Services Research*, *48*(6 PART2), 2134–2156. https://doi.org/10.1111/1475-6773.12117

Fetters, M. D., & Freshwater, D. (2015). The 1 + 1 = 3 Integration Challenge. In *Journal of Mixed Methods Research* (Vol. 9, Issue 2, pp. 115–117). SAGE Publications Inc. https://doi.org/10.1177/1558689815581222

Fischer, A., Greiff, S., & Funke, J. (2011). The Process of Solving Complex Problems. *The Journal of Problem Solving*, *4*(1), 19–42.

Flening, E., Asplund, F., & Edin Grimheden, M. (2021). Measuring professional skills misalignment based on early-career engineers' perceptions of engineering expertise. *European Journal of Engineering Education*, 1–27. https://doi.org/10.1080/03043797.2021.1967883

Flick, U. (2014). *An Introduction to Qualitative Research* (5th ed.). Sage. https://doi.org/10.1057/9780230239517_1

Friedman, A., Kohler, B., Gunalp, P., Boone, A. P., & Hegarty, M. (2020). A computerized spatial orientation test. *Behavior Research Methods*, *52*(2), 799–812. https://doi.org/10.3758/s13428-019-01277-3

Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. Advances in Methods and Practices in Psychological Science, 2(2), 156–168. https://doi.org/10.1177/2515245919847202

Funke, J. (2010). Complex problem solving: a case for complex cognition? *Cognitive Processing*, *11*(2), 133–142. https://doi.org/10.1007/s10339-009-0345-0

Funke, J. (2012). Complex problem solving. In *Encyclopedia of the Sciences of Learning* (Issue 3, pp. 682–685). Springer.

Galton, F. (1879). Psyhcometric experiments. *Brain*, *2*(2), 149–162.

Galton, F. (1880). Visualised numerals. *Nature*, *21*(533), 252–256.

Gaughran, W. (2002). Cognitive modelling for engineers. *ASEE Annual Conference Proceedings*, 10569–10581. https://doi.org/10.18260/1-2--10547

Gigerenzer, G. (2001). The Adaptive Toolbox. In G. Gigerenzer & R. Selten (Eds.), *Bounded Rationality - The Adaptive Toolbox* (36–50). The MIT Press.

http://pubman.mpdl.mpg.de/pubman/item/escidoc:2102373/component/escidoc:2102372/GG_Adaptive_Toolbox_2001.pdf

Gigerenzer, G. (2008). Why Heuristics Work. *Perspectives on Psychological Science*, *3*(1), 20–29. https://doi.org/10.1111/j.1745-6916.2008.00058.x

Gjoreski, M., Luštrek, M., Gams, M., & Gjoreski, H. (2017). Monitoring stress with a wrist device using context. *Journal of Biomedical Informatics*, *73*, 159–170. https://doi.org/10.1016/j.jbi.2017.08.006

Gomez Puente, S. M., Eijck, van, M. W., & Jochems, W. M. G. (2014). Exploring the effects of design-based learning characteristics on teachers and students. *International Journal of Engineering Education*, *30*(4), 916–928.

Gómez Puente, S. M., Van Eijck, M., & Jochems, W. (2013). Empirical validation of characteristics of design-based learning in higher education. *International Journal of Engineering Education*, *29*(2), 491–503.

Gómez Puente, S. M., van Eijck, M., & Jochems, W. (2015). Professional development for design-based learning in engineering education: a case study. *European Journal of Engineering Education*, *40*(1), 14–31. https://doi.org/10.1080/03043797.2014.903228

Gravetter, F. J., & Wallnau, L. B. (2014). Introduction to the t statistic. In *Essentials of statistics for the behavioral sciences* (8th ed.).

Greiff, S., & Funke, J. (2009). *Measuring Complex Problem Solving: The MicroDYN approach*. https://www.researchgate.net/publication/242072342

Guay, R. B. (1976). *Purdue Spatial Visualization Tests*.

Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Annals of Family Medicine*, *13*(6), 554–561. https://doi.org/10.1370/afm.1865

Hadgraft, R. G., & Kolmos, A. (2020). Emerging learning environments in engineering education. *Australasian Journal of Engineering Education*, *25*(1), 3–16. https://doi.org/10.1080/22054952.2020.1713522

Hair Jr., J., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate Data Analysis* (7th ed.). Pearson Education Limited.

Hambrick, D. Z., Altmann, E. M., & Burgoyne, A. P. (2018). A knowledge activation approach to testing the circumvention-of-limits hypothesis. *American Journal of Psychology*, *131*(3), 307–321. https://doi.org/10.5406/amerjpsyc.131.3.0307

Hambrick, D. Z., Libarkin, J. C., Petcovic, H. L., Baker, K. M., Elkins, J., Callahan, C. N., Turner, S. P., Rench, T. A., & LaDue, N. D. (2012). A test of the circumvention-of-limits hypothesis in scientific problem solving: The case of geological bedrock mapping. *Journal of Experimental Psychology: General*, *141*(3), 397–403. https://doi.org/10.1037/a0025927

Hambrick, D. Z., Macnamara, B. N., Campitelli, G., Ullén, F., & Mosing, M. A. (2016). Beyond Born versus Made: A New Look at Expertise. *Psychology of Learning and Motivation - Advances in Research and Theory*, *64*, 1–55. https://doi.org/10.1016/bs.plm.2015.09.001

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the predictive power of domain-specific experience and knowledge in skilled performance. *Current Directions in Psychological Science*, *20*(5), 275–279. https://doi.org/10.1177/0963721411422061

Handley, S. J., Capon, A., Copp, C., & Harper, C. (2002). Conditional reasoning and the Tower of Hanoi: The role of spatial and verbal working memory. *British Journal of Psychology*, *93*(4), 501–518. https://doi.org/10.1348/000712602761381376

Hanney, R., & Savin-Baden, M. (2013). The problem of projects: understanding the theoretical underpinnings of project-led PBL. *London Review of Education*, *11*(1), 7–19. https://doi.org/10.1080/14748460.2012.761816

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

Harwood, J. (2006). Engineering education between science and practice: rethinking the historiography. *History and Technology*, *22*(1), 53–79. https://doi.org/10.1080/07341510500497210

Hatta, T., Narita, K., Yanagihara, K., Ishiguro, H., Murayama, T., & Yokode, M. (2020). Crossover Mixed Analysis in a Convergent Mixed Methods Design Used to Investigate Clinical Dialogues About Cancer Treatment in the Japanese Context. *Journal of Mixed Methods Research*, *14*(1), 84–109. https://doi.org/10.1177/1558689818792793

Henriques, R., Paiva, A., & Antunes, C. (2013). On the need of new methods to mine electrodermal activity in emotion-centered studies. In L. Cao, Y. Zeng, A. L. Symeonidis, V. I. Gorodetsky, P. S. Yu, & M. P. Singh (Eds.), *Agents and data mining interaction, 7607*, 203–215). Springer. https://doi.org/10.1007/978-3-642-36288-0_18

Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C., & Roberts, W. (2001). Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *128*(4), 719–734. https://doi.org/10.1016/S1095-6433(01)00278-1

Horn, J. (1985). Remodelling old models of intelligence. In B. Wolman (Ed.), *Handbook of Intelligence: Theories, Measurements, and Applications* (pp. 267–300). Wiley.

Horn, J. (1988). Thinking about human abilities. In J. Nesselroade & R. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 645–685). Academic Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*(5), 253–270.

Jaarsveld, S., & Lachmann, T. (2017). Intelligence and Creativity in Problem Solving: The Importance of Test Features in Cognition Research. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00134

Jamil, M., Tariq, R. H., Shami, P. A., & Univetsity, Z. (2012). Computer-based vs Paper-based examinations: Perceptions of University teachers. *TOJET: The Turkish Online Journal of Educational Technology*, *11*(4), 371–381.

Jenkinson, J. (2009). Measuring the Effectiveness of Educational Technology: what are we Attempting to Measure? *Electronic Journal of E-Learning*, *7*(3), 273–280.

Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educational Technology Research and Development*. https://doi.org/10.1007/BF02299613

Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. *Educational Technology*, *63*, 63–85.

Jonassen, D. H. (2010). *Learning to Solve Problems: A Handbook for Designing Problem-Solving Learning Environments*. Routledge.

Jonassen, D. H. (2015). Engineers as problem solvers. In A. Johri & B. M. Olds (Eds.), *Cambridge Handbook of Engineering Education Research* (pp. 103–118). Cambridge University Press. https://doi.org/10.1017/CBO9781139013451.009

Jonassen, D. H., & Hung, W. (2008). *All Problems are Not Equal: Implications for Problem-Based All Problems are Not Equal: Implications for Problem-Based Learning Learning*. 6–28. https://doi.org/10.7771/1541-5015.1080

Jonassen, D. H., & Hung, W. (2015). All Problems Are Not Equal : Implications for Problem-Based Learning. In A. Walker, H. Leary, C. Hmelo-Silver, & P. A. Ertmer (Eds.), *Essential Readings in Problem-Based Learning* (pp. 17–42). Purdue University Press: West Lafayette, Indiana.

Jonassen, D. H., Strobel, J., & Lee, C. B. (2006). Everyday problem solving in engineering: Lessons for engineering educators. *Journal of Engineering Education*, *95*(2), 139–151. https://doi.org/10.1002/j.2168-9830.2006.tb00885.x

Julià, C., & Antolí, J. Ò. (2016). Spatial ability learning through educational robotics. *International Journal of Technology and Design Education*, *26*(2), 185–203. https://doi.org/10.1007/s10798-015-9307-2

Julià, C., & Antolì, J. Ò. (2017). Enhancing Spatial Ability and Mechanical Reasoning through a STEM Course. In *International Journal of Technology and Design Education* (Vol. 28, Issue 4, pp. 957–983).

Kalyuga, S. (2011). Cognitive Load Theory: How Many Types of Load Does It Really Need? *Educational Psychology Review*, *23*(1), 1–19. https://doi.org/10.1007/s10648-010-9150-7

Kalyuga, S., & Plass, J. L. (2017). Cognitive Load as a Local Characteristic of Cognitive Processes: Implications for Measurement Approaches. In R. Z. Zheng (Ed.), *Cognitive load measurement and application: A theoretical framework for meaningful research and practice* (pp. 59–74). Routledge. https://doi.org/10.4324/9781315296258-5

Kalyuga, S., & Singh, A.-M. (2016). Rethinking the Boundaries of Cognitive Load Theory in Complex Learning. *Educational Psychology Review*, *28*(4). https://doi.org/10.1007/s10648-015-9352-0

Kamaruzaman, F. M., Hamid, R., Mutalib, A. A., & Rasul, M. S. (2019). Comparison of Engineering Skills with IR 4.0 Skills . *International Journal of Biomedical Engineering*, *15*(10), 15–28. https://doi.org/10.3991/ijoe.v15i10.10879

Keighrey, C., Flynn, R., Murray, S., & Murray, N. (2020). A Physiology-based QoE Comparison of Interactive Augmented Reality, Virtual Reality and Tablet-based Applications. *IEEE Transactions on Multimedia*, 1–9. https://doi.org/10.1109/tmm.2020.2982046

Keith, N., & Ericsson, K. A. (2007). A Deliberate Practice Account of Typing Proficiency in Everyday Typists. *Journal of Experimental Psychology: Applied*, *13*(3), 135–145. https://doi.org/10.1037/1076-898X.13.3.135

Kell, H. J., & Lubinski, D. (2013). Spatial Ability: A Neglected Talent in Educational and Occupational Settings. *Roeper Review*, *35*(4), 219–230. https://doi.org/10.1080/02783193.2013.829896

Khooshabeh, P., Hegarty, M., & Shipley, T. F. (2011). Individual Differences in Mental Rotation : Piecemeal vs Holistic Processes. *Proceedings of the Annual Meeting of the Cognitive Science Society* .

Kimbell, R., Stables, K., Wheeler, T., Wosniak, A., & Kelly, V. (1991). *The Assessment of Performance in Design and Technology: The Final Report of the APU Design and Technology Project 1985-1991*.

Kimbell, R., Wheeler, T., Miller, S., Bain, J., Wright, R., & Stables, K. (2004). *Assessing Design Innovation final report*. *December*.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, *21*(1), 31–42. https://doi.org/10.1007/s10648-008-9095-2

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. R. (2018). From Cognitive Load Theory to Collaborative Cognitive Load Theory. *International Journal of Computer-Supported Collaborative Learning*, *13*(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y

Kivunja, C., & Kuyini, A. B. (2017). Understanding and Applying Research Paradigms in Educational Contexts. *International Journal of Higher Education*, *6*(5), 26. https://doi.org/10.5430/ijhe.v6n5p26

Kozhevnikov, M., Motes, M. A., & Hegarty, M. (2007). *Spatial Visualization in Physics Problem Solving*. *31*, 549–579.

Kozhevnikov, M., & Thornton, R. (2006). Real-Time Data Display, Spatial Visualization Ability, and Learning Force and Motion Concepts. *Journal of Science Education and Technology*, *15*(1), 111–132. https://doi.org/10.1007/s10956-006-0361-0

Kulasegaram, K. M., Grierson, L. E. M., & Norman, G. R. (2013). The roles of deliberate practice and innate ability in developing expertise: evidence and implications. *Medical Education*, *47*(10), 979–989. https://doi.org/10.1111/MEDU.12260

Lantada, A. D., Bayo, A. H., & Sevillano, J. D. J. M. (2014). Promotion of Professional Skills in Engineering Education: Strategies and Challenges. *International Journal of Engineering Education*, *30(B)*(6), 1–14.

Leandro Cruz, M., Saunders-Smits, G. N., & Groen, P. (2020). Evaluation of competency methods in engineering education: a systematic review. *European Journal of Engineering Education*, *45*(5), 729–757. https://doi.org/10.1080/03043797.2019.1671810

Leeson, H. v. (2006). The Mode Effect: A Literature Review of Human and Technological Issues in Computerized Testing. *International Journal of Testing*, *6*(1), 1–24. https://doi.org/10.1207/s15327574ijt0601_1

Leppink, J. (2017). Cognitive load theory: Practical implications and an important challenge. *Journal of Taibah University Medical Sciences*, *12*(5), 385–391. https://doi.org/10.1016/j.jtumed.2017.05.003

Leppink, J., Paas, F., van der Vleuten, C. P. M., van Gog, T., & van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Leppink, J., Paas, F., van Gog, T., van der Vleuten, C. P. M., & van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, *30*, 32–42. https://doi.org/10.1016/j.learninstruc.2013.12.001

Leppink, J., & van Merrienboer, J. J. G. (2015). The Beast of Aggregating Cognitive Load Measures in Technology-Based Learning. *Educational Technology & Society*, *18*(4), 230–245.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*, 764–766. https://doi.org/10.1016/J.JESP.2013.03.013

Li, J., & Fu, S. (2012). A Systematic Approach to Engineering Ethics Education. *Science and Engineering Ethics*, *18*(2), 339–349. https://doi.org/10.1007/s11948-010-9249-8

Lin, H. (2016). Influence of design training and spatial solution strategies on spatial ability performance. *International Journal of Technology and Design Education*, *62*, 123–131. https://doi.org/10.1007/s10798-015-9302-7

Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex Differences in Mental Rotation and Line Angle Judgments Are Positively Associated with Gender Equality and Economic Development Across 53 Nations. *Archives of Sexual Behavior*, *39*(4), 990–997. https://doi.org/10.1007/s10508-008-9460-8

Litzinger, T. A., Vanmeter, P., Firetto, C. M., Passmore, L. J., Masters, C. B., Turns, S. R., Gray, G. L., Costanzo, F., & Zappe, S. E. (2010). A cognitive study of problem solving in statics. *Journal of Engineering Education*, *99*(4), 337–353. https://doi.org/10.1002/j.2168-9830.2010.tb01067.x

Litzinger, T. A., Vanmeter, P., Firetto, C. M., Passmore, L. J., Masters, C. B., Turns, S. R., Gray, G. L., Costanzo, F., & Zappe, S. E. (2011). Engineering Education and the Development of Expertise. *Journal of Engineering Education*, *100*(1), 123–150. https://doi.org/10.1002/J.2168-9830.2011.TB00006.X

Liu, Y., & Du, S. (2018). Psychological stress level detection based on electrodermal activity. *Behavioural Brain Research*, *341*, 50–53. https://doi.org/10.1016/j.bbr.2017.12.021

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence, 4*, 181–248. Lawrence Erlbaum Associates, Inc.

Lohman, D. F. (1989). *Human Intelligence: An Introduction to Advances in Theory and Research, 59*(4). http://rer.aera.net

Lohman, D. F. (1994). Spatial Ability. *In R. J. Sternberg (Ed.), Encyclopaedia of Intelligence. 2, 1000–1007. New York, NY: Macmillan.*

Lohman, D. F. (1996). Spatial Ability and G. *Human Abilities: Their Nature and Assessment* , 97–116.

Lotz, C., Sparfeldt, J. R., & Greiff, S. (2016). Complex problem solving in educational contexts – Still something beyond a “good g”? *Intelligence*, *59*, 127–138. https://doi.org/10.1016/j.intell.2016.09.001

Lubinski, D. (2004). Introduction to the Special Section on Cognitive Abilities: 100 Years After Spearman’s (1904) &quot;’General Intelligence,’ Objectively Determined and Measured&quot; *Article in Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.86.1.96

Lubinski, D. (2010). Spatial ability and STEM: A sleeping giant for talent identification and development. *Personality and Individual Differences*, *49*, 344–351. https://doi.org/10.1016/j.paid.2010.03.022

Mackenzie, N., & Knipe, S. (2006). Research Dilemmas: Paradigms, methods and methodology. *Issues in Educational Research*, *16*(2), 1–13.

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions A Meta-Analysis. *Psychological Science*, *25*(8), 1608–1618. https://doi.org/10.1177/0956797614535810

Maeda, Y., & Yoon, S. Y. (2016). Are Gender Differences in Spatial Ability Real or an Artifact? Evaluation of Measurement Invariance on the Revised PSVT:R. *JOURNAL OF PSYCHOEDUCATIONAL ASSESSMENT*, *34*(4), 397–403. https://doi.org/10.1177/0734282915609843

Maeda, Y., Yoon, S. Y., Kim-Kang, G., & Imbrie, P. K. (2013). Psychometric properties of the revised PSVT:R for measuring First Year Engineering students’ spatial ability. *International Journal of Engineering Education*, *29*(3), 763–776.

Mainali, B. P. (2012). Higher Order Thinking in Education. *Academic Voices: A Multidisciplinary Journal*, *2*(1), 5–10. https://doi.org/10.3126/AV.V2I1.8277

Malmqvist, J. (2012). A comparison of the CDIO and EUR-ACE quality assurance systems. *International Journal of Quality Assurance in Engineering and Technology Education (IJQAETE)*, *2*(2), 9–22.

Malmqvist, J., Edström, K., Gunnarsson, S., & Östlund, S. (2006). The application of CDIO Standards in the evaluation of Swedish engineering degree programmes. *World Transactions on Engineering and Technology Education*, *5*(2), 361–364.

Malmqvist, J., Östlund, S., & Edström, K. (2006). Using integrated programme descriptions to support a CDIO programme design process. *World Transactions on Engineering and Technology Education*, *5*(2), 259–262.

Martin-Dorta, N., Saorin, J. L., de La, J., Cant, T., Melián Díaz, D., & Meier, C. (2009). Do Video Games Improve Spatial Abilities of Engineering Students. *International Journal of Engineering Education*, *25*(6), 1194–1204.

Martín-Gutiérrez, J., Contero, M., & Alcañiz, M. (2015). Augmented Reality to Training Spatial Skills. *Procedia Computer Science*, *77*, 33–39. https://doi.org/10.1016/j.procs.2015.12.356

Martín-Gutiérrez, J., & González, M. M. A. (2017). Ranking and predicting results for different training activities to develop spatial abilities. In *Visual-spatial Ability in STEM Education* (pp. 225–239). Springer, Cham. https://doi.org/10.1007/978-3-319-44385-0

Mayer, R. E., & Sims, V. K. (1994). For Whom Is a Picture Worth a Thousand Words? Extensions of a Dual-Coding Theory of Multimedia Learning. In *Journal of Educational Psychology* (Vol. 86, Issue 3).

Mayer, R. E., & Wittrock, M. C. (2006). Problem solving. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 287–304). New York: Macmillan.

McDonald, A. S. (2002). The impact of individual differences on the equivalence of computer-based and paper-and-pencil educational assessments. *Computers & Education*, 39(3), 299-312.

McGill, R. J., & Dombrowski, S. C. (2019). Critically Reflecting on the Origins, Evolution, and Impact of the Cattell-Horn-Carroll (CHC) Model. *Applied Measurement in Education*, *32*(3), 216–231. https://doi.org/10.1080/08957347.2019.1619561

McGrath, D. (2000). The Bologna Declaration and Engineering Education in Europe. *Institution of Engineers in Ireland Available at Mie. Uth. Gr.*

McGrew, K. S. (2005). *The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future*.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Meehl, P. (2006). The power of quantitative thinking. In *In N. Waller, L. Yonce, W. Grove, D. Faust, & M. Lenzenweger (Eds.), A Paul Meehl reader: Essays on the practice of scientific psychology.* (pp. 433–444).

Meinz, E. J., Hambrick, D. Z., Hawkins, C. B., Gillings, A. K., Meyer, B. E., & Schneider, J. L. (2012). Roles of domain knowledge and working memory capacity in components of skill in Texas Hold'Em poker. *Journal of Applied Research in Memory and Cognition*, *1*(1), 34–40. https://doi.org/10.1016/j.jarmac.2011.11.001

Metz, S., Sorby, S. A., & Jarosewich, T. (2016). Spatial skills training impacts retention of engineering students - Does this success translate to community college students in technical education? *Proceedings of the American Society for Engineering Education Annual Conference and Exposition, New Orleans, LA*. https://doi.org/10.18260/p.25853

Mitchell, J. E., & Smith, J. (2008). Case study of the introduction of problem-based learning in electronic engineering. *International Journal of Electrical Engineering Education*, *45*(2), 131–143.

Mohler, J. L. (2008). A review of spatial ability research. *Engineering Design Graphics Journal*, *72*(3), 19–30.

Mohler, J. L., & Miller, C. L. (2008). Improving Spatial Ability with Mentored Sketching. *The Engineering Design Graphics Journal*, *72*(1), 19–27.

Moreno, M., & Guidetti, M. (2018). Do we plan through gestures? Evidence from children, adolescents and adults in solving of Tower of Hanoï Task. *Universitas Psychologica*, *17*(2). https://doi.org/10.11144/Javeriana.upsy17-2.pgec

Mowbray, F. I., Fox-Wasylyshyn, S. M., & El-Masri, M. M. (2019). Univariate outliers: A conceptual overview for the nurse researcher. *Canadian Journal of Nursing Research*, 51(1), 31-37.

Murray, J. K., Studer, J. A., Daly, S. R., McKilligan, S., & Seifert, C. M. (2019). Design by taking perspectives: How engineers explore problems. *Journal of Engineering Education*, *108*(2), 248–275. https://doi.org/10.1002/jee.20263

Nair, C. S., Patil, A., & Mertova, P. (2009). Re-engineering graduate skills-a case study. *European Journal of Engineering Education*, *34*(2), 131–139. https://doi.org/10.1080/03043790902829281

Newcombe, N. (2017). Harnessing Spatial Thinking to Support Stem Learning. *OECD Education Working Papers*, *161*, 1–51. https://doi.org/10.1787/7d5dcae6-en

Newcombe, N. S., & Shipley, T. F. (2015). Thinking About Spatial Thinking: New Typology, New Assessments. In *Studying Visual and Spatial Reasoning for Design Creativity,* 179–192. Springer Netherlands. https://doi.org/10.1007/978-94-017-9297-4_10

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-hall.

Nourbakhsh, N., Wang, Y., Chen, F., & Calvo, R. A. (2012). Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. *Proceedings of the 24th Australian Computer-Human Interaction Conference*, 420–423. https://doi.org/10.1145/2414536.2414602

Olkun, S. (2003). Making connections improving spatial abilities with engineering drawing activities. *International Journal for Mathematics Teaching and Learning*, 1–10. https://doi.org/10.1501/0003624

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2014). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*, *47*(4), 1343–1355. https://doi.org/10.3758/s13428-014-0543-2

Ouwehand, K., Kroef, A. van der, Wong, J., & Paas, F. (2021). Measuring Cognitive Load: Are There More Valid Alternatives to Likert Rating Scales? *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.702616

Paas, F. (1992). Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach. *Journal of Educational Psychology*, *84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, *38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, *32*, 1–8. https://doi.org/10.1023/b:truc.0000021806.17516.d0

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Paas, F., van Merrienboer, J., & Darabi, &. (2005). A Motivational Perspective on the Relation between Mental Effort and Performance: Optimizing Learner Involvement in Instruction. *Etr&D-Educational Technology Research and Development*, *53*(3), 25–34. https://doi.org/10.1007/BF02504795

Paas, F., & van Merriënboer, J. J. G. (1994). Variability of Worked Examples and Transfer of Geometrical Problem-Solving Skills: A Cognitive-Load Approach. *Journal of Educational Psychology*, *86*(1), 122–133. https://doi.org/10.1037/0022-0663.86.1.122

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-Load Theory: Methods to Manage Working Memory Load in the Learning of Complex Tasks. *Current Directions in Psychological Science*, *29*(4), 394–398. https://doi.org/10.1177/0963721420922183

Paletta, L., Pittino, N. M., Schwarz, M., Wagner, V., & Kallus, K. W. (2015). Human Factors Analysis Using Wearable Sensors in the Context of Cognitive and Emotional Arousal. *Procedia Manufacturing*, *3*, 3782–3787. https://doi.org/10.1016/j.promfg.2015.07.880

Paskins, Z., Sanders, T., Croft, P. R., & Hassell, A. B. (2017). Exploring the Added Value of Video-Stimulated Recall in Researching the Primary Care Doctor–Patient Consultation: A Process Evaluation. *International Journal of Qualitative Methods*, *16*(1), 1–11. https://doi.org/10.1177/1609406917719623

Passow, H. J., & Passow, C. H. (2017). What Competencies Should Undergraduate Engineering Programs Emphasize? A Systematic Review. *Journal of Engineering Education*, *106*(3), 475–526. https://doi.org/10.1002/jee.20171

Patil, A., Codner, & G., & Codner, G. (2007). Accreditation of engineering education: review, observations and proposal for global accreditation. *European Journal of Engineering Education*, *32*(6), 639–651. https://doi.org/10.1080/03043790701520594

Patsenko, E. G., & Altmann, E. M. (2010). How Planful Is Routine Behavior? A Selective-Attention Model of Performance in the Tower of Hanoi. *Journal of Experimental Psychology: General*, *139*(1), 95–116. https://doi.org/10.1037/a0018268

Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse mental rotations test-different versions and factors that affect performance. *Brain and Cognition*, *28*(1), 39–58.

Poh, M.-Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assesment of electrodermal activity. *IEEE Transactions on Biomedical Engineering*, *57*(5), 1243–1252. https://doi.org/10.1109/tbme.2009.2038487

Posada-Quintero, H. F., & Chon, K. H. (2020). Innovations in electrodermal activity data collection and signal processing: A systematic review. In *Sensors* (Vol. 20, pp. 1–18). https://doi.org/10.3390/s20020479

Pouw, W. T. J. L., de Nooijer, J. A., van Gog, T., Zwaan, R. A., & Paas, F. (2014). Toward a more embedded/extended perspective on the cognitive function of gestures. *Frontiers in Psychology*, *5*(APR), 1–14. https://doi.org/10.3389/fpsyg.2014.00359

Pouw, W. T. J. L., Mavilidi, M. F., van Gog, T., & Paas, F. (2016). Gesturing during mental problem solving reduces eye movements, especially for individuals with lower visual working memory capacity. *Cognitive Processing*, *17*(3), 269–277. https://doi.org/10.1007/s10339-016-0757-6

Prados, J. W., Peterson, G. D., & Lattuca, L. R. (2005). Quality assurance of engineering education through accreditation: The impact of engineering criteria 2000 and its global influence. *Journal of Engineering Education*, *94*(1), 165–184. https://doi.org/10.1002/j.2168-9830.2005.tb00836.x

Prokop, M., Pilař, L., & Tichá, I. (2020). Impact of think-aloud on eye-tracking: A comparison of concurrent and retrospective think-aloud for research on decision-making in the game environment. *Sensors*, *20*, 1–21. https://doi.org/10.3390/s20102750

Ramey, K. E., & Uttal, D. H. (2017). Making Sense of Space: Distributed Spatial Sensemaking in a Middle School Summer Engineering Camp. *Journal of the Learning Sciences*, *26*(2), 277–319. https://doi.org/10.1080/10508406.2016.1277226

Reid, C., Keighrey, C., Murray, N., Dunbar, R., & Buckley, J. (2020). A novel mixed methods approach to synthesize EDA data with behavioral data to gain educational insight. *Sensors (Switzerland)*, *20*(23), 1–14. https://doi.org/10.3390/s20236857

Reynolds, T. S., & Seely, B. E. (1993). Striving for Balance: A Hundred Years of the American Society for Engineering Education. *Journal of Engineering Education*, *82*(3), 136–151. https://doi.org/10.1002/j.2168-9830.1993.tb00092.x

Roca-González, C., Martin-Gutierrez, J., García-Dominguez, M., & Carrodeguas, M. del C. M. (2017). Virtual technologies to develop visual-spatial ability in engineering students. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(2), 441–468. https://doi.org/10.12973/eurasia.2017.00625a

Rudolph, J., Greiff, S., Strobel, A., & Preckel, F. (2018). Understanding the link between need for cognition and complex problem solving. *Contemporary Educational Psychology*, *55*, 53–62. https://doi.org/10.1016/j.cedpsych.2018.08.001

Saleh, H., & Lamsali, H. (2020). Fundamental General Skills And Engineering Skills As An Important Skills For Engineering Graduates Employability: A Fundamental Study. *International Journal of Scientific & Technology Research*, *9*(2), 3370–3373. www.ijstr.org

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive Functioning as a Potential Mediator of Age-Related Cognitive Decline in Normal Adults. *Journal of Experimental Psychology: General*, *132*(4), 566–594. https://doi.org/10.1037/0096-3445.132.4.566

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality and Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Savin-Baden, M. (2008). Problem-Based Learning in Electronic Engineering: Locating Legends or Promising Problems? *International Journal of Electrical Engineering Education*, *45*(2), 96–109. https://doi.org/10.7227/IJEEE.45.2.2

Savin-baden, M. (2014). Using Problem-Based Learning : New Constellations for the 21 st Century. *Journal on Excellence in College Teaching*, *25*, 197–219.

Schäfer, T., & Schwarz, M. A. (2019). The Meaningfulness of Effect Sizes in Psychological Research: Differences Between Sub-Disciplines and the Impact of Potential Biases. Frontiers in Psychology, 10, 813. https://doi.org/10.3389/fpsyg.2019.00813

Schiff, R., & Vakil, E. (2015). Age differences in cognitive skill learning, retention and transfer: The case of the Tower of Hanoi Puzzle. *Learning and Individual Differences*, *39*, 164–171. https://doi.org/10.1016/J.LINDIF.2015.03.010

Schmidt, M. W., Kowalewski, K. F., Trent, S. M., Benner, L., Müller-Stich, B. P., & Nickel, F. (2020). Self-directed training with e-learning using the first-person perspective for laparoscopic suturing and knot tying: a randomised controlled trial. *Surgical endoscopy*, 34(2), 869-879.

Schneider, J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll (CHC) Model of Intelligence. In *Contemporary Intellectual Assessment: Theories, Tests, and Issues.* (3rd ed., pp. 99–144).

Schneider, J., & McGrew, K. S. (2018). The Cattell-Horn-Carroll theory of cognitive abilities. In D. Flanagan & E. McDonough (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues.* (pp. 73–163). New York: The Guilford Press.

Schoenfeld, A. H. (1983). The Wild, Wild, Wild, Wild, Wild World of Problem Solving (A Review of Sorts). *For the Learning of Mathematics*, *3*(3), 40–47.

Schraw, G. (2001). Promoting General Metacognitive Awareness. In *Metacognition in learning and instruction: Theory, research and practice.* (Vol. 19, pp. 3–16). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-2243-8_1

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting Self-Regulation in Science Education: Metacognition as Part of a Broader Perspective on Learning. *Research in Science Education*, *36*, 111–139. https://doi.org/10.1007/s11165-005-3917-8

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology*. https://doi.org/10.1002/acp.2350090605

SEFI. (2016). *Developing Graduate Engineering Skills*.

Setz, C., Arnrich, B., Schumm, J., Marca, R. la, Tr, G., & Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable EDA device. *Technology*, *14*(2), 410–417. https://doi.org/10.1109/TITB.2009.2036164

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics Made Easy: An Effort-Reduction Framework. *Psychological Bulletin*, *134*(2), 207–222. https://doi.org/10.1037/0033-2909.134.2.207

Shearman, R. (2007). Bologna: Engineering the right outcomes. *International Journal of Electrical Engineering Education*, *44*(2), 97–100.

Sheppard, S. D., Macatangay, K., Colby, A., & Sullivan, W. M. (2008). *Educating Engineers: Designing for the Future of the Field*. www.josseybass.com.

Shuman, L. J., Besterfield-Sacre, M., & McGourty, J. (2005). The ABET “Professional Skills” - Can They be Taught? Can They be Assessed? *Journal of Engineering Education*, *94*, 41–55.

Simarro, C., & Couso, D. (2021). Engineering practices as a framework for STEM education: a proposal based on epistemic nuances. *International Journal of STEM Education*, *8*(53), 1–12. https://doi.org/10.1186/s40594-021-00310-2

Solhjoo, S., Haigney, M. C., McBee, E., van Merrienboer, J. J. G., Schuwirth, L., Artino, A. R., Battista, A., Ratcliffe, T. A., Lee, H. D., & Durning, S. J. (2019). Heart Rate and Heart Rate Variability Correlate with Clinical Reasoning Performance and Self-Reported Measures of Cognitive Load. *Scientific Reports*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-50280-3

Son, J., & Park, M. (2011). Estimating cognitive load complexity using performance and physiological data in a driving simulator. *Adjunct Proceedings of the Automotive User Interfaces and Interactive Vehicular Applications Conference*, November 29-Decmber 2. http://www.auto-ui-cogload.unh.edu/wp-content/uploads/2011/11/clw2011_son_park.pdf

Sorby, S. A. (1999). Developing 3-D Spatial Visualization Skills. *Engineering Design Graphics Journal*, *63*(2), 21–32.

Sorby, S. A. (2005). Impact of Changes in Course Methodologies on Improving Spatial Skills. *Journal for Geometry and Graphics*, *9*(1), 99–105.

Sorby, S. A. (2007). Developing 3D spatial skills for K-12 students. *Australasian Journal of Engineering Education*, *13*(1), 1–11.

Sorby, S. A. (2009). Educational Research in Developing 3-D Spatial Skills for Engineering Students. *International Journal of Science Education*, *31*(3), 459–480. https://doi.org/10.1080/09500690802595839

Sorby, S. A., & Baartmans, B. J. (1996). A course for the Development of 3-D Spatial Visualization Skills. *Engineering Design Graphics Journal*, *60*(1), 13–20.

Sorby, S. A., & Baartmans, B. J. (2000). The Development and Assessment of a Course for Enhancing the 3-D Spatial Visualization Skills of. *Journal of Engineering Education*, *89*(3), 301–307.

Sorby, S. A., Casey, B. M., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences*, *26*, 20–29. https://doi.org/10.1016/j.lindif.2013.03.010

Sorby, S. A., Drummer, T., & Molzon, R. (2006). Experiences in using spatial skills testing instruments with younger audiences. *Journal for Geometry and Graphics*, *10*(2), 227–235.

Sorby, S. A., Nevin, E., Mageean, E., Sheridan, S., & Behan, A. (2014). Initial investigations into spatial skills as predictors of success in first-year STEM programmes. *SEFI Annual Conference 2014*, 111–117. https://doi.org/10.1109/

Sorby, S. A., & Veurink, N. (2019). Preparing for STEM: Impact of spatial visualization training on middle school math performance. *Journal of Women and Minorities in Science and Engineering*, *25*(1), 1–23. https://doi.org/10.1615/JWomenMinorScienEng.2018024516

Spearman, C. (1904). “General Intelligence,” objectively determined and measured. *The American Journal of Psychology.*, *15*(2), 201–292.

Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction*, *17*(2), 219–234. https://doi.org/10.1016/j.learninstruc.2007.01.012

Stieff, M., & Uttal, D. (2015). How Much Can Spatial Training Improve STEM Achievement? *Educational Psychology Review*, *27*(4), 607–615. https://doi.org/10.1007/s10648-015-9304-8

Strobel, J., & Barneveld, A. van. (2015). PBL Effectiveness , Tensions , and Practitioner Implications. In A. Walker, H. Leary, C. Hmelo-Silver, & P. A. Ertmer (Eds.), *Essential Readings in Problem-Based Learning* (pp. 355–372). Purdue University Press: West Lafayette, Indiana.

Strobel, J., & van Barneveld, A. (2009). When is PBL More Effective? A Meta-synthesis of Meta-analyses. *Interdisciplinary Journal of Problem-Based Learning*, *3*(1), 44–58. https://doi.org/10.7771/1541-5015.1046

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123–138.

Sweller, J. (2015). In Academe, What Is Learned, and How Is It Learned? *Current Directions in Psychological Science*, *24*(3), 190–194. https://doi.org/10.1177/0963721415569570

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory. In J. M. Spector & S. P. Lajoie (Eds.), *Cognitive Load Theory*. Springer. https://doi.org/10.1017/CBO9780511844744

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. In *Educational Psychology Review* (Vol. 31, Issue 2, pp. 261–292). https://doi.org/10.1007/s10648-019-09465-5

Thammasan, N., Stuldreher, I. v., Schreuders, E., Giletta, M., & Brouwer, A.-M. (2020). A usability study of physiological measurement in school using wearable sensors. *Sensors*, *20*, 1–24. https://doi.org/10.3390/s20185380

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago University Press.

Tricot, A., & Sweller, J. (2014). Domain-Specific Knowledge and Why Teaching Generic Skills Does Not Work. In *Educational Psychology Review* (Vol. 26, Issue 2, pp. 265–283). https://doi.org/10.1007/s10648-013-9243-1

Trochim, W., & Donnelly, J. (2006). *The research methods knowledge*.

Trofatter, C., Kontra, C., Beilock, S., & Goldin-Meadow, S. (2015). Gesturing has a larger impact on problem-solving than action, even when action is accompanied by words. *Language, Cognition and Neuroscience*, *30*(3), 251–260. https://doi.org/10.1080/23273798.2014.905692

Tzuriel, D., & Egozi, G. (2010). Gender Differences in Spatial Ability of Young Children: The Effects of Training and Processing Strategies. *Child Development*, *81*(5), 1417–1430.

UNESCO. (2010). Engineering: Issues, Challenges and Opportunities for Development. In *UNESCO Report*. http://unesdoc.unesco.org/images/0018/001897/189753e.pdf

Uttal, D. H., & Cohen, C. A. (2012). Spatial Thinking and STEM Education. When, Why, and How? In *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 57, pp. 147–181). https://doi.org/10.1016/B978-0-12-394293-7.00004-2

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139*(2), 352–402. https://doi.org/10.1037/a0028446

Uttal, D. H., Miller, D. I., & Newcombe, N. S. (2013). Exploring and Enhancing Spatial Thinking: Links to Achievement in Science, Technology, Engineering, and Mathematics? *Current Directions in Psychological Science*, *22*(5), 367–373. https://doi.org/10.1177/0963721413484756

Valentine, A., Belski, I., & Hamilton, M. (2017). Developing creativity and problem-solving skills of engineering students: a comparison of web- and pen-and-paper-based approaches. *European Journal of Engineering Education*, *42*(6), 1309–1329. https://doi.org/10.1080/03043797.2017.1291584

van Barneveld, A., & Strobel, J. (2011). Reports from teaching practice: Experiences and management of tensions encountered with PBL implementations in the early years of undergraduate engineering education. *Proceedings of the Research in Engineering Education Symposium*. https://www.researchgate.net/publication/286179547

van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, *43*(1), 16–26. https://doi.org/10.1080/00461520701756248

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*(2), 147–177. https://doi.org/10.1007/s10648-005-3951-0

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*(2), 599–604. https://doi.org/10.2466/pms.1978.47.2.599

Vernon, P. (1950). *The structure of human abilities*. Wiley.

Veurink, N. L., & Hamlin, A. J. (2011). Spatial visualization skills: Impact on confidence and success in an engineering curriculum. *2011 ASEE Annual Conference & Exposition*.

Veurink, N., & Sorby, S. A. (2012). Comparison of Spatial Skills of Students Entering Different Engineering Majors. *Engineering Design Graphics Journal (EDGJ) Copyright*, *76*(3). http://www.edgj.org

Villanueva, I., Campbell, B. D., Raikes, A. C., Jones, S. H., & Putney, L. A. G. (2018). A Multimodal Exploration of Engineering Students’ Emotions and Electrodermal Activity in Design Activities. *Journal of Engineering Education*, *107*(3), 414–441. https://doi.org/10.1002/jee.20225

Villanueva, I., Husman, J., Christensen, D., Youmans, K., Khan, M. T., Vicioso, P., Lampkins, S., & Graham, M. C. (2019). A cross-disciplinary and multi-modal experimental design for studying near-real-time authentic examination experiences. *Journal of Visualized Experiments*, *151*. https://doi.org/10.3791/60037

Villanueva, I., Valladares, M., & Goodridge, W. (2016). Use of galvanic skin responses, salivary biomarkers, and self-reports to assess undergraduate student performance during a laboratory exam activity. *Journal of Visualized Experiments*, *108*. https://doi.org/10.3791/53255

Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: a meta-analysis. *Psychonomic Bulletin & Review*, *18*(2), 267–277. https://doi.org/10.3758/s13423-010-0042-0

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial Ability for STEM Domains: Aligning Over 50 Years of Cumulative Psychological Knowledge Solidifies Its Importance. *Journal of Educational Psychology*, *101*(4), 817–835. https://doi.org/10.1037/a0016127

Wang, M.-T., & Degol, J. L. (2017). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. In *Educational Psychology Review* (Vol. 29, Issue 1, pp. 119–140). https://doi.org/10.1007/s10648-015-9355-x

Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not Lack of Ability but More Choice. *Psychological Science*, *24*(5), 1–6. https://doi.org/10.1177/0956797612458937

Wang, Y. (2007). The Theoretical Framework of Cognitive Informatics. *Cognitive Informatics and Natural Intelligence*, *1*(1), 1–27.

Wang, Y., & Chiew, V. (2010). On the cognitive process of human problem solving. *Cognitive Systems Research*, *11*(1), 81–92. https://doi.org/10.1016/j.cogsys.2008.08.003

Wang, Y., Member, S., Wang, Y., Patel, S., & Patel, D. (2006). A Layered Reference Model of the Brain (LRMB). *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, *36*(2), 124–133. https://doi.org/10.1109/TSMCC.2006.871126

Ward, A. E., & Thiriet, J. (2010). How well aligned are the views of generic competence development between Electrical and Information Engineering students, their faulty and their employers? *9th International Conference on Information Technology Based Higher Education and Training (ITHET)*, 67–74.

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial Ability: A Neglected Dimension in Talent Searches for Intellectually Precocious Youth. *Journal of Educational Psychology*, *99*(2), 397–420. https://doi.org/10.1037/0022-0663.99.2.397

Welsh, M. C., & Huizinga, M. (2005). Tower of Hanoi disk-transfer task: Influences of strategy knowledge and learning on performance. *Learning and Individual Differences*, *15*(4), 283–298. https://doi.org/10.1016/J.LINDIF.2005.05.002

Welsh, M. C., Satterlee-Cartmell, T., & Stine, M. (1999). Towers of Hanoi and London: Contribution of working memory and inhibition to performance. *Brain and Cognition*, *41*(2), 231–242. https://doi.org/10.1006/brcg.1999.1123

World Economic Forum. (2020). *The Future of Jobs Report*.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence*, *40*(1), 1–14. https://doi.org/10.1016/J.INTELL.2011.11.003

Xu, X., Kim, E. S., & Lewis, J. E. (2016). Sex difference in spatial ability for college students and exploration of measurement invariance. *Learning and Individual Differences*, *45*, 176–184. https://doi.org/10.1016/j.lindif.2015.11.015

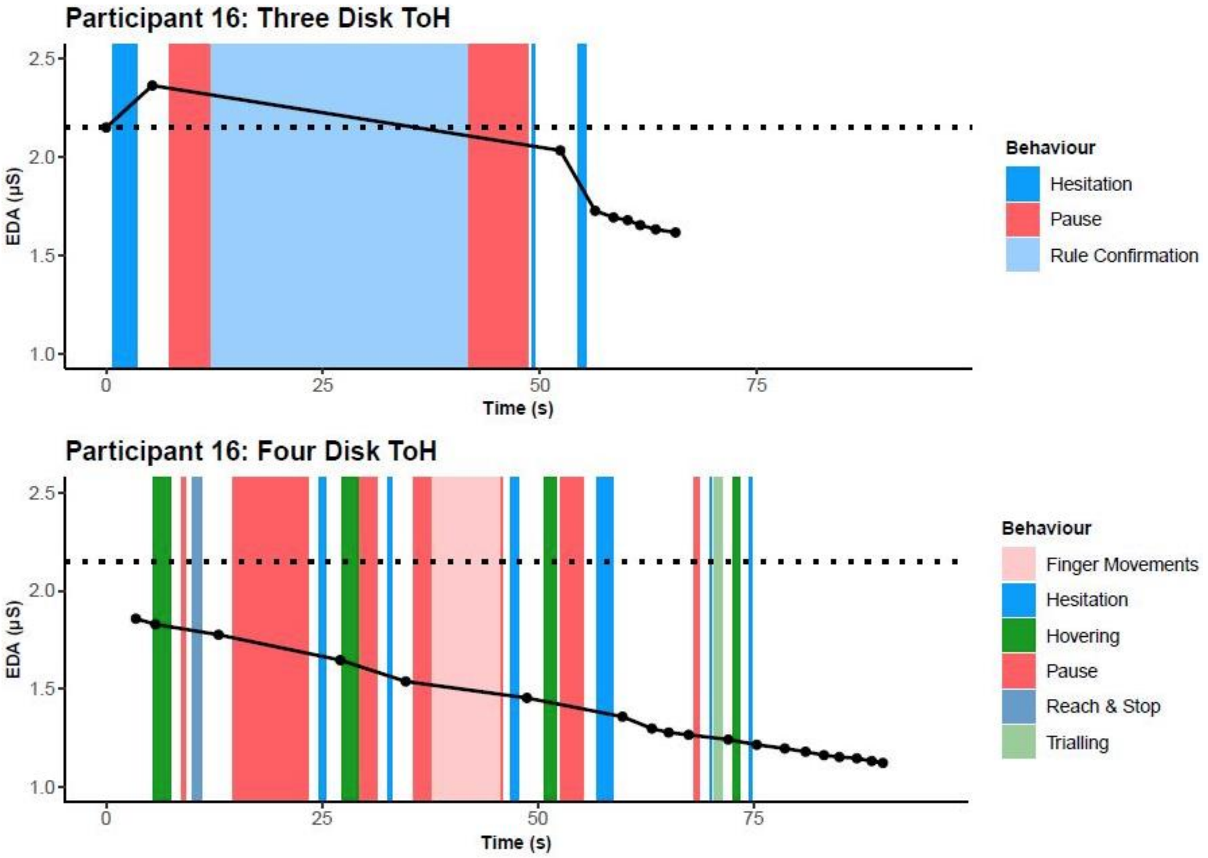
Yilmaz, H. B. (2009). On the development and measurement of spatial ability. *International Electronic Journal of Elementary Education*, *1*(2), 83–96.

# Appendices

### *Appendix I - Joint performance, EDA, and behavioural graphs of participants in descending order of spatial ability*

To evaluate the insight that could be gained from combining the EDA, inductively coded behaviours, and problem-solving performance data, joint display graphs were produced for all participants in the preliminary study. The horizontal dashed line on each graph represents the individuals baseline EDA. The first point on the graph represents when participants were instructed to begin the problem with each subsequent point indicating the next move, time it occurred, and EDA at that moment for the participant. The thickness of the coloured vertical lines is determined by the time spent exhibiting an observed behaviour with the colours representing the theme that each behaviour is coded into. The behaviour theme colour codes are as follows:

- Blue: “indecision”
- Green: “progress”
- Red: “stop”


Participant 16: Three Disk ToH

Participant 16: Four Disk ToH

*Figure 34. Participant 16 joint display of EDA data, performance (moves and time), and behaviours.*

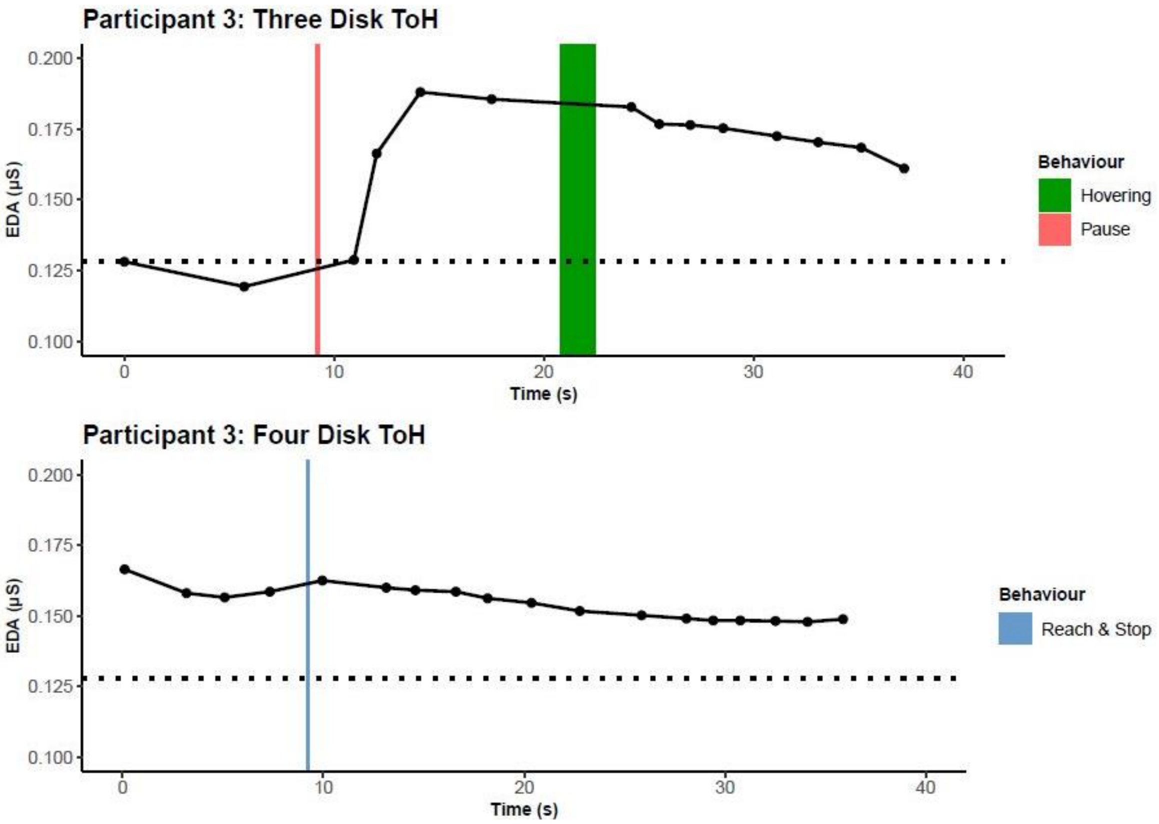


*Figure 35. Participant 3 joint display of EDA data, performance (moves and time), and behaviours.*

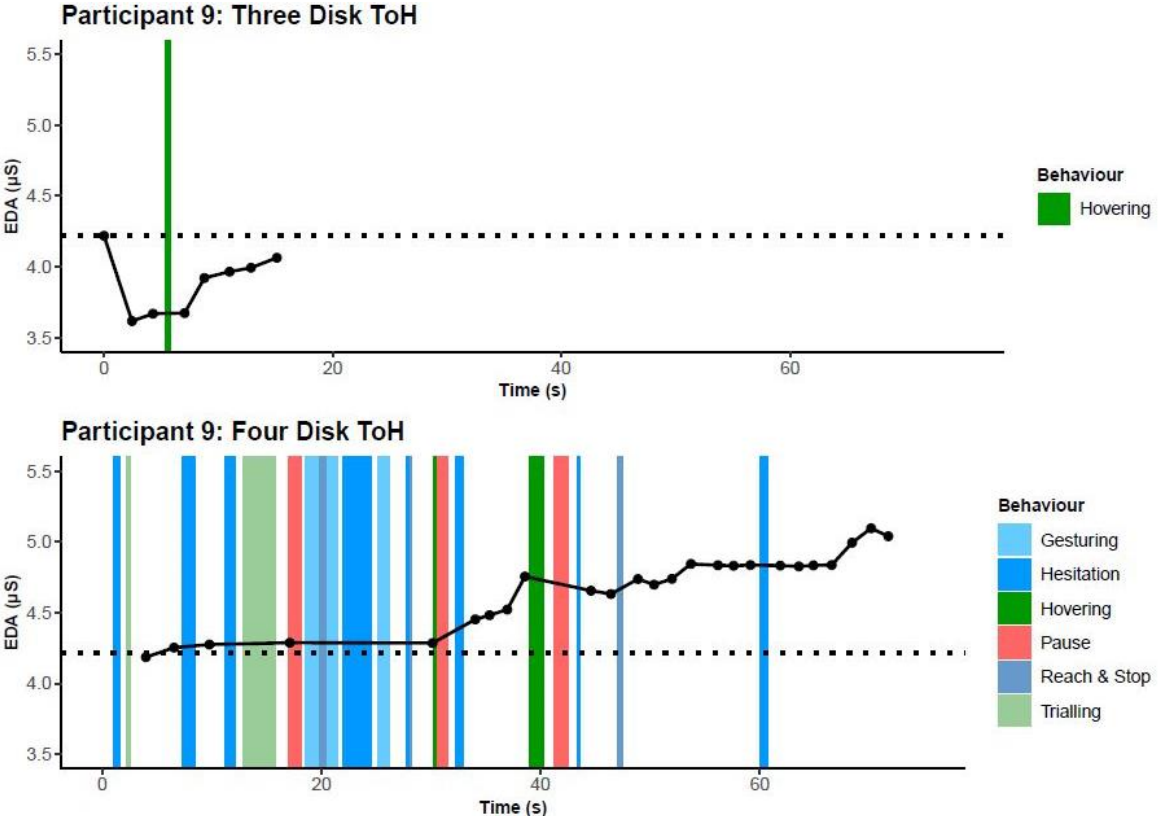
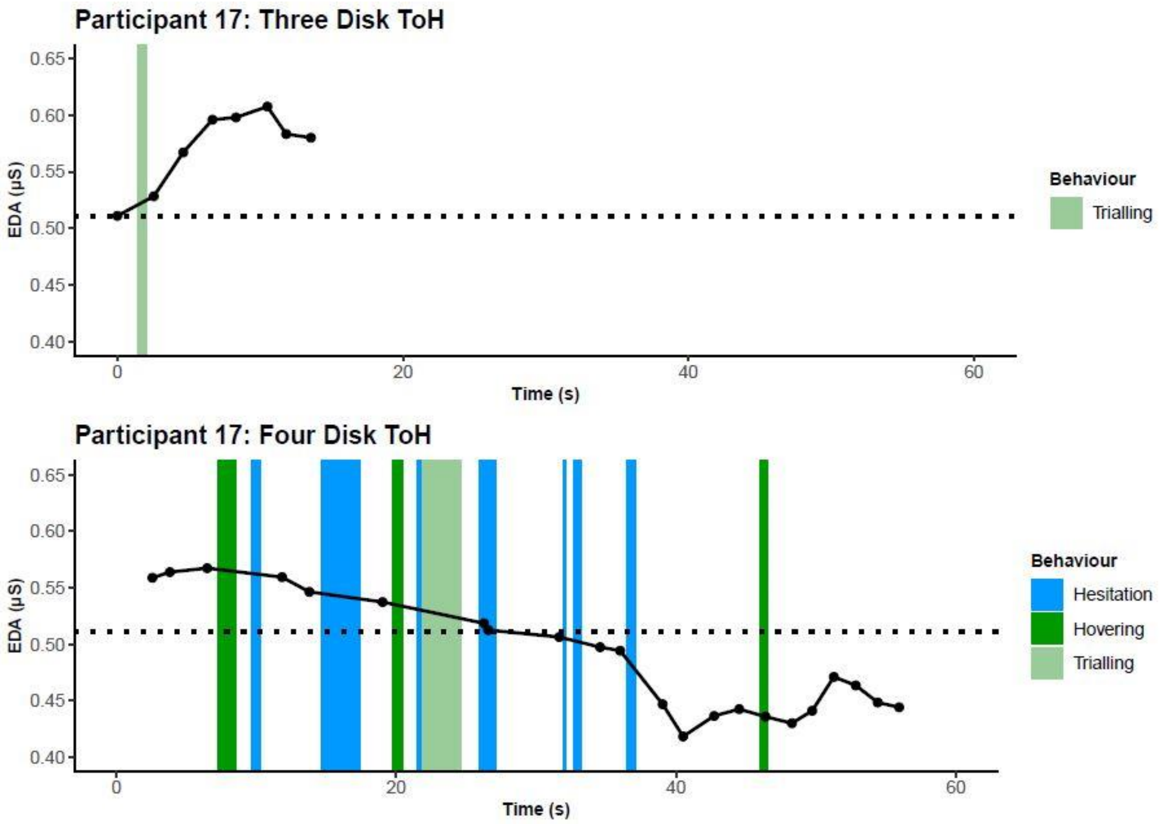


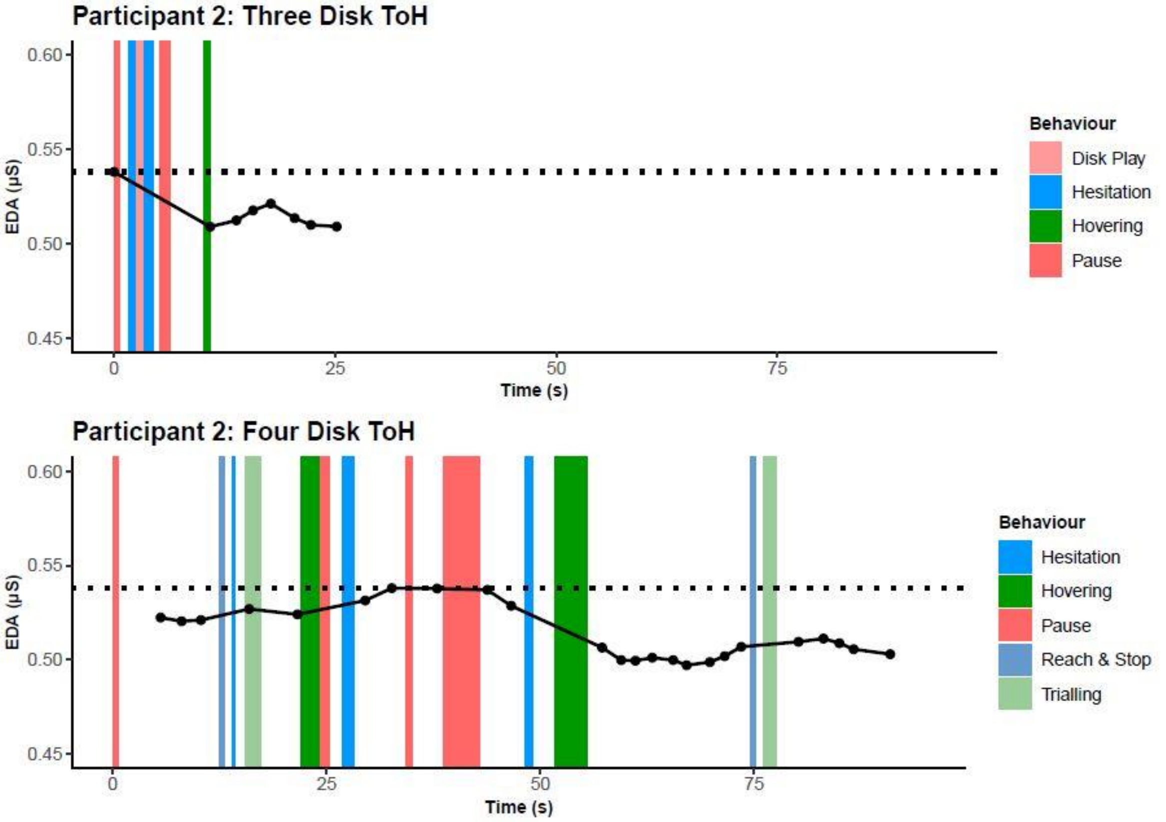
*Figure 36. Participant 9 joint display of EDA data, performance (moves and time), and behaviours.*



Participant 17: Three Disk ToH

Participant 17: Four Disk ToH

*Figure 37. Participant 17 joint display of EDA data, performance (moves and time), and behaviours.*



Participant 2: Three Disk ToH

Participant 2: Four Disk ToH

*Figure 38. Participant 2 joint display of EDA data, performance (moves and time), and behaviours.*

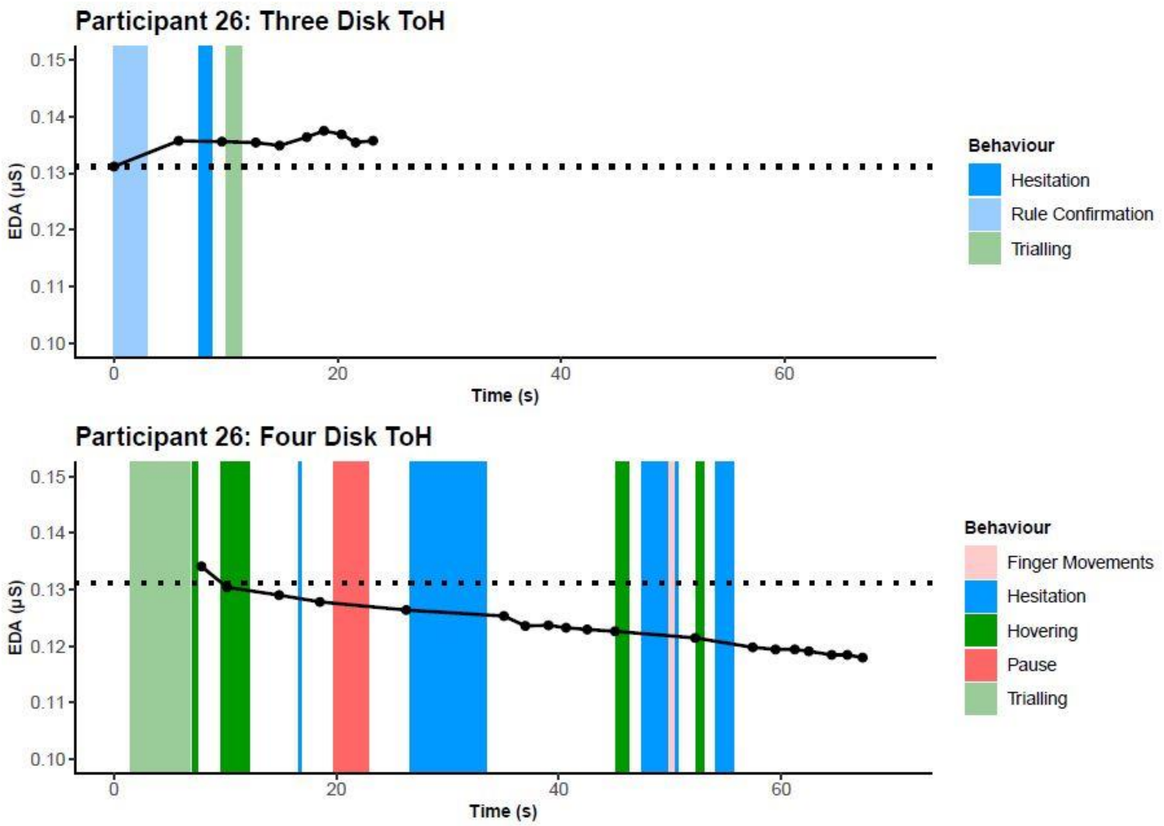


*Figure 39. Participant 26 joint display of EDA data, performance (moves and time), and behaviours.*

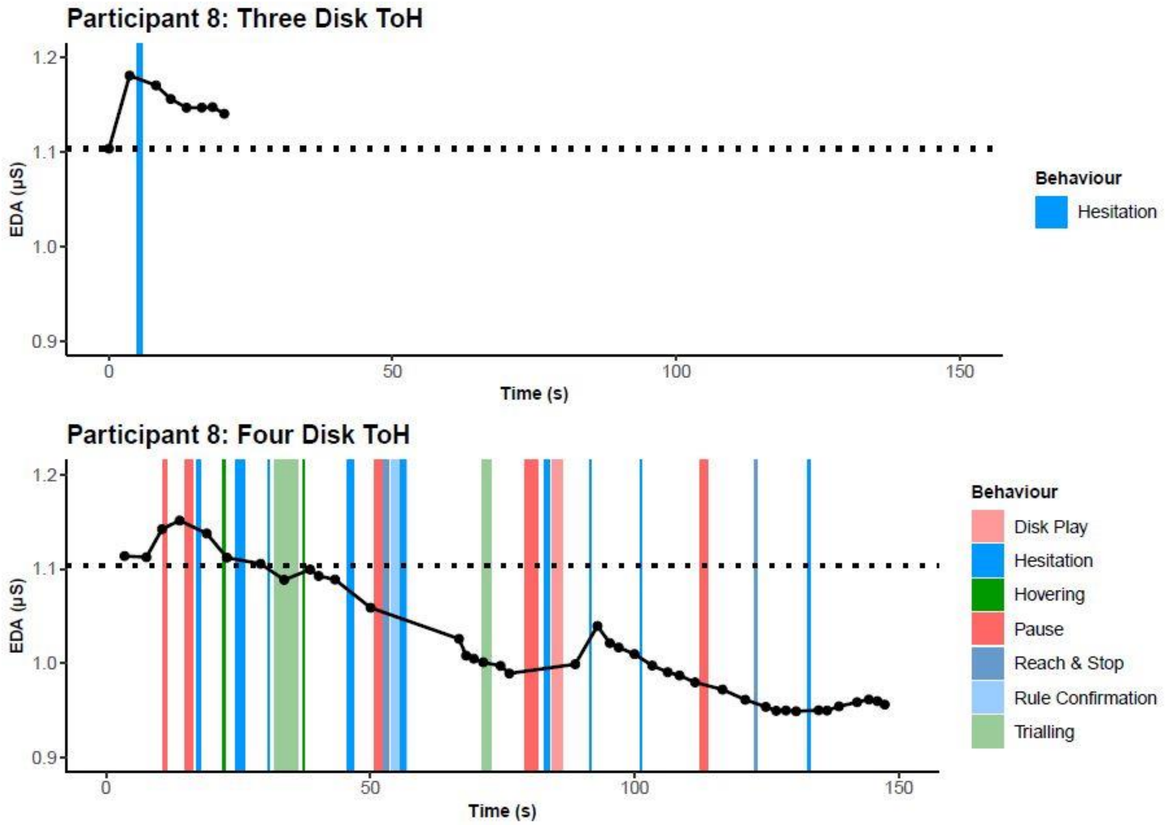


*Figure 40. Participant 8 joint display of EDA data, performance (moves and time), and behaviours.*

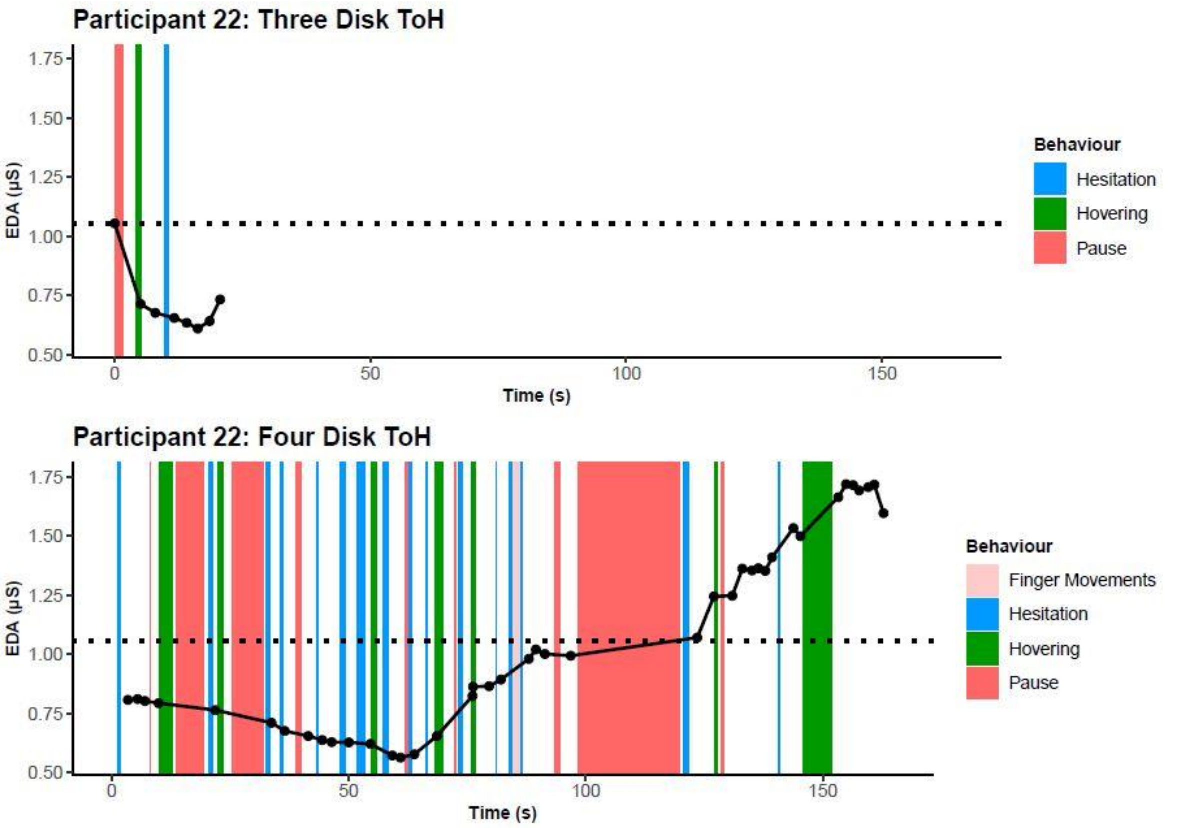


*Figure 41. Participant 22 joint display of EDA data, performance (moves and time), and behaviours.*

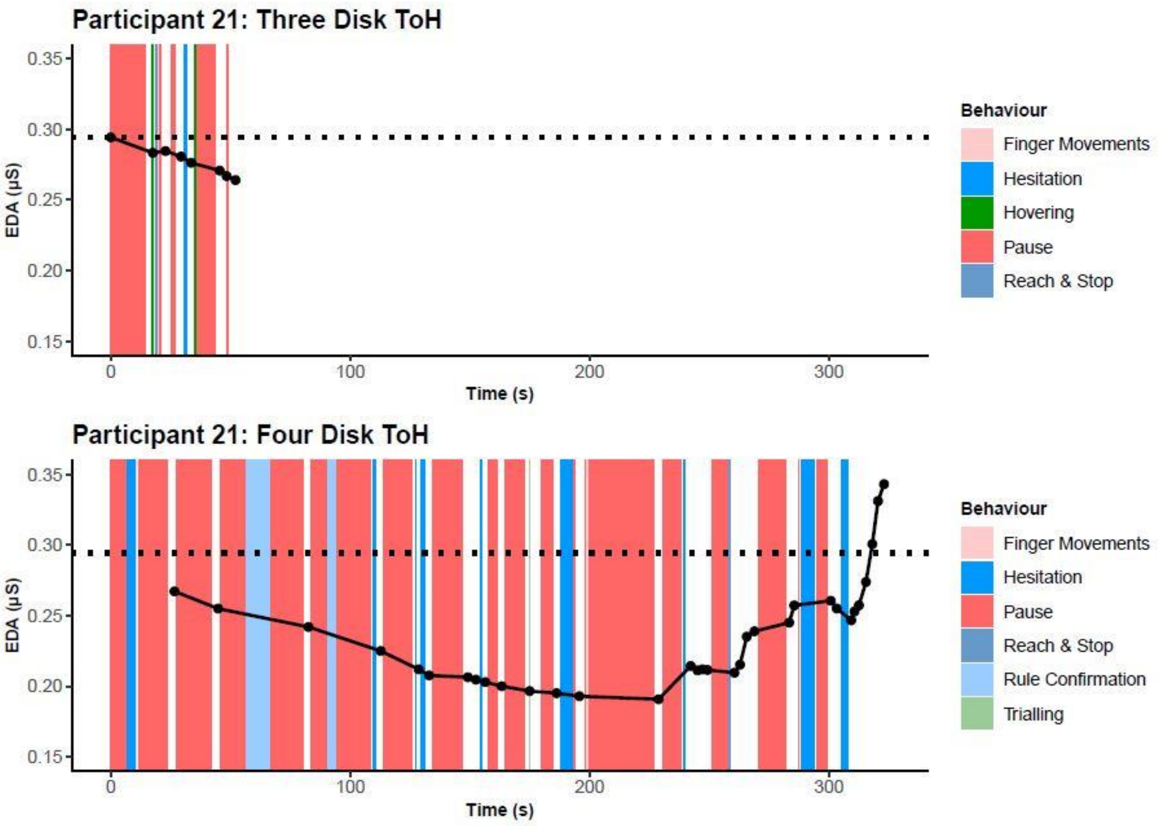


*Figure 42. Participant 21 joint display of EDA data, performance (moves and time), and behaviours.*

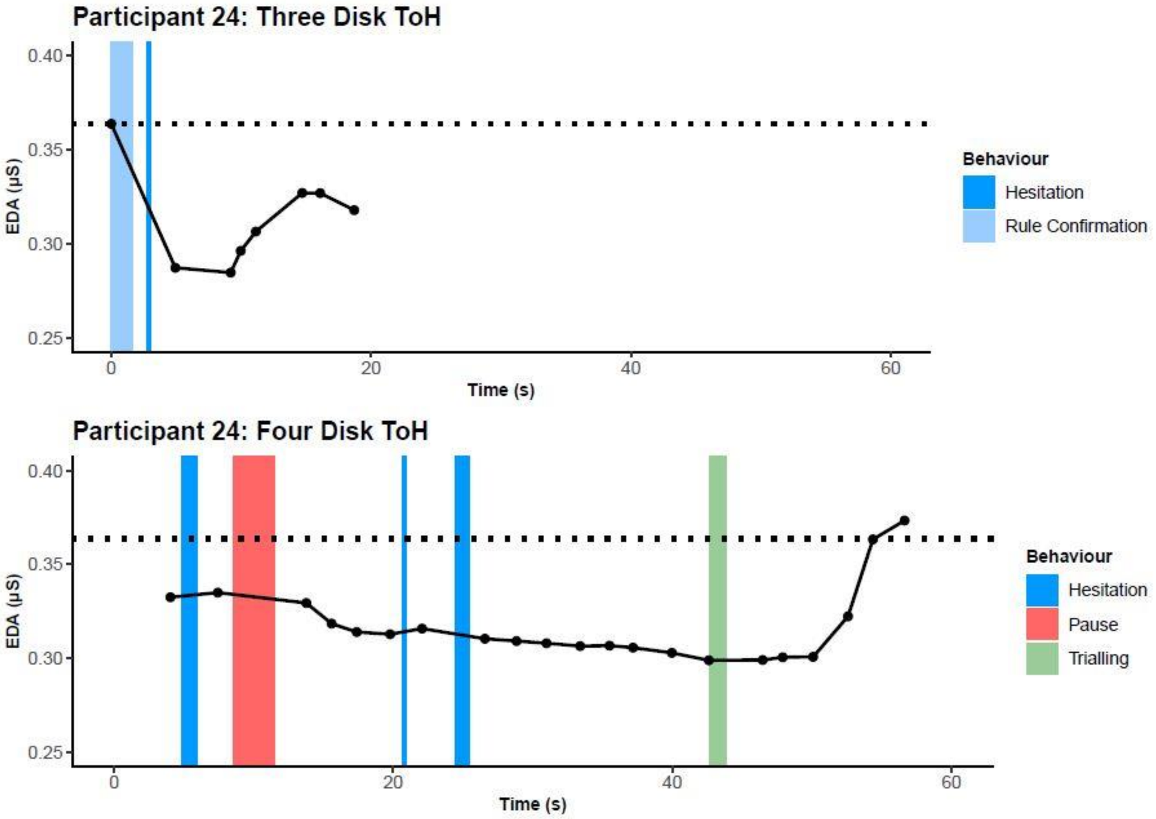


*Figure 43. Participant 24 joint display of EDA data, performance (moves and time), and behaviours.*

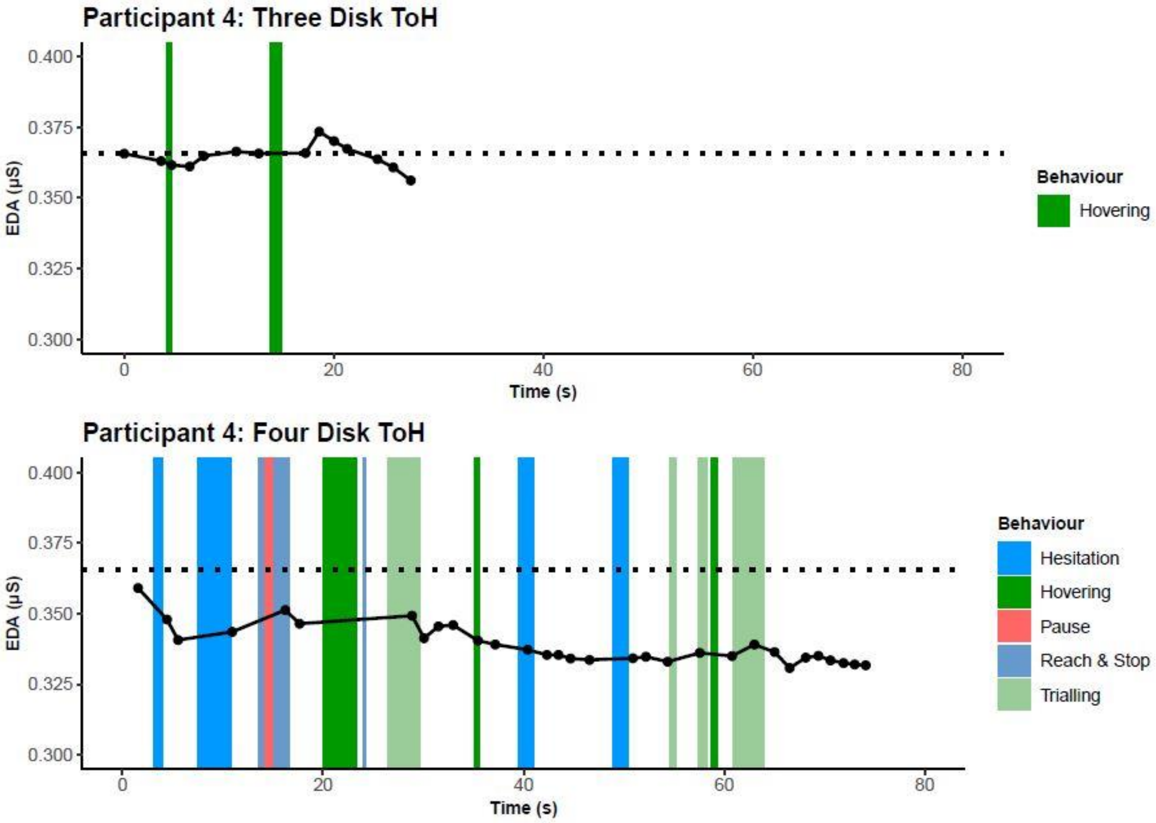
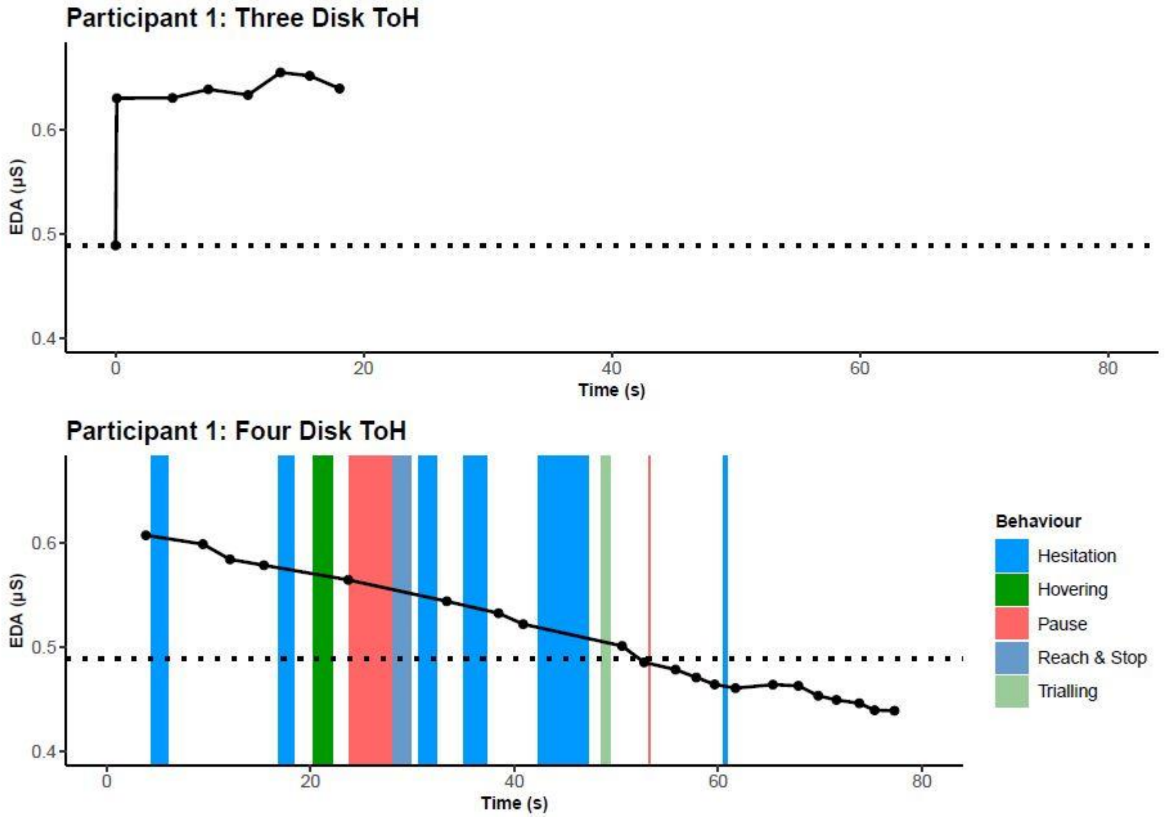


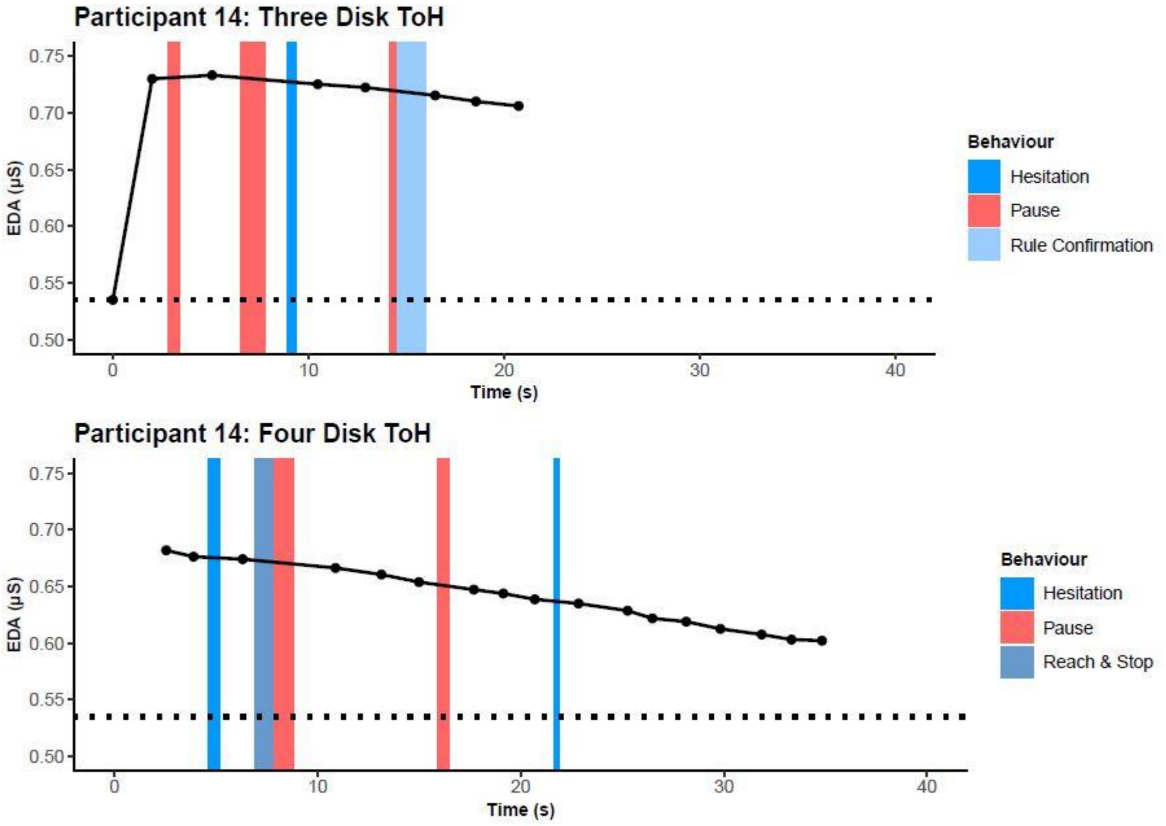
*Figure 44. Participant 4 joint display of EDA data, performance (moves and time), and behaviours.*



Participant 1: Three Disk ToH

Participant 1: Four Disk ToH

*Figure 45. Participant 1 joint display of EDA data, performance (moves and time), and behaviours.*



Participant 14: Three Disk ToH

Participant 14: Four Disk ToH

*Figure 46. Participant 14 joint display of EDA data, performance (moves and time), and behaviours.*

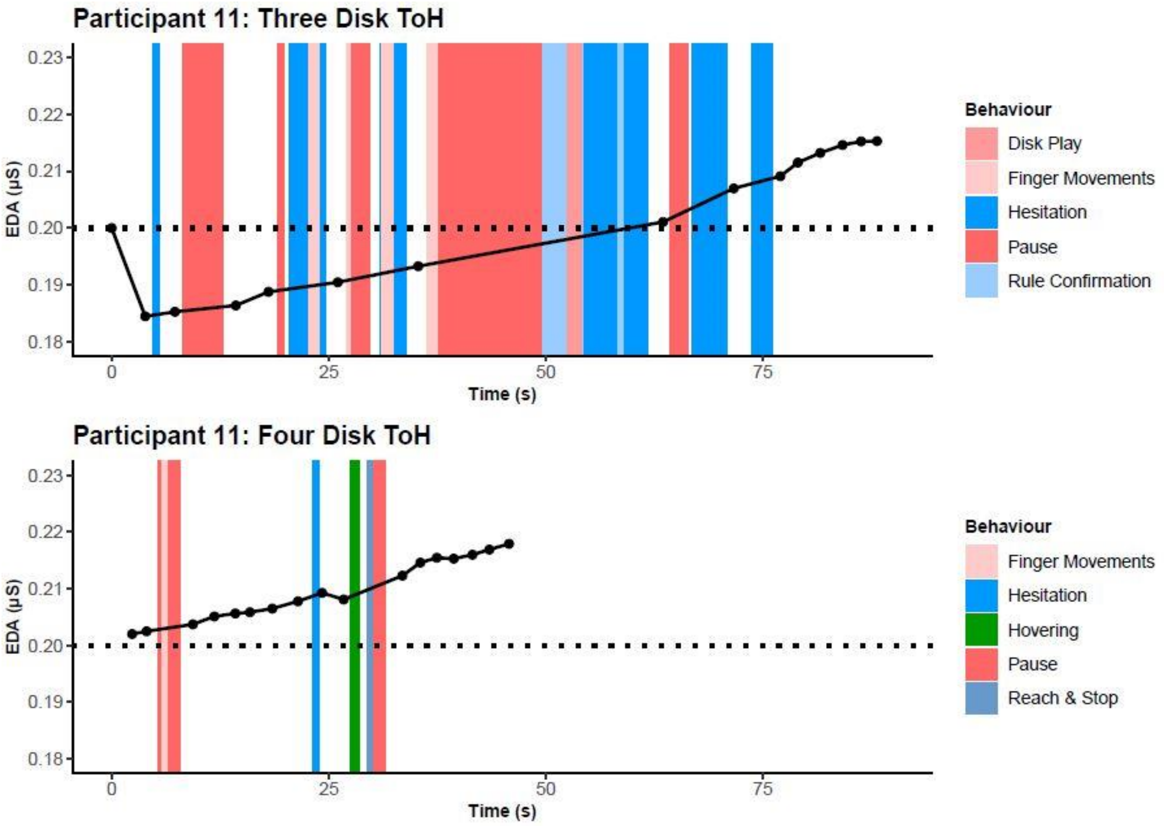


*Figure 47. Participant 11 joint display of EDA data, performance (moves and time), and behaviours.*

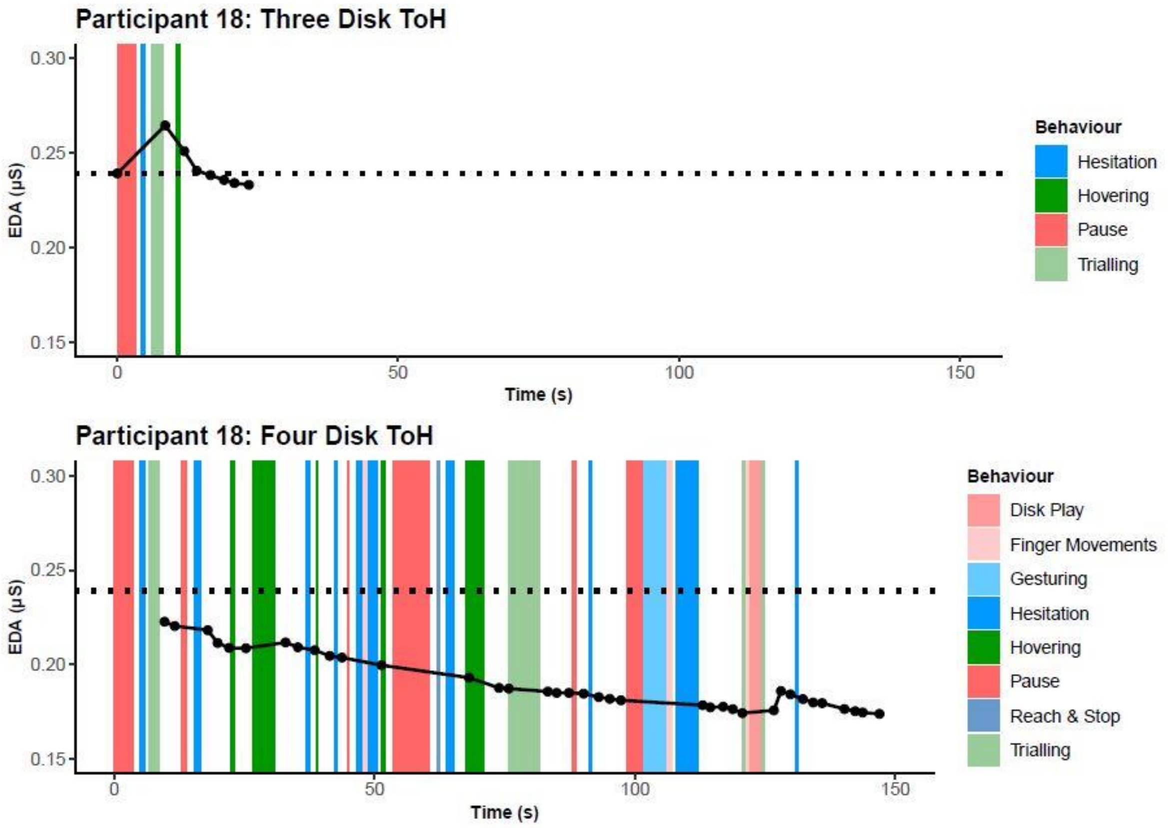


*Figure 48. Participant 18 joint display of EDA data, performance (moves and time), and behaviours.*

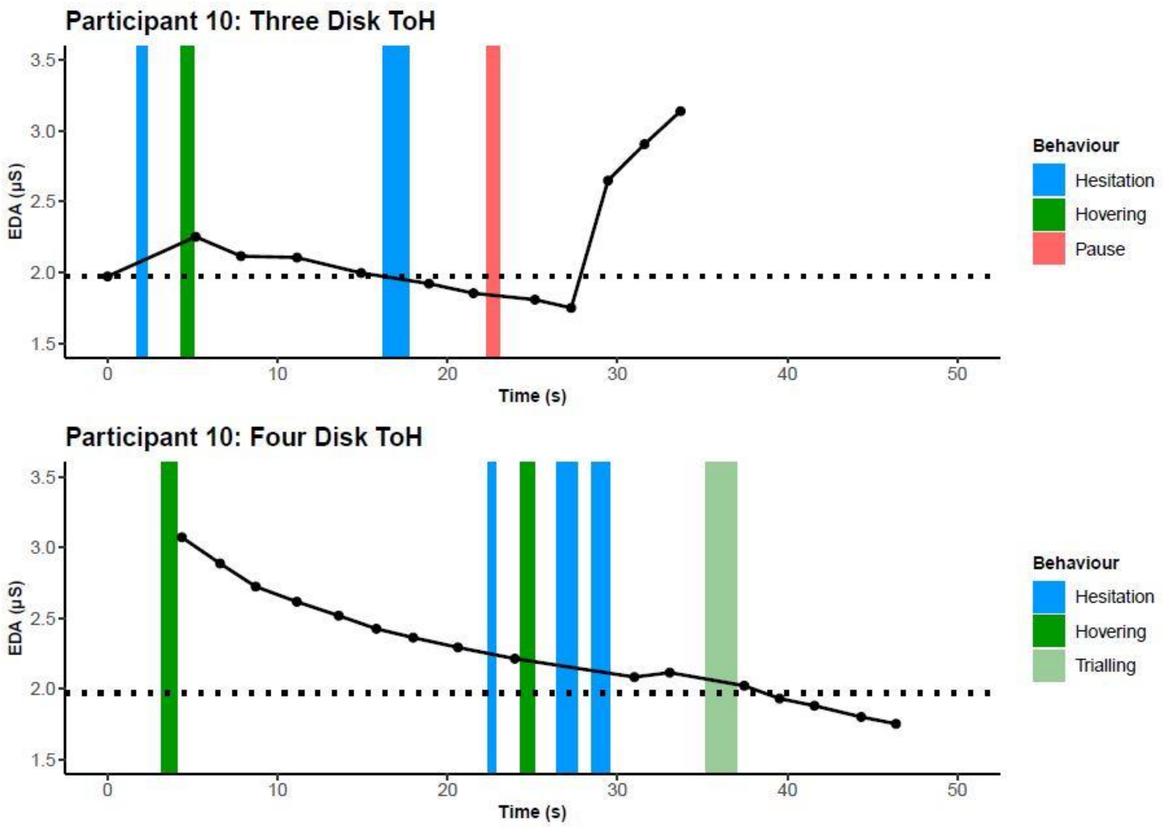


*Figure 49. Participant 10 joint display of EDA data, performance (moves and time), and behaviours.*

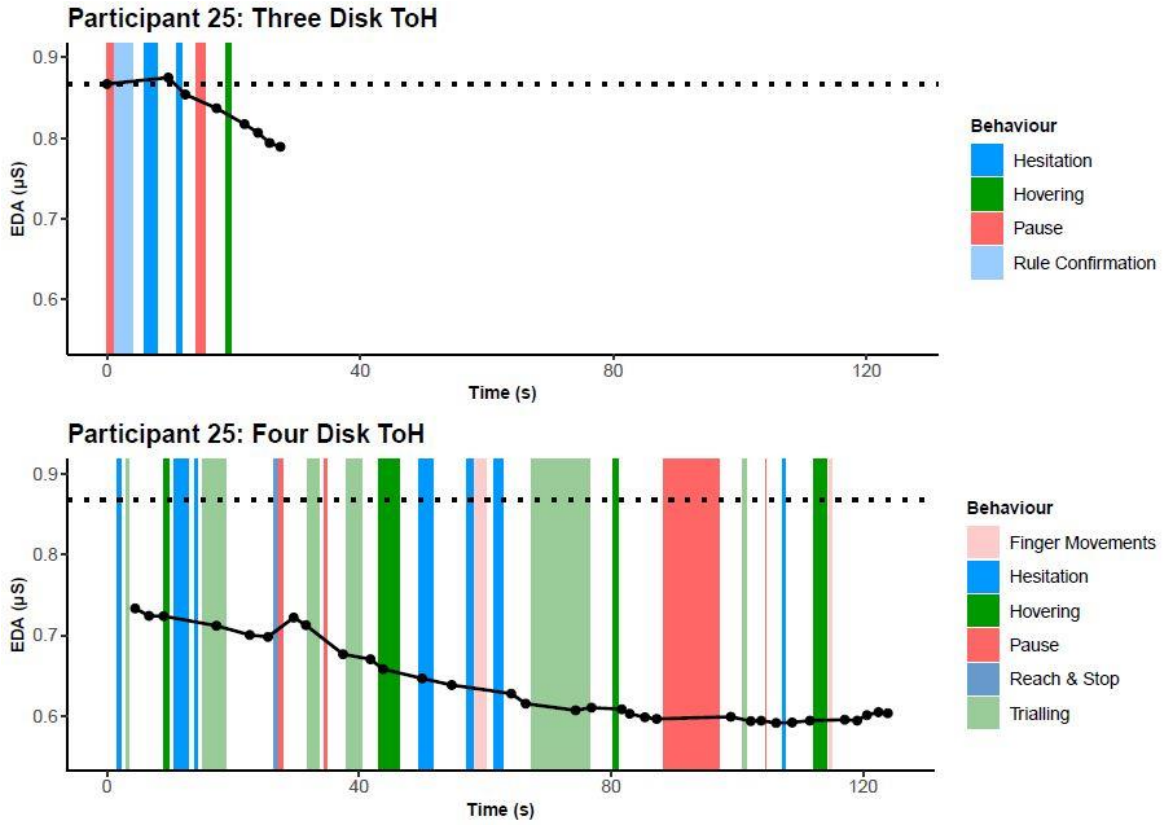
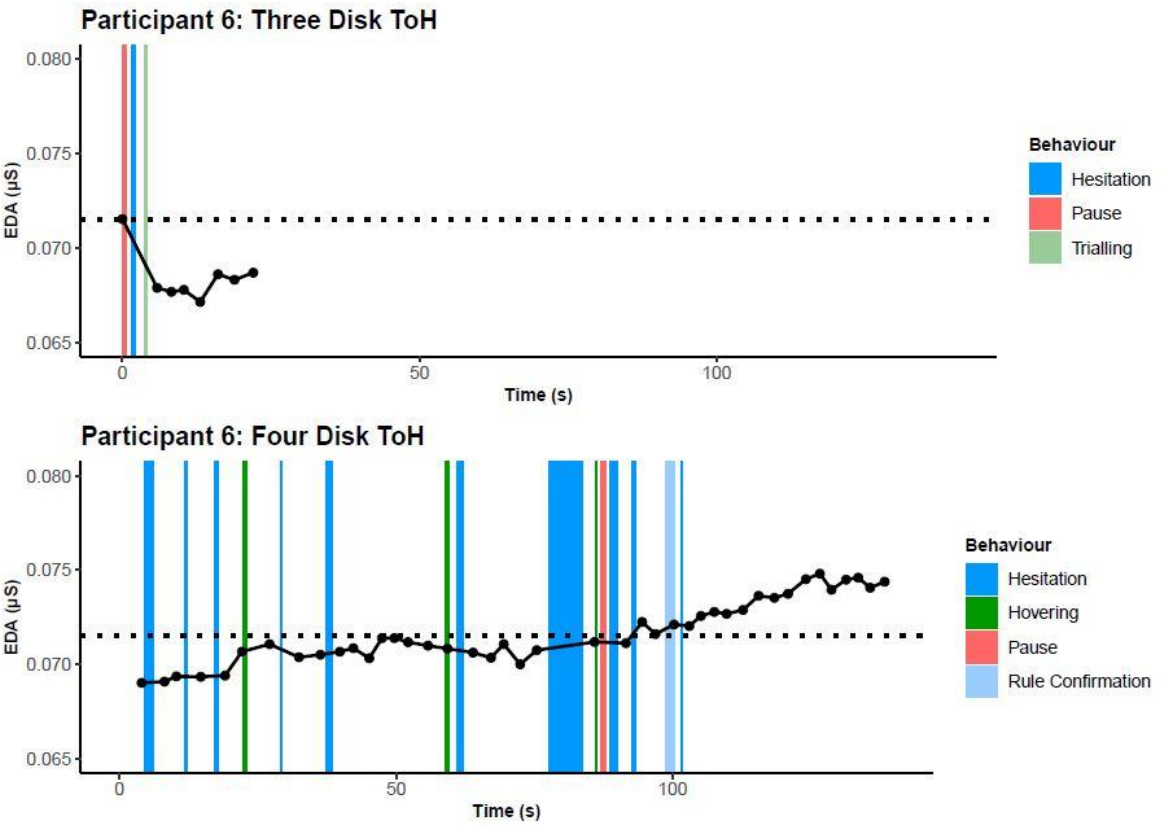
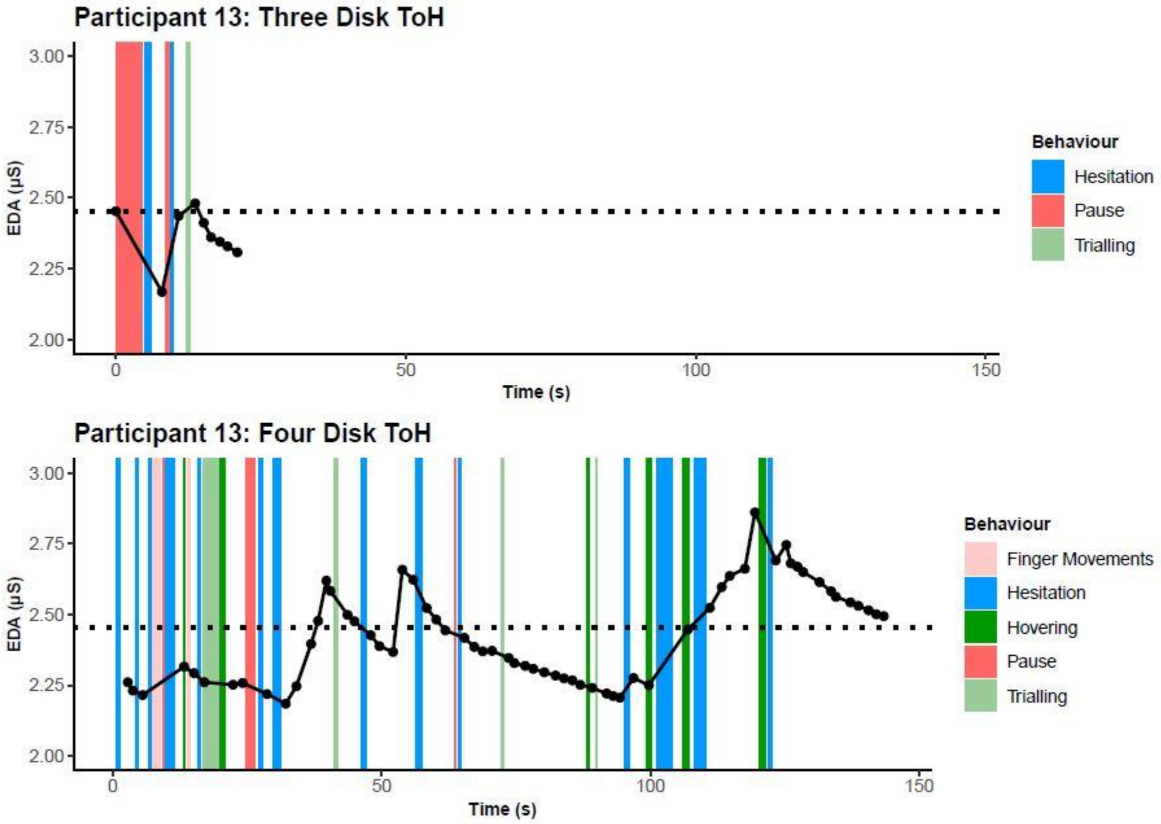


*Figure 50. Participant 25 joint display of EDA data, performance (moves and time), and behaviours.*


Participant 6: Three Disk ToH

Participant 6: Four Disk ToH

*Figure 51. Participant 6 joint display of EDA data, performance (moves and time), and behaviours.*


Participant 13: Three Disk ToH

Participant 13: Four Disk ToH

*Figure 52. Participant 13 joint display of EDA data, performance (moves and time), and behaviours.*

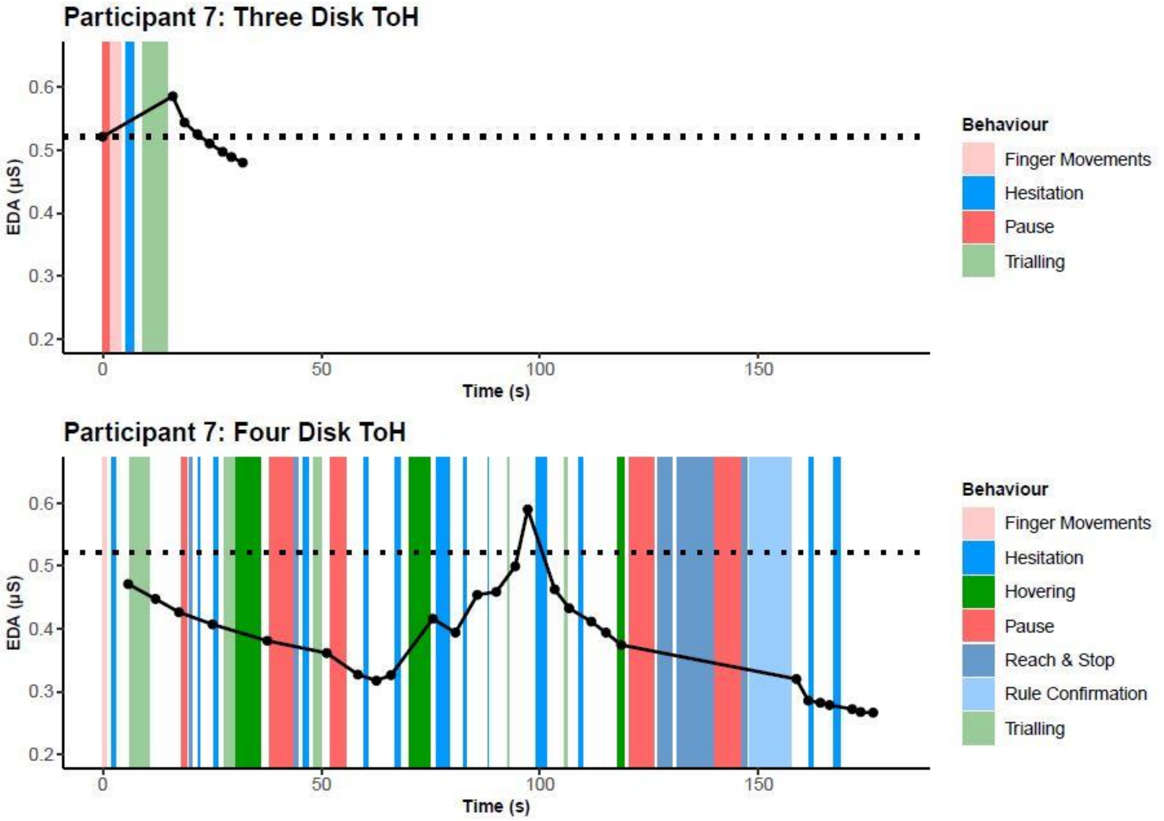


*Figure 53. Participant 7 joint display of EDA data, performance (moves and time), and behaviours.*

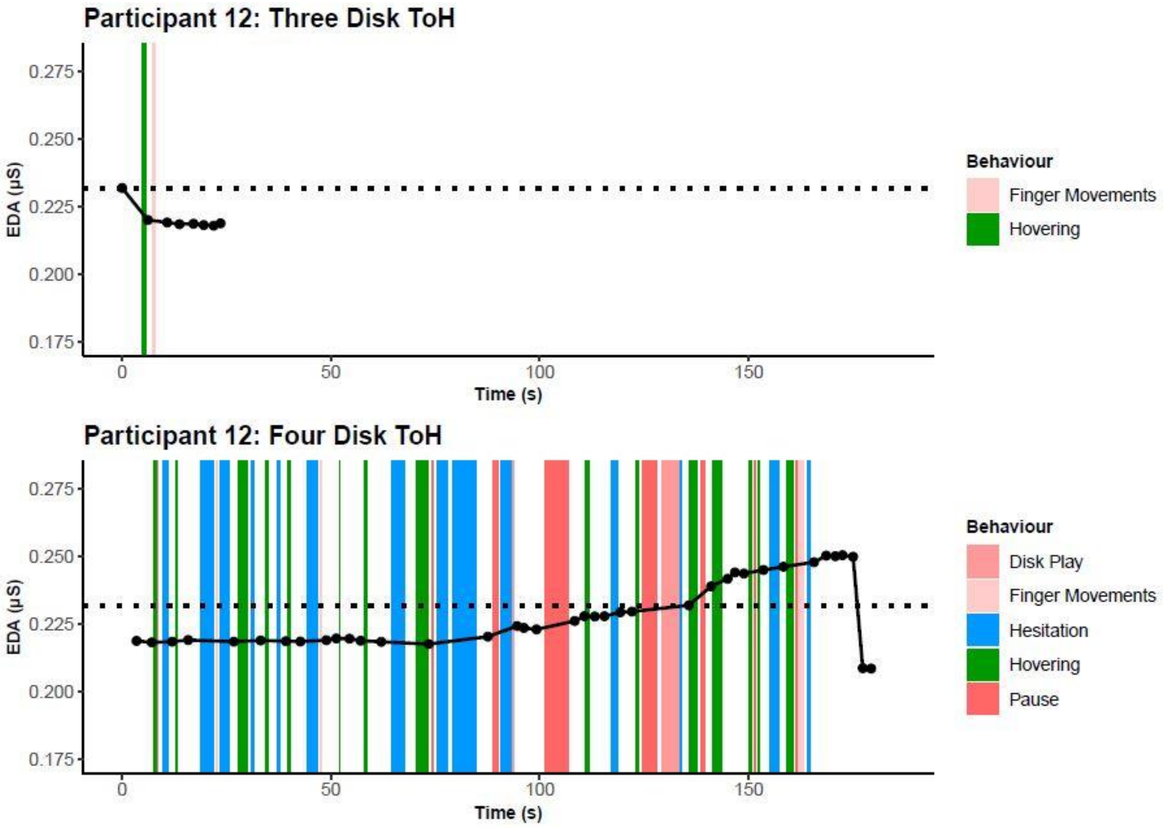


*Figure 54. Participant 12 joint display of EDA data, performance (moves and time), and behaviours.*

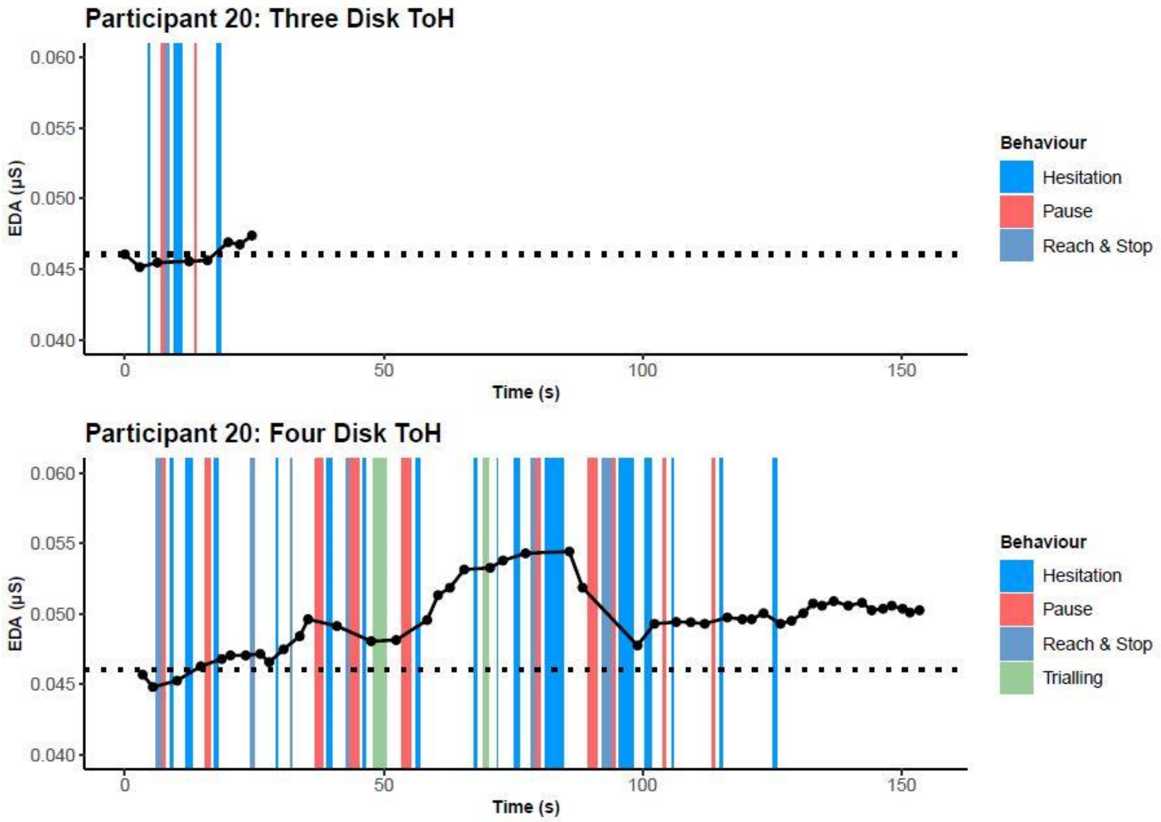


*Figure 55. Participant 20 joint display of EDA data, performance (moves and time), and behaviours.*