Exploring the Root of Sensorimotor Expertise in Dental Education

By

Ahmed Mohammed Bakur Balkhoyor

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Faculty of Medicine and Health

School of Dentistry

January 2020

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The paper listed below has been submitted to Journal of Surgical Education and is currently under review. It forms the basis of Chapter 5 of this thesis. The candidate is the primary author and the named co-authors provided support through review and proofreading prior to publication. The paper has also been disseminated through a conference presentation by the candidate, detailed below:

Balkhoyor A, Mir R, Mirghani I, Pike T, Biyani C, Lodge P, Mon-Williams MA, Manogue, M Mushtaq F. **Exploring the presence of core skills for surgical practise through simulation**. *In the International Association for Dental Research (IADR) Conference, 2018, London, England*.

Dedication

This thesis is dedicated to my beloved wife, Doaa, who has been a constant source of support and encouragement during the preparation of this work. I am truly thankful for having you in my life.

This work is also dedicated to my parents, Abeer and Mohammad, who have always loved me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve. This thesis is also dedicated to my aunts, Afaf and Omalsaad for their unconditional love and generous support every step of the way.

And finally, dedicated to my precious children: Sakha’a, Mohammed, and Jood for all the evenings and weekends that I missed playing with them!

Acknowledgements

First, and foremost, all praise is to Allah, the Most Gracious, the Most Merciful, for his guidance, mercy and countless blessings.

There are a number of people without whom this thesis might not have been written, and to whom I am greatly indebted. First, I would like to thank the Ministry of Education in Kingdom of Saudi Arabia and Um Alqura University Faculty of Dentistry for given me the chance of study in University of Leeds. I gratefully acknowledge the financial support and funding of my MClinDent and PhD scholarship by the Government through the Saudi Arabian Cultural Bureau (SACB) in London. The professional support provided by the SACB team throughout my studies in the United Kingdom is sincerely appreciated.

I would like to thank my thesis advisors, Professor Michael Manogue, Professor Mark Mon-Williams and Dr. Faisal Mushtaq for their contributions to this work. Over the years, each has given me superb scientific guidance, many insightful suggestions and demonstrated a sincere interest in my work. I am fortunate to have such a group of intelligent researchers. They have always been willing to help and gave me the best of advice.

I would like to thank Dr Faisal Mushtaq in particular, for his constructive comments and suggestions on my work; his kind support, constant encouragement, and continuous guidance throughout the write up of this project. I started this work knowing nothing and without his constant constructive feedback this work would have been remained nothing but a dream.

The atmosphere at the University of Leeds, especially within the Perception Action Cognition research group (PAC Lab) and the Immersive

Cognition research group (ICON lab) community, is incredibly collaborative. I would like to extend my appreciation to the people who have fostered this wonderful learning environment.

I would like to thank my beloved family for the continuous support throughout my studies especially my sibling's Rafef, Hanouf, Atouf, Elaf and Anas. I would also like to thank my brothers in law Shade Saber, Makki Maymani, Rayan Bawakid and Mohammed Habhab for their endless support.

Finally, to all of my lovely friends in Jeddah, London, and Leeds I would like to thank you for your support and encouragement. The individuals that I consider friends are too numerous to name but there are a few that cannot go unmentioned, I would specifically like to recognize Hisham Komo, Waleed Taju, Hassan Abed, Abdulaziz Bakhsh, Fatima Nazer, Elham Alsahafi, Mohammed Holdar and Nour for their friendship and support.

Thank You

Ahmed Mohammed Bakur Balkhoyor

Leeds, 01/01/2020

Abstract

Humans are capable of remarkable feats of sensorimotor control, but it is a matter of common observation that some individuals are able to far exceed the capabilities of the general population in specific tasks. These individuals are often labelled experts in their domain, and it is well established that achieving such mastery requires many years of training and in some cases, it can be a life-long pursuit. In dentistry, we are faced with highly specific challenges with time constraints and importantly, all trainees who wish to practice must reach a level of expertise that allows them to operate safely with patients. To understand the development of expertise in dental training, and thus inform training protocols to support learning, this thesis examines the putative mechanisms underlying highly skilled performance. Using state of the art haptic virtual reality (VR) technology, this thesis first examines fundamental differences between experts and novice dentists across a variety of simulated dental tasks. We find that learning over a considerable period of time appear to be quantitively similar in expert and novice data set. We also found that expert performers can use their well-learned sensorimotor skill in a flexible manner to solve new tasks, have superior motor economy and shorter planning times. Moreover, the recruitment of cognitive control scales with the degree of behavioural adjustment following error commission. Having qualitatively and quantitively captured differences as a function of expertise, we ask whether haptic VR technology can be used to accelerate the learning process. Across two experiments, we show that actively manipulating user error during training can have a positive impact on learning and that task difficulty levels need to be tailored to an individual's ability for valid and reliable assessments of trainees on the way to expertise. Taken together, this work presents a comprehensive

examination of dental expertise and highlights the utility of haptic virtual reality technology in supporting the transition from novice to expert.

Table of Contents

List of Figures

Chapter 1 Introduction

1.1 Overview

Sensorimotor skills are an essential part of dental practice. The acquisition of sensorimotor skills starts early on in the preclinical years and requires continued commitment by the practitioner to develop the fine motor skills that allows them to precisely control dental instruments and navigate around delicate oral tissues in a small, contained oral cavity.

My own interest in dental educational research started from my personal experiences as a dental student at an undergraduate level in King Abdul-Aziz University and subsequently at a specialist training level at Queen Mary University of London. Reflecting on the undergraduate and postgraduate clinical dental education process has made me realise that learning dentistry has been one of the most challenging tasks I have undertaken as well as most rewarding. It is clear to me that mastering dentistry takes years of practise and hard work, both on the part of the student as well as the tutors. It is therefore imperative that these efforts be carried out in as efficient a manner as possible. Thus, for good progress to be made, it is perhaps more important that a student practices well than it is that he or she practises often. To help identify exactly what it means to practice well, I set out to identify the characteristic features that set apart expert dentists from trainees. Specifically, the experimental chapters in this thesis set out to address the two fundamental research questions:

1. What are the behavioural and cognitive features that are associated with the development of expertise?

2. How can we use Haptic VR dental simulators to support the development of expertise in undergraduate dental students?

1.2 Outline

The thesis is structured into nine chapters with six empirical chapters presented in the style of manuscripts. In Chapter 2, we explore the relevant literature on expertise, focussing on the general principles of sensorimotor skills and how technology can be used to enhance learning. This is complemented by a focused review of the literature as it relates to the empirical investigations in the following chapters.

In Chapter 3, we perform a longitudinal examination of the performance of an expert and novice in dentistry over an extended period of simulated practice and examine their learning curves.

In Chapter 4, we investigate differences in planning and efficiency between novice and expert performers in a sensorimotor task.

Chapter 5 investigates the specificity and transfer of surgical skills developed by expert dentists in comparison to laparoscopic surgeons on simulated surgical and dental tasks.

In Chapter 6, we explore whether a pattern of oscillatory brain activity known as frontal theta, a potential biomarker for cognitive control processes, could be used to differentiate between novice and experienced dental surgeons.

Chapter 7 and Chapter 8 are concerned with the role of error and task difficulty in classification and accelerating learning. In Chapter 7, we manipulate task difficulty by increasing the minimum acceptable performance threshold.

In Chapter 8, we directly manipulate a drilling task by changing the haptic profile of target objects to examine how haptic simulators can be implemented to help sensorimotor learning.

Finally, Chapter 9 provides a general discussion of the findings, discusses limitations, presents opportunities for future work. This thesis concludes with a consideration of how this body of work has contributed to furthering our understanding of the highlighted issues.

Chapter 2 Literature Review

2.1 Overview

The acquisition of sensorimotor skills is central to safe and effective dentistry (Rodger, Tang and White, 2016). As such, trainee dentists undergo intensive theoretical and practical training, with fine sensorimotor skills most often learned through simulation-based training using a phantom head simulator with plastic teeth (Perry *et al.*, 2015). In addition to this emphasis on sensorimotor skills, the process of dental education is unique in comparison to other healthcare professions for two key reasons.

First, it is only in dentistry that irreversible procedures are routinely undertaken on members of the public by students in the first half of their training (Ross, 2004; Leinster, 2009). Dental students start providing supervised treatment to real patients relatively early in their career (3rd or 4th year in most dental schools) compared to other healthcare professionals. This demands clinically acceptable level of highly specific sensorimotor skills, such as hand-eye-finger coordination, precise instrument handling and other skills to perform dental procedures safely and effectively from a very early age.

Second, dental education requires bespoke clinical training environments and does not borrow from other healthcare services (*Dental Education at the Crossroads: Challenges and Change*, 1995) and as such, has developed a unique path in supporting students in acquiring the level of expertise required safe and efficient practice.

Within recent years, the education of dental students has undergone several modifications. Some of these changes fundamentally transformed the approach towards delivery of information and dissemination of skills (e.g. with VR haptic simulators, dental students can have instant, objective, and visual feedback that permits enhanced self-assessment) (Buchanan, 2004; Jasinevicius *et al.*, 2004). Irrespective of any changes in the dental education, however large or small, it remains the case that a graduate dentist should be able to make decisions based on qualitative values, to be proficient in diagnosis and to be able to devise and execute treatment competently. Indeed, the Quality Assurance Agency for Higher Education in Dentistry state:

> *"On graduation dental students should have developed a holistic view of patient care, accept their professional responsibilities, and acknowledge their limitations. They should have demonstrated an appropriate level of competence to deal with complex issues both systematically and creatively, make sound judgements based on available data, and have acquired a commitment to continuing professional development"* (AAHE, 2002).

To achieve these goals, dental education has, more than any other health specialty, been reliant on simulators and technology throughout its history (Levine, 2013). While simulation is needed to facilitate the transition into the dental clinic and to enhance the students' experience through inclusion of a wide range of simulated patient and cases scenarios (Hollis, Darnell and Hottel, 2011), these skills are largely contingent on being able to execute treatment competently and this requires a core set of sensorimotor abilities.

This thesis will examine how these abilities develop and manifest in expertise and, to provide context, this introductory chapter includes an overview of the principles of motor learning as motor learnings an integral component in understanding the processes underlying sensorimotor skill acquisition, retention, and transfer to simulated and real environments. The second section of this chapter includes an overview of research on technology enhanced learning in medicine and dentistry, focussing specifically on simulation. Finally, this chapter closes with an overview of the specific dental simulation technology implemented in the experimental chapters reported later in this thesis.

2.2 Defining Expertise

Before we move on to the empirical work reported in this thesis on expertise, it is worth examining what expertise might mean from various perspectives. The Oxford English Dictionary defines expertise as “expert knowledge or skill in a particular subject” (Oxford English Dictionary, 2017). This is the most common and broadest interpretation and one we encounter in everyday language (e.g. Stephen Hawking was an expert physicist while Mozart showed incredible expertise in music). Whilst this type of definition is one that is easy to understand, providing a formal framework for expertise that can be subjected to experimental investigations is a little trickier.

There are, in fact, many definitions of expertise in the scientific literature. Most often, experts are defined by their qualifications and track record (Wielinga, Bredeweg and Breuker, 1988). However, this approach has some noteworthy weaknesses. In particular, degrees often relate to declarative knowledge that is not necessarily the same as procedural experience. Take for example the theoretical knowledge acquired in medical schools and compare that to the clinical knowledge used in medical practice (Burgman *et al.*, 2011; Gobet, 2015).

Expertise is also defined through experience and the amount of time an individual has spent in a specific domain, with the expert performance being a reflection of a long period of deliberate practice (Ericsson, Krampe and Tesch-Römer, 1993). However, there is a surprisingly weak correlation between amount of the time an individual has spent in practice and level of expertise (Ericsson and Charness, 1994). For example, recent work has shown that deliberate practice accounts for only 29.9% of the variance in expertise in musical skills (Hambrick *et al.*, 2014).

Gobet defined an expert as “someone who obtains results that are vastly superior to those obtained by the majority of the population” (Gobet, 2013, 2015). This definition can be applied to domains where most individuals have a high level of natural ability (e.g. walking- an activity that the majority of healthy adults are capable of performing). It can also be applied to experts themselves: a super-expert is “somebody whose performance is vastly superior to the majority of experts” (Gobet, 2013).

Finally, expertise might also be defined by the neural and cognitive processes underlying performance (Debarnot *et al.*, 2015). Recent insights from neuroscience have shown that many brain structures are recruited during task performance, but only activity in regions related to domain-specific knowledge distinguish experts from novices (Buschkuehl, Jaeggi and Jonides, 2012; Guida *et al.*, 2012). Furthermore, compared to novices, experts exhibit a decrease in the overall volume of brain activation and this is coupled with a relative increased intensity of activation in specific brain regions necessary for the execution of the task (Jäncke, Shah and Peters, 2000; Münte, Altenmüller and Jäncke, 2002; Lotze *et al.*, 2003). In this way, experts are able to prioritise task relevant information, and this may be the product of their superior levels of performance.

2.3 Principles of Sensorimotor Learning

Sensorimotor learning is a continuous process that allows us to adapt to new or changing environments, acquire new movement skills and perceptive abilities, and recover from debilitating conditions (Tresilian, 2012). While the majority of the adult human population are able to effectively interact with the environment around us using sensorimotor control, there are some contexts, such as dentistry, where an exceptionally high degree of sensorimotor skill is required. Practising dentistry is a particularly challenging sensorimotor control task as it requires exceptional fine motor skills to precisely control an instrument to navigate around delicate oral tissue, in a small oral cavity.

To preface this introduction (and the experimental work in subsequent chapters), it is worth making a distinction between some terms. Sensorimotor skill, whether it is dentistry or golf is often defined by the quality of movement produced by the performer (Tresilian, 2012; Krakauer *et al.*, 2019). Sensorimotor performance on the other hand, is an observable attempt to perform a motor task that can be influenced by a number of factors such as stress (Gheorghe, Panouillères and Walsh, 2018), fatigue (Aune, Ingvaldsen and Ettema, 2008), and motivation (Wulf and Lewthwaite, 2016). Finally, sensorimotor learning is a set of complex process (perception, cognition and action) involving changes in an individual’s internal processes that determine the persons capability to perform a motor task (Tresilian, 2012; Krakauer *et al.*, 2019).

The relationship between sensorimotor performance and learning may be explained by considering the concept of implicit learning and the schema of learning stages. Implicit learning is defined as the process by which a learner improves performance by practice until the correct performance of the motor skill becomes automatic (Shmuelof, Krakauer and Mazzoni, 2012). These concepts of sensorimotor performance and learning are central to Fitts & Posner's 3 stage model of motor learning (Fitts and Posner, 1967). According to this, now classic, theory of sensorimotor skill acquisition, learning can be divided into three stages: an early cognitive stage, an intermediate associative or integrative stage and finally, an automatic or autonomous stage of performance.

To contextualise Fitts & Posner's model for the present topic, consider learning to drill a tooth - a skilled motor task that all dental students learn in their preclinical years. The initial cognitive stage is characterised by erratic performance as the trainee is required to learn the mechanics of the task (e.g. learning to grasp the instruments, how to get the finger support, and which part of the tooth to drill and which not). At this stage, performing the task requires significant cognitive effort. With prolonged practice, the trainee progresses into the integrative stage, at which point performance becomes more refined as the learner is able to apply their knowledge (e.g. holding the instrument in the correct position, and the student focusing on the caries, checking which part of the tooth is to be removed before drilling). The autonomous stage is the stage when the task is no longer cognitively demanding, and it can be carried out with low cognitive effort.

2.3.1 Components of Sensorimotor Learning

Since Fitts & Posner’s model was first introduced, significant advances have been made in understanding the mechanisms involved in sensorimotor learning. Researchers who study sensorimotor control have simplified the processes involved into four interacting components (Wolpert and Flanagan, 2010).

The first component in sensorimotor control and learning is the effective and efficient extracting of sensory information, which allows the performer to determine when, where, and how to use their sensory receptors (Wolpert, Diedrichsen and Flanagan, 2011). This gathering of task-relevant sensory information is an essential process in which the learner can decipher which sensory information to process and how to extract the most relevant information in an efficient manner (Wolpert, Diedrichsen and Flanagan, 2011). For example, before drilling a tooth with caries lesion, dental students need to gather sensory information about the size of the decay, softness of the caries dentin, and proximity to the pulp tissue.

Second, the learner must understand key features of the task such as the normal anatomy and biology of the tooth and the oral environment, and how that can impact on their action (Braun *et al.*, 2010; Wolpert, Diedrichsen and Flanagan, 2011). For example, in drilling a tooth with caries, the student must learn the transformation between muscle commands and the motion of the drilling bur, learn how to credit errors to different aspects of our performance and determine how the context (such as different in density or saliva conditions) affects the task.

Third, the learner needs to set up different classes of parameter to optimize sensorimotor performance. This will generate appropriate motor commands to achieve the task goals (for example, optimizing the power with which the foot presses the foot-pedal to control the speed of drilling).

The final component is making effective decisions based on the experience. Most sensorimotor tasks involve a sequence of decision-making processes based on the gathering of sensory information. This includes when to make the next movement and which movement to make. Being able to effectively adapt the pressure and power of drilling according to the pulp anatomy is a clear example of when this component is necessary (Braun *et al.*, 2010; Wolpert, Diedrichsen and Flanagan, 2011).

2.3.2 Learning Signals

Sensorimotor learning may also be classified by the type of information that the motor system uses as a learning signal. This topic is often separated into error-based learning, reinforcement learning and observational learning (Wolpert, Diedrichsen and Flanagan, 2011). These process are central to reducing uncertainty to minimise the errors, and maximising performance efficiency (Todorov and Jordan, 2002) and we discuss them in more detail next.

2.3.2.1 Error-based Learning

Learning from errors is a basic tenet of motor skill acquisition (Miall and Wolpert, 1996; Diedrichsen *et al.*, 2010). Simply put, the differences between the goal and the actual sensory information (a prediction error) drive changes in future action. Specifically, when motor errors are detected by sensory systems, this information is used to guide and update motor commands for subsequent motor actions (Seidler *et al.*, 2013). This information not only tells us that we missed the goal, but also specifies the particular way in which the target was missed. However, to effectively use this information, the nervous system needs to correlate the error with respect to each component of the motor command (Wolpert and Flanagan, 2010).

The relationship between feedback and feedforward control can be observed during the acquisition of a novel task e.g. learning to use indirect vision to drill a tooth. New skills do not have enough of a motor history for an accurate forward input, which results in a large prediction error along with jerky and inaccurate movements. As the student becomes used to performing the action, smaller errors occur and thus, fewer adjustments. Actions become faster, smoother, and more accurate as skill level increases (see Chapter 3 for an empirical demonstration of this phenomenon). Theoretically speaking, error-based learning is a highly effective learning process, but once the average error is zero it does not provide a mechanism to further improve the performance (Wolpert, Diedrichsen and Flanagan, 2011).

2.3.2.2 Reinforcement Learning

A second critical signal for sensorimotor learning relates to information indicating the relative success and failure of an action. Humans are highly capable of tracking the value of sensory input and varying their behaviour on the basis of motor history (McDougle *et al.*, 2016). When a sequence of actions results in an outcome, how do we determine which actions should get credit for the successful outcome? For example, a novice dentist drilling a deep caries lesion may try several angles and instruments before cleaning the tooth. Some of those actions may bring it closer to the pulp or to drill into the healthy part, while others contribute towards task success. How does the dentist learn which actions are optimal and which suboptimal? Reinforcement learning studies show how performers can learn to behave so as to maximise the rewards and minimise the losses (Wolpert, Diedrichsen and Flanagan, 2011).

In real-world actions, the underlying cause of an unsuccessful attempt is sometimes ambiguous: accidentally drilling into the tooth pulp could occur because the dentist made a poor choice about where to drill or failed to properly execute the drilling. Thus, learning to act so as to maximise success and minimise failure requires the ability to predict future outcome success through accurately evaluating current successes and failures (Maia, 2009). Reinforcement-learning systems within the brain are capable of resolving these problems to optimise behaviour (Mushtaq *et al.*, 2016; McDougle *et al.*, 2019) and reinforcement is key to enhancing the consolidation of a motor behaviour (Abe *et al.*, 2011; Huang *et al.*, 2011). However, notably, reinforcement signals do not give information about the direction of required behavioural change (Wolpert, Diedrichsen and Flanagan, 2011). Thus, the sensorimotor system needs to try out different possibilities to gradually improve performance. In addition, because the signal (the success /failure) provides less information than error-based learning, the learning is typically slower (Wolpert, Diedrichsen and Flanagan, 2011).

2.3.2.3 Observational Learning

Observation of others is an important source of information in learning sensorimotor skills (Wolpert, Diedrichsen and Flanagan, 2011). The learner typically requires observing an expert model physically performing a task, after which the learner attempts to mimic the action they have just observed. Studies have shown that watching another person perform an action engages similar, or at least overlapping, sensorimotor representations of the observed action (Buccino, Binkofski and Riggio, 2004; Cattaneo and Rizzolatti, 2009). Numerous studies have provided evidence that people can learn and extract information about what movements to make, and in what sequence, by observing other peoples' actions (Heyes and Foster, 2002). Indeed, some work has shown that people can also learn how to compensate their movement through observation (Mattar and Gribble, 2005).

Despite the absence of physical involvement of the motor system in trial and error learning during observational learning, this process of sensorimotor learning might involve error-based learning as the observer compares the prediction to actual outcomes and can use the error to update the sensorimotor system (Wolpert, Diedrichsen and Flanagan, 2011; Roberts *et al.*, 2014).

Observation alone can't substitute for physically performing the task, but it is particularly beneficial when used as an adjunct to actual practice (Weeks and Anderson, 2000). Observational learning already plays a significant role in healthcare training, through demonstrations of procedures or the opportunity to observe surgical or dental procedures in the operating room (Harris *et al.*, 2018). During observation, the observer tends to produce predictive eye movements similar to the performer, by following the objects before they are interacted with (Flanagan and Johansson, 2003), indicating effective gathering of sensorimotor information. Additionally, understanding the key features of the task has been demonstrated in a learning of hand movement sequences (Blandin, Lhuisset and Proteau, 1999). Indeed, task strategies in sensorimotor tasks can be learned directly from the performer, leading to improvement in decision-making skills (Harris *et al.*, 2018).

2.3.3 Selection and Execution in Sensorimotor Control

An important characteristic of skill learning is the improvements in accuracy and/or speed that comes with training (Bassett *et al.*, 2015). It is generally agreed that this process is the manifestation of a set of a hierarchical processes (Diedrichsen and Kornysheva, 2015). At the neural level, movements are generated through the interaction of different representational levels in the motor cortical neurons, traversing from movement goals (selection level) down to the specification of the actual muscle commands (execution level).

When we attempt a movement, the selection process activates a spatiotemporal pattern that is specific for muscle activity (Churchland *et al.*, 2012). This process is time-consuming because it needs to consider multiple factors and then select the most appropriate set of motor actions, or a ‘motor primitive’ - a spatiotemporal pattern of muscle activity that occurs across a range of complex movements and is encoded in the spinal cord and the primary motor cortex (Hick, 1952; Diedrichsen and Kornysheva, 2015). At the execution level, stable spatiotemporal patterns of muscle activity are produced with the outcome being muscle activity.

Central to the idea that sensorimotor skill is hierarchically represented is motor chunking (Lashley, 1951). Motor chunking is the “segregation of long sequences of movements into subparts, and concatenation of motor responses into groups of responses” (Diedrichsen and Kornysheva, 2015). Movements that are grouped into one chunk are retrieved faster and more accurately than when the selection level triggers them individually, with the advantage that acquired chunks can be used in the context of novel sequences. In this context, sensorimotor learning means effortful selection of single movement elements to their combined fast and accurate production of action. Hierarchical representations also allow generalisation and the flexible generation of novel behaviours and produce movements using less motor planning or preparation time (Wolpert, Diedrichsen and Flanagan, 2011; Diedrichsen and Kornysheva, 2015).

2.3.4 Generalisation of Sensorimotor Skill

Generalisation of a sensorimotor skill, in which the learning of a response in one situation influences the response in another can also be defined as the gain (or loss) in the capability for responding in a new task (the criterion task) as a function of practice or experience on some other task(s) (the transfer tasks), is sometimes used as a way of making presumption about basic behavioural mechanisms (Schmidt and Young, 1986; Adams, 1987).

Generalisation is a double-edged sword: if a small behavioural change is associated with a large alteration of the learning problem, then generalisation from prior learning will interfere with the new task and impair performance (Krakauer *et al.*, 2006). For example, when a naïve student drills a tooth using indirect vision, he/she has to do it so slowly to avoid unwanted generalization from the skilled learned in direct vision drilling. In contrast, an expert dentist can learn and access models for direct and indirect drilling independently.

In the literature, two types of generalisation have been addressed. First, the transfer component of generalisation has been investigated by practicing in one context and then testing in another (practice in task A→ perform task B), finding that transfer depends on two main things, the degree of similarity between the training and test episodes and on the ability of using a learned skill in a flexible manner. Second, the interference component of generalization has been investigated by trying to train participants to acquire and recall different motor context (perform task A→ practice in task B→ perform task A) (Adams, 1987; Krakauer, Ghez and Ghilardi, 2005; Krakauer *et al.*, 2006).

The topic of transfer becomes important when we want to understand how tasks contribute to, or interact with, each other in training situations, and it forms the basis of understanding such situations as those involving the use of simulators for learning some complex and lifesaving skills (Schmidt and Young, 1986). This behaviour can be explained by two distinct mechanisms by which the sensorimotor system might adjust its control parameters. First, generalisation could be a consequence of ‘learning to learn phenomenon” (discussed next), and second it could be the product of a smoother selection and execution process. Both mechanisms have been observed in sensorimotor control tasks showing that human can adapt and readapt to a sequence of similar tasks (Braun, Mehring and Wolpert, 2010; Gabriel, 2012; Diedrichsen and Kornysheva, 2015).

2.3.5 Structural Learning

Structural learning theory in the sensorimotor domain proposes that the characteristic features of a task can be abstracted from a set of examples, with the consequence that learning of similar tasks is facilitated (Braun *et al.*, 2010; Braun, Mehring and Wolpert, 2010). This "learning to learn" phenomenon is driven by reducing the uncertainty of the space that the learner has to search to adapt to novel tasks and explains the remarkable ability of humans to quickly adapt to new environments. For example, when we learn a motor skill, such as roller skating, we are more rapidly able to generalize to a novel task, such as ice skating (Braun *et al.*, 2009) in comparison to a novice who has not experienced roller skating.

In human sensorimotor research, studies show that participants who practiced walking tasks with a variety of distorting lenses performed better than participants practicing with a single distorting lenses on a novel walking task that required obstacle avoidance (Cohen, Bloomberg and Mulavara, 2005). Similarly, in a visuomotor task, Braun et al have shown that when participants are exposed to randomly varying tasks of the same structure, the sensorimotor system can extract the structure of the task and reduce the interference in the novel task (Braun *et al.*, 2009). These studies have also emphasized the importance of variability during training (given that this variability provides the learner with more exposure to potentially relevant structures in the task space that can be generalised to another related task) as well as explaining the role of structural learning in generalisation of sensorimotor skill.

2.3.6 Sensorimotor Performance and Learning Measures

While in common parlance the terms performance and learning are often used interchangeably, as mentioned earlier, they are two distinct terms in the field of sensorimotor control.

Performance is differentiated from learning by being defined as an interim change in sensorimotor action observed during or after practice. In contrast, learning is a set of complex processes (perception, cognition and action) associated with practise which lead to relatively permanent changes in the capability of producing skilled action (Tresilian, 2012; Shumway-Cook and Woollacott, 2014).

In preclinical dental courses, the purpose of training is not only to improve performance during practice, but also to facilitate the learning and transfer of the skills. The change in performance, either as an improvement or worsening level, might vary over time, which could be as a result of learning or some other factors such as motivation, and fatigue. To determine if the improvement is a long lasting learning effect, one may perform a retention test that provides a measure of the extent to which improvements made during the training phase are retained (Tresilian, 2012).

For the assessment of sensorimotor learning in dentistry there are two main approaches (Hauser and Bowen, 2009). The “scoring” approach involves measuring the final product without considering the rationale behind the student’s operative decisions. For example, a VR dental simulator may provide an objective measure of target area removal. This approach is important in assessment of sensorimotor skills that require precision and accuracy such as drilling. However, the provision of a percentage or volume of material removed inside or outside of a target area might not be a useful metric for the learner. This is because it cannot be used to indicate the causes of performance problems, which, as discussed earlier, is essential for motor learning (as well as for effective instruction) (Towers *et al.*, 2019).

The “kinematic” approach focusses on the motion of the learner by recording the movement of the hand while performing a sensorimotor task (such as the path length of the hand movement, smoothness, and velocity). This approach could provide an insight into the differences in behaviour between skilled and novice practitioners. The latest generation of haptic virtual reality systems are capable of providing both scoring and kinematic measurements automatically (Perry *et al.*, 2015) and we will make use of both approaches throughout the experimental work reported in this thesis.

2.3.7 Cognitive Control and Sensorimotor Learning

In this section we shift our focus to the cognitive mechanisms and the brain activity changes that underlie learning a sensorimotor skill.

Cognitive control refers to the processes that permit selection and prioritization of information processing in different cognitive domains to reach the capacity-limited conscious mind (Rabbi *et al.*, 2009; Wu *et al.*, 2016). In experimental psychology, cognitive control is often examined through recording the electrical activity generated by the brain and visible on the scalp. The observed signal is known as the electroencephalogram or EEG (Rabbi *et al.*, 2009). EEG can be used to continuously monitor levels of task engagement and mental workload in operational environments (Berka *et al.*, 2007; Holm *et al.*, 2009). EEG has a very high temporal resolution which allows the capture of some physiological changes that co-occur with continuous and intensive attention (da Silva, 2013; Cohen, 2017). Recent advances in wireless technology have started to make the recording of EEG outside experimental psychology laboratories possible. For example, a recent feasibility study used wireless EEG to monitor cognitive performance during venous cannulation. These authors found that collecting data on cognitive workload using EEG was possible and, moreover, brain activity was significantly greater in novice participants when compared with expert operators performing the same task (Lowe *et al.*, 2016).

Learning leads to neuronal recruitment – in other words, neurons not previously activated by the task become engaged. Early phases of learning are associated with increases in overall activity, followed by reductions in activity and neural variability in later phases (Minogue and Jones, 2006; Wolpert, Diedrichsen and Flanagan, 2011; Peterson and Robertson, 2013; Debarnot *et al.*, 2015). However, studies find that brain activity decreases after prolonged training (Costa, 2011; Diedrichsen and Kornysheva, 2015). Evidence also indicates that the consequences of sensorimotor skill acquisition and practise are often accompanied by considerable neuronal reorganisations within the motor and sensory brain areas (Jäncke, Shah and Peters, 2000). These signal decreases are interpreted as either a sign that a region has stopped to play a role in the production of the movement or it is also possible that the region continues to perform the same function, but has increased its efficiency and is able to use less neural activity (Costa, 2011; Diedrichsen and Kornysheva, 2015). Recent insights showed that many brain structures are recruited during task performance, but only activity in regions related to domain-specific knowledge distinguishes skilled from non-skilled performer (Peterson and Robertson, 2013; Debarnot *et al.*, 2015). For example, compared to non-musicians, professional pianists exhibit a decrease in the overall volume of brain activation in the primary and secondary motor area (M1, SMA, pre-SMA, and CMA) when performing bimanual and unimanual tapping tasks (Jäncke, Shah and Peters, 2000). We examine this relationship between cognitive control and performance of a dental task in Chapter 6.

2.4 Technology Enhanced Learning

Technology and learning have become closely intertwined in educational contexts. New technologies create learning opportunities that challenge traditional learning methods (Collins and Halverson, 2010). There is no doubt that in the 21st-century, technology has drastically impacted on education (Altbach, Reisberg and Rumbley, 2009). Communication of knowledge, distance learning, publication of journals, books, and e-books, and academic management are main areas sharing substantial changes according to The United Nations Educational, Scientific and Cultural Organization (Altbach, Reisberg and Rumbley, 2009). In recent years, the term ‘technology-enhanced learning’, or TEL, has become widely used in the UK and adopted by Higher Education Funding Council for England (Higher Education Funding Council for England, 2009) and the UK Higher Education Academy (HEA, 2009) to describe “the application of information and communication technologies to enhance teaching and learning” (Kirkwood and Price, 2014; Bayne, 2015).

Kirkwood and Price (Kirkwood and Price, 2014) characterise the desired enhancements that technology can provide for the learner. Technology can provide operational improvement by, for example, providing more flexibility for students and making the resources accessible anytime. Technology also provides qualitative and quantitative changes to the learning experience, for example through increasing engagement and achieving improved test scores or assessment grades. Indeed, evidence suggests that appropriate use of technology is leading to significant improvements in learning, teaching and assessment across the sector and this is translating into improved satisfaction, retention and achievement (Higher Education Funding Council for England, 2009). For example, preliminary data from a virtual plant cell project conducted by collaborators of our research group show that the assessment score for ‘recognising the cell and its components as 3D’ was 31% higher on average for students that had access to virtual plant lab, compared to Year 8 students who learnt the subject through conventional methods (*Virtual Plant Cell Evaluation*, 2019). Finally, emerging technologies clearly provide potential opportunities for enhancement and innovation in learning opportunities.

2.4.1 General Principle of TEL in Healthcare Education

The use of TEL in healthcare education has greatly expanded in the last decade (Thimbleby, 2013), and although there is plenty of evidence that technology enhanced learning has a place in dental education, it is often taken for granted that technologies can enhance learning without the commensurate evidence base to support these claims (Stein *et al.*, 2014). Technologies such as e-learning, smart phone applications, social media education, simulations and simulators, and other technologies provide a potential novel way for healthcare students, trainees and staff to acquire, develop and maintain the essential clinical and theoretical skills for safe and effective practice (Cook *et al.*, 2011). Education, training and the ongoing development of the healthcare students are integral to the improvement of patient outcomes, safety and experience. Indeed, wherever healthcare is delivered there must be ongoing teaching, learning, and evaluation.

All learners from novices to advanced practitioners can help to shape education process for their own needs, guided by general principle and qualitative and quantitative feedback data from faculty, clinicians, learners and even simulators. It is likely that technology will continue to reshape education in the years to come. Indeed, some authors believe that technology has already provoked significant changes in education and that there is a need to embrace such changes (Collins and Halverson, 2010). Thus, a framework with six principles developed by the UK's Department of Health (about TEL) was first released in 2008. This was followed by an updated version in 2011 which reviewed the current provision and use with input from students, trainees and staff, providing a guide for the use of TEL in healthcare education (Department of Health and Social Care, 2011). The framework focuses on "the use of technology as part of a managed learning process, appropriate expert supervision of students and trainees, particularly in clinical practice for ensuring patient safety" (Figure 2-1).

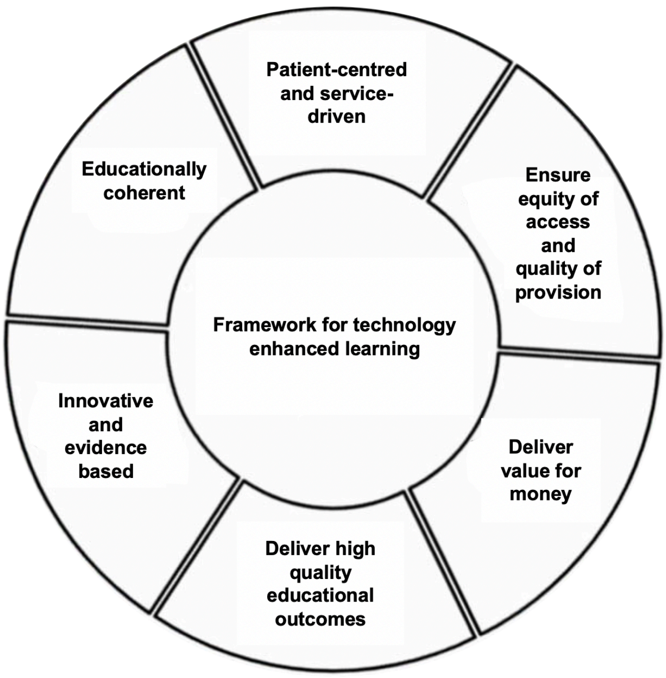


Figure 2-1 A framework for technology enhanced learning in healthcare education based on (Department of Health and Social Care, 2011).

The ultimate goal of healthcare technological applications is to equip the healthcare providers with the necessary skills for safe and effective patient care. Innovative technologies such as simulation and VR, have an important role to play as part of improving education. Reviews of postgraduate medical training have highlighted that trainees, in certain cases, feel that they are required to act beyond their level of competence (Department of Health and Social Care, 2011). Thus, trainees must have the opportunity to develop and improve their clinical skills, through TEL (such as clinical skills laboratories and simulated patient environments).

There is, however, a risk that technological applications are under-utilised. Curricula, learning frameworks, and applying principles from sensorimotor learning can play an important role in enhancing the implementation of innovative technologies. In addition to defining the learning outcomes and the assessments, curricula should describe how technological applications can support the development of knowledge and skills for each learning outcome (Department of Health and Social Care, 2011).

Before adopting new technologies to support learning there are a number of factors that need to be taken into consideration. For example, can they improve productivity, and are they adding extra value to the existing way of learning? The decisions of healthcare providers to use technological applications to support learning must be based on a clear understanding of the needs of learners and guided by the available evidence. Chapter 7 and Chapter 8 of this thesis will investigate how state of the art technology can be implemented in the assessment and learning of sensorimotor skill relevant to dentistry.

2.4.2 Simulators as an Example of TEL

Simulation is defined as “a methodology that replicates or amplifies real experiences with directed experiences using analogous tools or settings that imitate real world conditions, with the goal of learning and training, in an immersive and interactive mode” (Gaba, 2004; Littlewood, 2011). Simulation in healthcare education is not new. In fact, there are reports of simulation being used as a learning tool as early as the 19th century (e.g. teaching anatomy). In recent years, this has evolved into a distinct educational method (Bradley, 2006) and has been effectively used for training, assessment, and maintenance of various skills across diverse domains especially in complex professions which demand a high degree of precision and safety (Issenberg *et al.*, 2005).

Key to simulation is that it provides a standardised educational environment for training, assessment and maintenance of a wide range of skills across multiple healthcare disciplines, in a safe and ethical environment that enhances learning without jeopardizing patient safety (Gaba, 2004; Okuda *et al.*, 2009; Cheng *et al.*, 2016; Sevdalis *et al.*, 2016). This methodology can replicate real patient care scenarios in a controlled environment, to achieve pre-defined learning objectives, using artificial physical models, standardised patients or virtual reality devices for the purpose of improvement of individual and team performance in a health care system (Littlewood, 2011).

The majority of simulators in health care education are designed for the learning of procedural skills (minimally invasive surgery, obstetrics, tooth drilling), or soft skills (or non-technical skills) such as communication skills, team work, and decision making (Gaba, 2004). Simulation has become fully integrated into the clinical training of undergraduate and postgraduate students as well as for continuing professional development (Issenberg and Scalese, 2007). Beyond sensorimotor skill learning, simulation is needed to facilitate the transition into the clinic, to augment ergonomics and to enhance the students' preclinical experience through inclusion of a wide range of simulated patient scenarios emphasising a holistic approach to patient management (Hollis, Darnell and Hottel, 2011). Preclinical practice is essential for learning a sensorimotor skill and it may take several years to learn all the skills of a profession like dentistry. As such, simulation based research has grown exponentially in recent years to enhance the learning (Sevdalis *et al.*, 2016).

From G.V. Black's giant tooth models and Fergus's phantom head (Mason, 2005) to high fidelity virtual reality simulators and robotics, dental education has come a long way in the realism of the preclinical simulation experience, which continues to be an integral part of undergraduate dental education. Dental student education depends on high functioning teams and educators must ensure that technological approaches are used to support education process. Patients that are about to undergo a dental procedure want the dentist to be well-trained, competent, and experienced so that treatment can carried out quickly and effectively (Owen, 2016).

Today, phantom head simulators with typodont are considered the gold standard for undergraduate preclinical teaching, as well as for postgraduate skill training in most dental schools around the world (Gottlieb, Vervoorn and Buchanan, 2013). The phantom head simulator is a task trainer that facilitates the learning of fine sensorimotor skills and tooth preparation and restoration procedures in a safe environment (Fugill, 2013). They are reliable educational tools of relatively low initial cost that have been in use for a long time (Ben-Gal *et al.*, 2011). However, the plastic teeth used in these mannequins lack the real tactile sensation of natural layers of tooth structure (i.e. enamel and dentine) and there is a constant need for unit and handpiece technical maintenance as well as constant availability of disposable training resources (plastic teeth, burs, etc.).

With the continuous technological advances, computer assisted dental simulators have been developed based on virtual reality technology. In these systems, computer software is used to create a virtual environment that allows users to interact and navigate through similar challenges to those faced in real life.

Furthermore, the advancement of haptic technology has produced a step-change in the fidelity of VR dental simulators, fundamentally changing the way one interacts with virtual objects by providing realistic feel and touch sensation (Gottlieb, Vervoorn and Buchanan, 2013). Haptic VR simulators transfer the simulation experience, almost entirely, to the virtual world (i.e. no phantom head, plastic teeth, or real handpiece).

A unique feature of VR simulators is the availability of objective real-time feedback on student performance, in addition to the feasibility of iterative practice without the need for additional resources (plastic teeth, burs, etc.). Evidence is now accumulating that these types of simulators are particularly effective for formative assessment and evaluation that is facilitated by immediate and post practice feedback (e.g. video recordings), as well as in enhancing fine motor skill acquisition rate (Buchanan, 2001; Shahriari-Rad, 2013; Vervoorn *et al.*, 2015). Compared to traditional simulators, haptic VR simulators have also been reported to enhance the student learning via improved hand-eye coordination and self-reflection (Cox *et al.*, 2015) .

2.4.3 Tactile Sensation in the Human Hand and Haptic Technology

The “tactile sensation” in humans comprises three main parts - cutaneous, kinaesthetic and haptic - depending on site of sensory inputs (Dargahi and Najarian, 2004). The cutaneous sense receives sensory inputs from the receptors embedded in the skin and provides awareness of the outer surface of body. The kinaesthetic sense receives sensory inputs from the receptors within muscles, tendons, and joints which provides information about the static and dynamic body postures. Haptic sense refers to restoring sense of both tactile and force information. The haptic system uses significant information about objects and events both from cutaneous and kinesthetic systems. Haptics involves both action and reaction which is a two-way transfer of touch information to allow both “action for perception” and “perception for action” (Hagen *et al.*, 2008; Dahiya *et al.*, 2010; Culmer *et al.*, 2012; Tiwana, Redmond and Lovell, 2012).

In dentistry, we use our tactile information to understand, diagnose and treat disease (Macey *et al.*, 2018). Tactile sensing can be defined as a system that can measure a given property of an object or contact event, through physical contact between the system and the object to extract information such as temperature, vibration, softness, texture, shape and composition. A tactile sensor may measure one or more of these properties. Detection of a lesion using tactile sensors require the acquisition, processing, and display of tactile data (Tamè, Azañón and Longo, 2019). Touch can reveal shape (Longo and

Haggard, 2011), texture (Johnson and Hsiao, 1992), and other object properties. For example, consider the ease with which a dentist is able to perform a task such as caries removal. When drilling a tooth, the shape, size, location, colour and texture are transmitted to the brain from the sensory receptors. If the applied force is too great, the pulp could be perforated. A precise force needs to be applied and constant feedback of the measured applied forces keeps the pulp intact. In addition, a priori knowledge of the tooth's biophysical properties, such as the hardness of each tooth layer are also integrated into the cortical processing used for performing the task. Now, if the same task is to be performed using a simulator, then accurate sensory feedback is even more critical to provide the necessary feedback to explore and interact with objects (Dargahi and Najarian, 2004; Gwilliam *et al.*, 2010; Tiwana, Redmond and Lovell, 2012).

Originally, the word haptic comes from the Greek verb haptikos: to touch, implying the ability to touch and manipulate objects. In simulators, haptics is used to describe human-environment interaction via the sense of touch (Minogue and Jones, 2006). The need for implementing haptic touch has increased; especially due to the expectations of the virtual reality simulators. Having haptic feedback would enable analysis of tissue characteristics and pathological conditions. This interaction involves a two-way flow of data that allows the manipulation of the virtual objects in the virtual environment. Through force feedback and computer controlled haptic interface, haptic virtual objects are created which incorporate both cutaneous and kinaesthetic stimulation (Robles-De-La-Torre, 2010; Diego *et al.*, 2012). The haptic interface receives the motor signals followed by the generation of haptic feedback as a response. The combined integration of haptic, visual and auditory modalities result in “a degree of immersion” which facilitate the manipulation of virtual objects in the virtual environment through multiple sensory levels (Mihelj and Podobnik, 2012).

These systems are primarily focused on replicating the 'feel' of performing procedures, but this does not necessarily translate to efficient training (Roy, Bakr and George, 2017). Haptic technology potentially has substantial utility in promoting learning by directly manipulating movement (Williams, Tremblay and Carnahan, 2016; Clamann and Kaber, 2018), but, thus far, no simulators have exploited this potential to accelerating learning. In existing systems, the tasks, quality and quantity of haptic feedback are typically generic and at most have graded levels of difficulty to be completed in sequential order on the basis of subjective (self or teacher) imposed timelines. Recent research suggests that learning processes can be accelerated through tailored delivery of haptic feedback. Specifically, evidence suggests that whilst performance might be enhanced through haptic guidance, haptic forces that disrupt performance ultimately benefit learning (Wei *et al.*, 2005; Matsuoka, Brewer and Klatzky, 2007; Abdollahi *et al.*, 2011; Shirzad and Van Der Loos, 2012). We explore this novel application of haptics for motor learning in Chapter 8.

2.4.4 The Simodont Haptic VR Dental Trainer®

The Simodont® is one currently available haptic virtual reality dental simulator. This simulator was developed through a collaboration between Moog Inc. (Nieuw-Vennep, Amsterdam, The Netherlands) a company with expertise in flight haptic simulation and the Academic Centre for Dentistry, Amsterdam, the Netherlands). It is an educational VR simulator with a 3D display and high-resolution haptics, allowing the user to use tools to interact with virtual tooth models in the virtual space. All of the experiments described in this thesis make use of this simulator and, to contextualise that work, a detailed overview of the system is provided here.

2.4.4.1 Simodont® Hardware

The hardware of this device consists of a small screen (5” size with a 60 Hz refresh rate and 800 X 600 resolution) located in front of the trainee so that it simulates the patients’ head position. The screen supports a high-resolution stereo image facilitated by 3D projection. Magnification (zoom in/out) of the 3D display is possible up to 200% with full rotation of the virtual models around in the 3D display using the controller.

Underneath the screen is a physical handpiece with a handle that can be used as a virtual dental mirror. A foot pedal is used to activate the handpiece to start drilling (Bakker *et al.*, 2010). The speed of the virtual handpiece is controlled using the foot pedal. Once the participant presses the foot pedal the handpiece operates with realistic air rotor sound and the dental bur starts revolving. Once the bur comes in contact with the block or the virtual tooth the cutting takes place, providing that the participant presses on a specific area of the virtual tooth. When a virtual tooth is cut, multimodal simultaneous visual, audio and tactile feedback are received (De Boer, Wesselink and Vervoorn, 2013; De Boer *et al.*, 2015).

To obtain the 3D stereoscopic vision, the simulator is equipped with two digital multimedia projectors from LGTM (type HS101, resolution 800X600), which operate simultaneously resulting in projection of two images superimposed onto the screen through a polarizing filter. The trainee needs to wear passive polarised glasses for the image to be perceived as one 3D image, (Figure 2-2) shows the components.

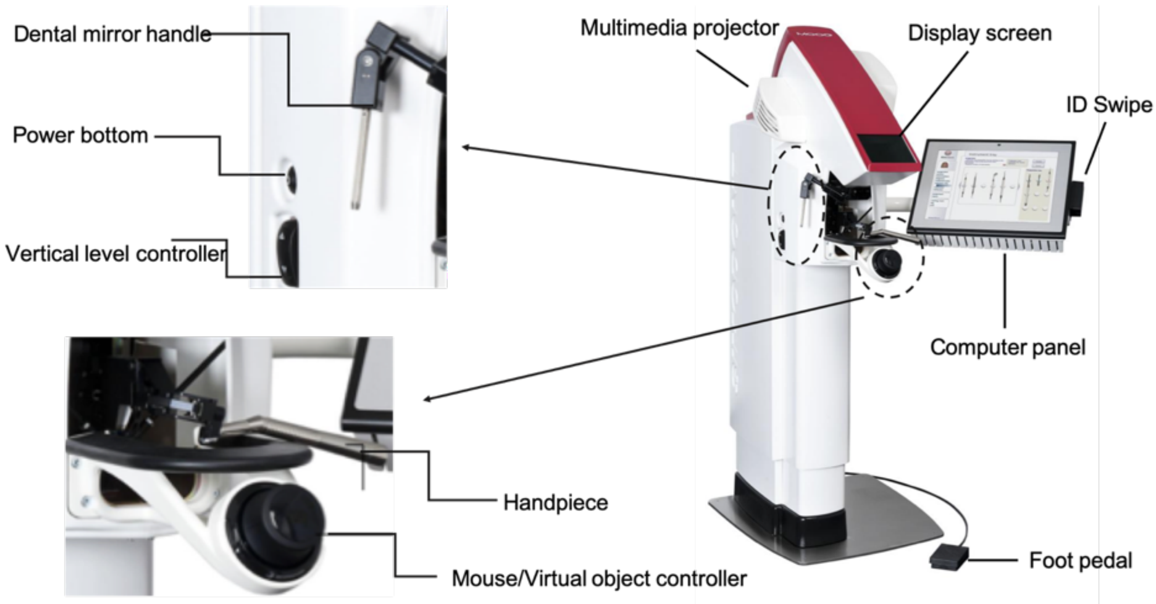


Figure 2-2 Simodont® simulator device with labelled components. Original image source from Moog (Moog Inc. 2011).

2.4.4.2 Simodont® Software

The Simodont® software, known as “courseware”, consists of lesson programs and modules with range of manual dexterity exercises, clinical operative dentistry procedures which includes caries removal, cavity preparation, crown and bridge preparation, and access cavity with varied levels of difficulty. The manual dexterity module offers automatic evaluation and records the real-time kinematics of student performance and can be seen on the attached computer screen (which includes the percentage of the target removed, the percentage of errors done to the sides and bottom of the leeway and container of the shape, total time, drilling time, and hand movement). Thus, participants are able to monitor their progress in real-time.

The available teeth library is derived from real extracted teeth (De Boer, Wesselink and Vervoorn, 2013). The virtual teeth library is expandable and editable, allowing for the addition of various shapes and sizes of virtual teeth with and without pathology with unlimited practice possibilities using dental cases of varied complexities, contributed by educators and researchers in some dental schools including ACTA and Leeds School of Dentistry.

2.4.4.3 The Simodont® Haptic Interface

The haptic interface provides force feedback based on the admittance control paradigm of a robotic arm, Haptic Master, developed by (MOOG). The simulator responds to force exerted by the user, leading to a sense that the user is interacting with an object of equal mass. The simulation of tooth cutting, and collision detection runs via the haptic interface so that the realistic force feedback in tooth cutting simulation is computed within only 1 millisecond (Bakker *et al.*, 2010). The haptic technology of the system allows for models to have varying density which changes the feeling of the interaction between the tool and the tooth. Parts of a model with low-density feel soft and are more easily removed (less force required) using the drill, and high-density materials are harder and require more force from the user to remove.

2.5 Summary

The purpose of this introduction was to prime the reader on (i) core concepts related to sensorimotor learning (from a behavioural, cognitive and neural perspective); (ii) provide a context in which simulation technology, designed to enhance sensorimotor skill acquisition operates; and (ii) finally, provide detailed coverage on the simulation technology that will be used in the following chapters to capture the processes underlying skilled dentistry and explore how such technology can be used to assess and accelerate learning.

Chapter 3 Examining the Limits of Learning in Dental Training

3.1 Abstract

Introduction: The surgical education literature makes extensive use of "learning curves" to identify when a performance plateau has been reached and analysis typically stops at this point. But do these plateaus best describe the limits of an individual's capabilities, or are continual improvements possible through extended practice? To address this question, we conducted a case study with one expert dental surgeon and one novice over 7 months of extensive practice.

Methods: The two participants, matched by age, gender and qualification level were asked to repeatedly perform a simple and advanced dental task on a high-fidelity virtual reality haptic dental simulator every weekday for seven months (n= 600 attempts), with their preferred and non-preferred hand. We examined whether their learning curves of performance changes (time and error) would be best described by an exponential function (indicating continual improvements in performance) or a power function (indicating performance tailing off over time).

Results: An exponential model provided a better fit for the expert (R^2 μ= .53, range = .339 - 673) and novice (R^2 μ = .25, range = .088 - .327) in the completion time relative to the power function (expert: R^2 μ =.29, range = .14 - .338; novice: R^2 μ =.173, range = .026 - .288). In contrast, for error rates, the power function accounted for more variance for the expert (R^2 μ =.219, range = .04 -.309) and novice (Power R^2 μ =.337, range = .228 - .443) relative to the

exponential function (expert: R^2 μ = .055, range = .009 -.074; novice: R^2 μ = .137, range = .123 - .146).

Conclusion: We found that while participants were unable to lower error without compromising time (and thus, this measure was best explained by a power law), there were small, ever increasing improvements in completion time for both novice and expert. These data show that performance can continue to improve long after supposed plateaus have been reached.

3.2 Introduction

There is an old adage that practice makes perfect. Whilst it is obvious to state that practice supports learning and, in most cases, produces relatively permanent changes in behaviour (Sands, 2017), describing the nature of human learning and identifying when perfection might be reached (or, in the case of surgery, a threshold level of competence to perform the task safely) is a century old discussion.

Over the previous century, there have been a number of theories of skill acquisition that have been proposed to account for the processes underlying human learning. Perhaps the most renowned is Fitts and Posner's (Fitts and Posner, 1967) three stage model of skill acquisition. Fitts and Posner proposed that the nonlinear changes observed in performance over time are due to individuals' transition from a cognitive stage (where there is a large amount of trial and error), through to an associative stage (where actions are much more targeted), and then finally, transition into a stage of learning referred to as autonomous (where performance appears to be flexible and easy) (Bryan and Harter, 1897; Fitts and Posner, 1967; Heathcote, Brown and Mewhort, 2000).

The first, "cognitive" stage involves the understanding of how to perform the task and how to evaluate the performance. This stage is characterised by rapid improvements in performance. This is related to the greater room for improvement, as the task is completely new, so there will be dramatic changes in performance (Fitts and Posner, 1967). As the learner practises, the degree to which there is room for improvement becomes smaller (Duong, Gardner and Rucker, 2010; Wolpert, Diedrichsen and Flanagan, 2011; Rodger, Tang and White, 2016). The associative stage is largely concerned with refining the baseline skills acquired during the cognitive phase. Performance becomes more consistent as the learner settles on a strategy. Finally, the autonomous phase takes place on a much longer time scale than the previous two phases, in which the motor skill become largely automatic and requires very little attention to perform (Fitts and Posner, 1967).

The primary assumption of most contemporary theories of motor skill acquisition is that learning is the result of the acquisition of more appropriate representations of actions for the desired goal (Newell, 1991; Wolpert and Flanagan, 2010). Specifically, the improved performance over time comes from the acquisition of prescriptions for action that specify the movement dynamics relevant to the task demands (Newell, 1991).

In healthcare, learning curves have often been used as an assessment tool for the adoption of new surgical procedures or technologies (Ramsay *et al.*, 2002; Dubrowski, 2005; Wulf, Shea and Lewthwaite, 2010; Harrysson *et al.*, 2014; Ben-Gal *et al.*, 2017). In research designed to demonstrate how learning curves can describe proficiency improvements associated with deliberate practice of radiograph interpretation, Pusic et al. explored how much practice is enough? Pusic et al. used learning curves to identify the inflection point (the point at which the rate of learning slows from an initial rapid phase to a slower phase during which each successive unit of practice results in less learning) (Pusic, Pecaric and Boutis, 2011). Harrysson et al.'s systematic review (including 592 studies) of learning curves in minimally invasive surgery shows that time and intraoperative outcome are used as metrics in learning curve. Time is the most commonly used proxy for the learning curve (508 study, 86%).

Intraoperative outcomes were used in (316 study, 53%) (Harrysson *et al.*, 2014).

In dentistry, the issue of more critically analysing learning curves surfaced recently when Ben-Gal et al. (Ben-Gal *et al.*, 2017) concluded that 12 weeks of practice (typically the length of a teaching semester in the UK) was not sufficiently long for students' learning to reach a plateau and thus be able to differentiate between students on the basis of their exhibited performance.

Long before their application in healthcare, learning curves were synonymous with a term referred to as the "power law of practice" (Wright, 1936). The power of law of practice states that the logarithm of the time taken to complete a task decreases linearly with the number of attempts made to complete said task (Anderson, 1982; Heathcote, Brown and Mewhort, 2000). A key theoretical (and practical) implication of the power law of practice is that it indicates that performance will eventually tail off to a point where no further improvements are possible. More recently, evidence that an exponential function may be better describe data from human learning experiments has begun to accumulate (Heathcote, Brown and Mewhort, 2000). In contrast to the power law, an exponential law indicates that the amount of performance improvement gained with each attempt decreases, but that improvements (however small) remain possible through extended practice, with a constant learning rate relative to the amount left to be learned.

To examine whether an exponential or power law would best describe the processes involved in skill acquisition for dentistry, we asked one novice participant with no experience of dental training, and one qualified dentist with extensive experience of clinical practice, to practice daily on a virtual reality dental simulator over a period of 7 months.

We expected that the novice participant would have a higher starting point for both time to complete the task, and frequency of errors but would show faster learning (due to the greater space for improvement). We also expected to find that the expert would be able to more quickly adapt to the task demands, with faster times and fewer errors. Our primary focus, however, was after this initial steep reduction in performance and how performance varied over a protracted period of practice. Specifically, our objective was to identify whether the learning curves for these individuals at different ends of the dental training spectrum would be best described by an exponential or power law. If the power law best described the data, this would indicate participants had reached a limit, with further training providing no significant changes in improvement. In contrast, if the exponential law best fit the data, it would indicate participants were showing (ever smaller) improvements over time.

3.3 Materials and Methods

3.3.1 Participants

One qualified dentist (male, 32 Years) undertaking a postgraduate degree and a non-specialist (an individual with no experience of using handpieces in clinical dentistry), matched by age, gender and education level (male, 32 Years) took part in the study. Both participants were right-handed and provided informed consent. The study was approved by the ethics committees based in the School of Psychology at the University of Leeds, United Kingdom (Reference number: 17-0166; date approved: 30-May-2017).

3.3.2 Experimental Protocol

The participants performed two tasks with different levels of complexity. The tasks were selected from the manual dexterity module in The Simodont Haptic VR Dental Trainer® (Figure 2-2), the virtual reality haptic dental simulator described in and elsewhere (Mirghani et al., 2016; Al-Saud et al., 2017).

The participants completed the drilling of a simple straight shape task (Figure 3-1, a), and a more advanced cross-shape (Figure 3-1, b). Each shape comprised three zones: a target zone- which must be removed by the participant; Leeway zones (side and bottom) which surrounded the target zone - participants were instructed to avoid removing this if possible; and container zones (sides and bottom), which were represented by a block surrounding the abstract shape that participants were told they must avoid during target removal.

The goal of each trial was to use the connected dental handpiece to remove 99% of the target 'red zone' in the middle of a block whilst attempting to minimise removal of leeway zones (the ‘safe’ outer areas of the block) as much as possible. Real-time feedback on performance was presented on a computer monitor attached to the device throughout the task. The feedback information included a percentage score for each of the following: target (task completion percentage), error scores (leeway bottom, leeway sides, container bottom and container sides) and drill time (in seconds).

To avoid confounding order effects, we counterbalanced the task shape and hand order using a Latin squares design that gave us four possible options; simple task performed by the preferred hand, advanced task performed by the preferred hand, simple task performed by the non-preferred hand, and advanced task performed by the non-preferred hand (Figure 3-1, c). The participants performed the experimental tasks every weekday for seven months.

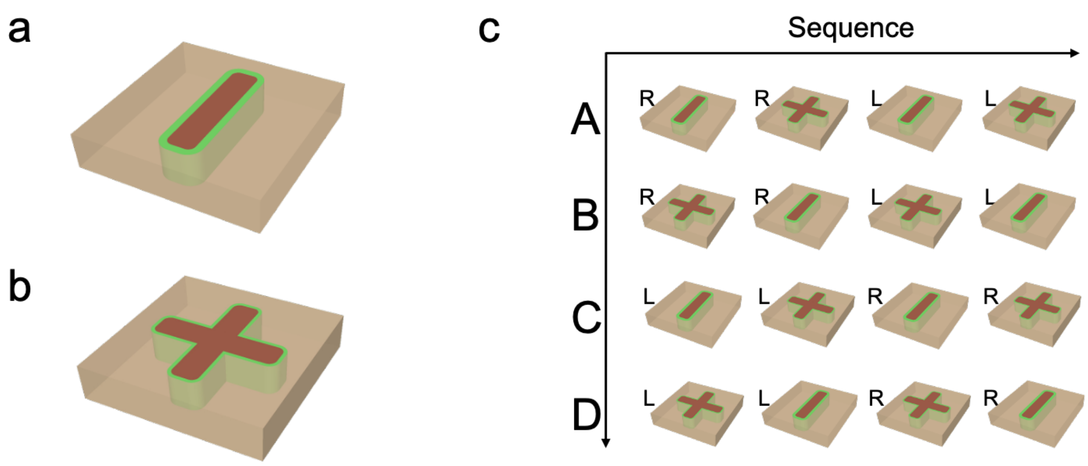


Figure 3-1 Schematic drawing of the experimental setting and the dental tasks.(a) Straight - shape simple task. (b) Cross-shape advanced task. (c) The four option (A, B, C, and D) of counterbalanced between the task shape and hand order using the Latin squares, each row representing the sequence of practice in a day. R is refereed to right hand and L is refereed to left hand.

3.4 Statistical Analysis

For data analysis, we measured performance on two outcome variables. Time (in seconds) and total error (as a percentage). The total error score was calculated as a mean average of the leeway area removed from the side and bottom of the block.

The exponential (Equation 1) and power (Equation 2) equations applied to learning curves are of the following general form:

$$\mathrm{y} = a * e^{b*x} \tag{1}$$

$$\mathrm{y} = a * x^{b} \tag{2}$$

Where, (y) is some measure of learning, (e) is the base of the natural logarithm (2.718), and (x) is the number of trials of training.

The equations also contain two free parameters (a and b), which can be adjusted for each particular data set. Parameter (*a)* defines the predicted asymptote for performance, capturing a subject's initial performance. Parameter (*b)* is the decay parameter describing the learning rate. We compared the amount of variance explained (R^2) for each of the models, for each participant separately for each task (simple and advanced, for the preferred and non-preferred hand). We also averaged the (R^2) of all the four tasks for each participant and compare the value of exponential and power laws (Average exponential R^2 of the four tasks in comparison to average power R^2 of the four tasks for each participant). All the analyses were performed using MATLAB Version R2017b.

3.5 Result

We first examined the time to complete the task dependent variable in the expert participant dataset. We found that overall, the exponential function (R^2= .53) accounted for more variance than the power function (R^2=.29).

Next, we examined each task, performed by each hand separately. For the simple task, performed using the preferred hand, the R^2 was .339 for the exponential function and .14 for the power function (Figure 3-2, a). Similarly, for the simple task performed with the non-preferred hand, the R^2 was .557 for the exponential function and .305 for the power function (Figure 3-2, b).

For the advanced task performed with the preferred hand, the R^2 was .552 for the exponential function and .331 for the power function (Figure 3-2, c). For the advanced task performed with the non-preferred hand, the R^2 was .673 for the exponential function and .384 for the power function (Figure 3-2, d).

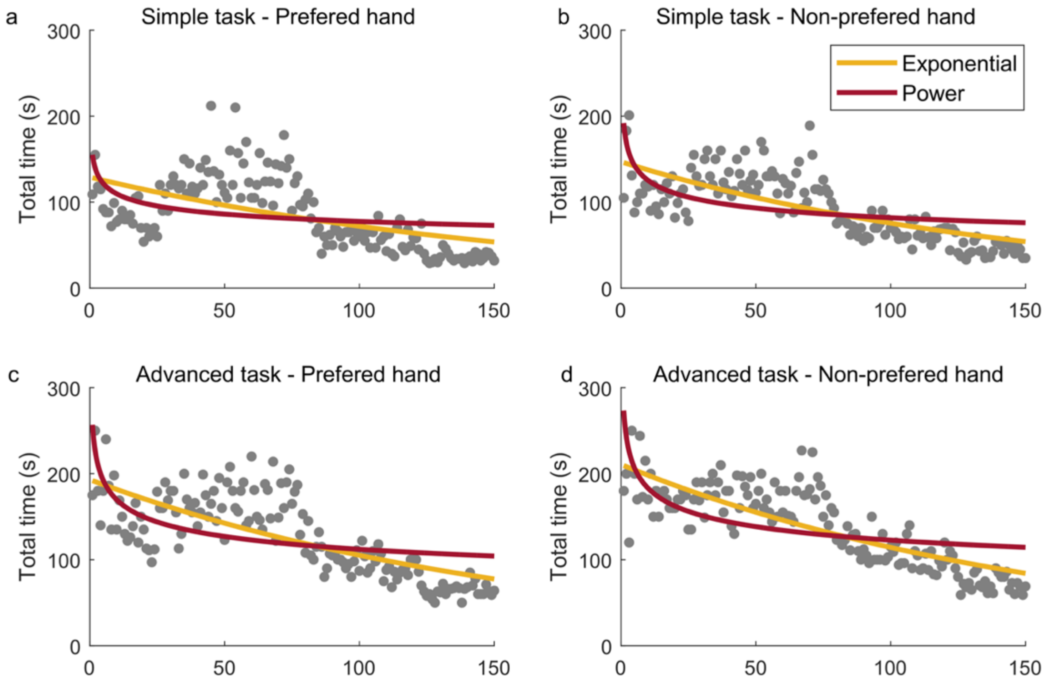


Figure 3-2 The result of the time variable in the expert data set. The grey dots representing the daily performance, the yellow line representing the exponential law curve fitting, and the red line representing the power law curve fitting (a) the simple task performed with the preferred hand, (b) the simple task performed with the non-preferred hand, (c) the advanced task performed with the preferred hand, (d) the advanced task performed with the non-preferred hand.

The novice dataset results mirrored the expert. The exponential function accounted for more variance (R^2 = .25) than the power function (R^2 = .173) for time to complete the task.

For the simple task performed with the preferred hand, the R^2 was .283 for the exponential function and .288 for the power function (Figure 3-3, a). For the simple task performed by the non-preferred hand, the R^2 was .088 for the exponential function and .026 for the power function (Figure 3-3, b).

For the advanced task performed with the preferred hand, the R^2 was .302 for the exponential function and .228 for the power function (Figure 3-3, c). For the advanced task performed with the non-preferred hand, the R^2 was .327 for the exponential function and .153 for the power function (Figure 3-3, d).

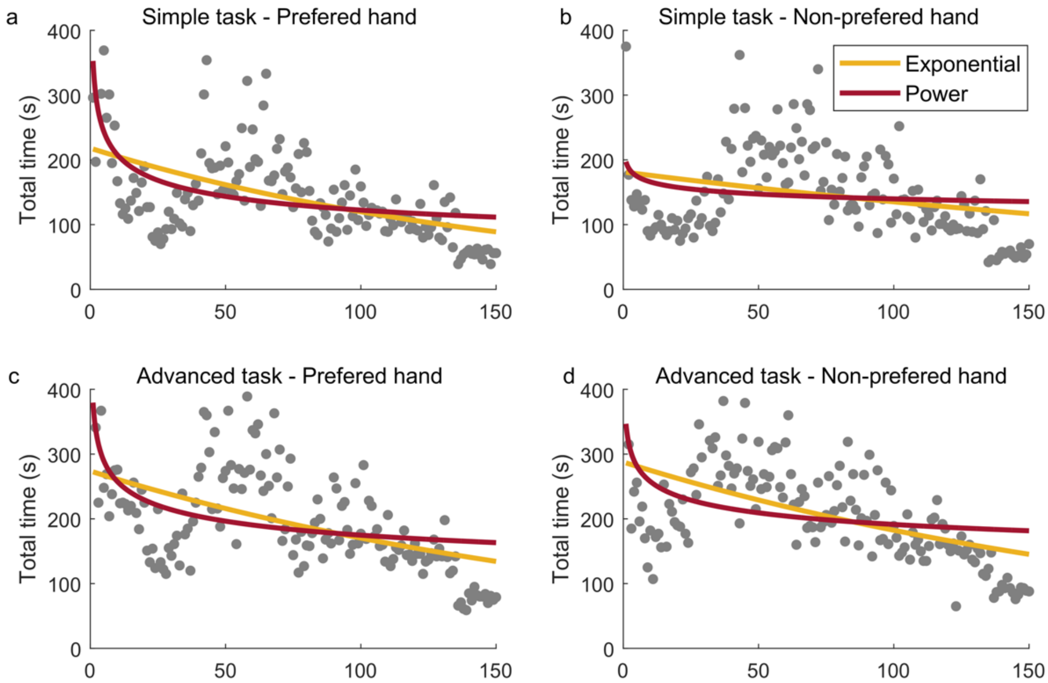


Figure 3-3 The result of the time variable in the novice data set. The grey dots representing the daily performance, the yellow line representing the exponential law curve fitting, and the red line representing the power law curve fitting (a) the simple task performed with the preferred hand, (b) the simple task performed with the non-preferred hand, (c) the advanced task performed with the preferred hand, (d) the advanced task performed with the non-preferred hand.

For the error measurements in the expert data set, in contrast to the time variable, the power function (R^2 = .219) accounted for more variance than the exponential function (R^2 = .055). We probed this in more detail by examining the separate tasks, performed with each hand.

For the simple task performed with the preferred hand, the R^2 was .095 for the exponential function and .309 for the power function (Figure 3-4, a). For the simple task performed with the non-preferred hand, the R^2 was .042 for the exponential function and .219 for the power function (Figure 3-4, b). For the advanced task performed with the preferred hand, the R^2 was .009 for the exponential function and .040 for the power function (Figure 3-4, c). For the advanced task performed with the non-preferred hand, the R^2 was .074 for the exponential function and .237 for the power function (Figure 3-4, d).

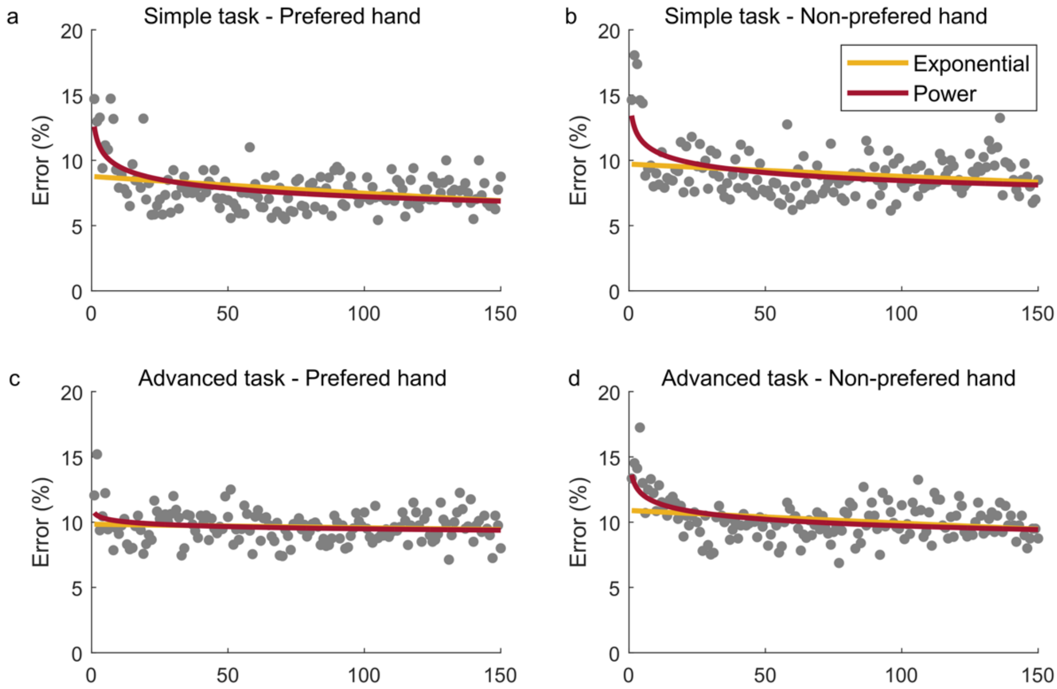


Figure 3-4 The result of the error variable in the expert data set. The grey dots representing the daily performance, the yellow line representing the exponential law curve fitting, and the red line representing the power law curve fitting (a) the simple task performed with the preferred hand, (b) the simple task performed with the non-preferred hand, (c) the advanced task performed with the preferred hand, (d) the advanced task performed with the non-preferred hand.

In the novice data set, the power function also accounted for more variance than the exponential function in the error variable. Overall, the average R^2 was .137 for the exponential function and .337 for the power function. For the simple task performed with the preferred hand, the R^2 was .123 for the exponential function and .285 for the power function (Figure 3-5, a).

For the simple task performed with the non-preferred hand, the R^2 was .138 for the exponential function and .392 for the power function (Figure 3-5, b). For the advanced task performed with the preferred hand, the R^2 was .141 for the exponential function and .228 for the power function (Figure 3-5, c). For the advanced task performed with the non-preferred hand, the R^2 was .146 for the exponential function and .443 for the power function (Figure 3-5, d).

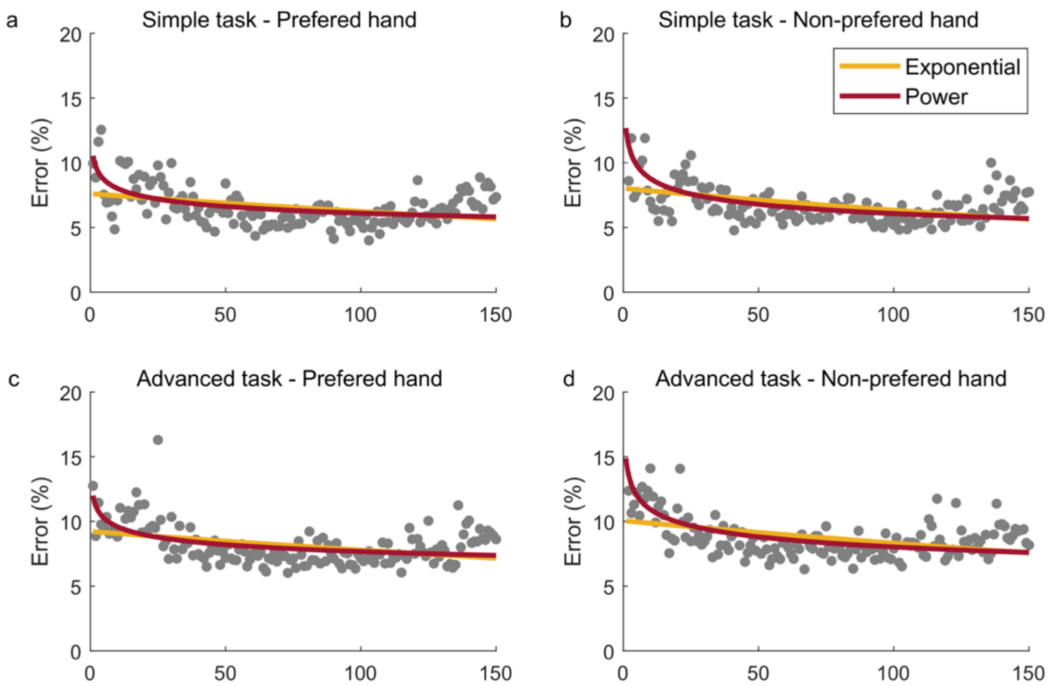


Figure 3-5 The result of the error variable in the novice data set. The grey dots representing the daily performance, the yellow line representing the exponential law curve fitting, and the red line representing the power law curve fitting (a) the simple task performed with the preferred hand, (b) the simple task performed with the non-preferred hand, (c) the advanced task performed with the preferred hand, (d) the advanced task performed with the non-preferred hand.

3.6 Discussion

Learning curves are often used to describe the changes produced through practice and, in the healthcare field, used as a tool to evaluate motor learning (Harrysson *et al.*, 2014). In the present study, we sought to probe the underlying nature of learning curves in order to better understand the relationship between practice and learning. Specifically, we asked a novice and expert dentist to practise repeatedly on a VR haptic simulator and examined their learning curve profiles by asking which of two non-linear functions (power vs. exponential) could best account for the data. We reasoned that if a power law was a superior fit, this would indicate that performance would hit a ceiling and tail off with no further improvements possible (Newell and Rosenbloom, 1981). However, if the exponential function provided the best account of the data, it would imply a constant learning rate that relatively decreased with further practice.

Overall, we found that the exponential function accounted for more variance than the power function in both the expert and novice data sets, for the time variable. In contrast, for the error measure, there was a reversal in pattern, with the power function accounting for more variance than the exponential function in all conditions. We examine the theoretical and practical implications of these results next.

Overall, the data from the present study align well with the general law of learning for the motor domain (Newell, Liu and Mayer-Kress, 2001). Recently, Heathcote et.al examined 40 sets of data representing 7,910 learning series from 475 subjects in 24 experiments from a wide range of tasks. When they directly compared power and exponential functions as possible forms for a general law of practice, they found the exponential functions provided better fits than power functions in all unaveraged data sets (Heathcote, Brown and Mewhort, 2000). Our data on time are consistent with these results, with even an expert, after prolonged practice, able to continually make improvements over time.

Interestingly, in our case, the power law did provide a better account of the error dependent variable. It is worth noting that error is zero-bound and there are both physical limits and potentially machine related limits – indeed, recent work has shown that even expert performers have non-zero error performance on this machine (Wierinck *et al.*, 2007; Mirghani *et al.*, 2016). Given the participant instructions to reach a very high level of target completion (i.e. 99%), it is unclear whether speed-accuracy trade-offs may be manifesting and accounting for these results (Hendee, 2001).

In the introduction, we described a well-known model of human learning, Fitts & Posner’s 3 stage model of skill acquisition and introduced the different stages of learning which achieved by the performers as they become able to produce movements using less motor planning or preparation time. The time taken to perform the task decreases and accuracy improves as the performers gain more skill. It is worth inspecting the novice participant data from this viewpoint. In the first 40 days of practice, we found a dramatic decrease in the time measurements followed by increase in time to complete the task for the next 40 days, with a gradual decrease after that. The data from the first 80 days align extremely well with the transition from an early stage- which is often characterised by inconsistency, hesitation and lack of confidence even with improve in performance- through to an associative stage- when performance starts to become more accurate and consistent. By the end of the experiment, given how well the participant was able to perform the tasks with their preferred hand and their non-preferred hand on both simple and difficult task, it seems probable that this stage of performance was characterised by the third and final automatic stage of learning.

Finally, we note that much of the research on human learning has employed cross-sectional designs over relatively short durations of time. These types of studies are important in finding differences among various groups, however, they lack the ability to trace changes within individuals over period of time. Whilst extensive long-term practice studies on a small number of individuals was the norm in the early stages of experimental psychology (Boring, 1954; Smith and Little, 2018), these types of studies are now rarely conducted. Despite the small sample sizes that these types of experiments often necessitate, there is much value that can be gained from this small-study approach when the data are interrogated and interpreted carefully (Hackshaw, 2009). The case study approach we used in this experiment has allowed us to better understand the mathematical relationship between practice and learning in a VR dental task.

Chapter 4 Differences in Planning and Efficiency as a Function of Expertise in a Sensorimotor Dental Surgery Task

4.1 Abstract

Aim: Most approaches discriminating between behavioural measures of performance focus on error frequency and time to complete a task. Here, we explore the value of taking measures of idle time (the difference between time spent drilling and time to complete the task) and hand trajectory path length. We hypothesise that these measures could provide useful insights into the processes involved in planning and efficiency, respectively during the task and compare across a group of experienced and novice participants on a dental surgery task.

Methods: We recorded behavioural performance from experienced dental students (N=41) enrolled on the dentistry programme the School of Dentistry and novice participants (N=38) enrolled on a variety of undergraduate programmes including Engineering, Psychology, and Medicine at the University of Leeds. Data were recorded while they performed a drilling dental task on a high-fidelity virtual reality dental simulator.

Results: We found differences in idle time [$F(1, 77) = 39.49$, $p < 0.0001$, $\eta_G^2 = 0.283$] and path length [$F(1, 77) = 43.77$, $p < 0.0001$, $\eta_G^2 = 0.362$] across the two groups. Novice participants (M = 94, 95% CI = [76.7, 111.2]) took significantly more time than experts (M = 27.6, 95% CI = [11, 44.2]) and their hands moved over longer distances (M = 2.38, 95% CI = [2.14, 2.63]) relative to experts (M = 1.27, 95% CI = [1.03, 1.50]). Time correlations revealed a significant positive

correlation with path length and idle time in both the novice ($p < 0.0001$) and expert groups ($p < 0.0001$). Error correlations revealed a significant positive correlation with path length in the novice group [$r = .35$, $n=39$, $p=.032$], and positively with idle time in the experienced group whilst completing the task [$r = .34$, $n=41$, $p=.029$].

Conclusion: The data indicate that experienced performers have superior economy and shorter planning times. We propose that idle time and path length could be useful adjuncts to the oft-reported metrics of time and error and that their inclusion in the dental education training program could be useful for mentors in providing more specific guidance to trainees to optimize the learning process.

4.2 Introduction

Expert sensorimotor performance requires extensive, intensive training (Ericsson and Lehman, 1996; Wierinck *et al.*, 2007). Through this process, individuals transition from slow, clumsy movement patterns to highly coordinated, fast and efficient execution of goal-directed behaviours (Newell, 1991; Schmidt and Wrisberg, 2000). Expertise in dental surgery requires a particularly unique set of skills that are honed over years of training that allow a clinician to create geometrical shapes in small dimensions, with limited workspace, visibility and distraction force by the tongue and cheek muscles (Dimitrijevic *et al.*, 2011; Gottlieb, Vervoorn and Buchanan, 2013).

The pre-clinical dental training is fundamental for novice dental students to gain familiarity with the dental operations and to master dexterous sensorimotor skills (Buchanan, 2001). It may be obvious but is worth stating that on average, an expert dentist can perform a dental task more accurately and faster than a novice (Wierinck *et al.*, 2007; Suebnukarn *et al.*, 2009). In recent work, we showed that when assessing dental students' performance in virtual reality simulators for every one unit increase in training year, performance on the composite error score decreased by an unstandardised beta coefficient value of 0.519 (Mirghani *et al.*, 2016). However, these measurements of the final product (error and time) do not consider the performer's strategy or efficiency while performing the task (Schmitz *et al.*, 2014; Towers *et al.*, 2019).

Idle time and path length are two potential measurements that can be used to differentiate between novice and expert performers in the planning and efficiency of sensorimotor execution. Idle time is characterized by a lack of movement of the performer's hands and may represent periods of motor planning or decision making (Oropesa *et al.*, 2011, 2013; D'Angelo, Rutherford, Ray, Laufer, *et al.*, 2015).

Examining idle time may reveal more about the relationship between sensorimotor performance and learning stage than time to completion alone (Fitts and Posner, 1967). Various models of sensorimotor skill acquisition propose that as an individual progresses through stages of learning, there is a reduction in the cognitive processes related to the sensorimotor planning (Wolpert, Diedrichsen and Flanagan, 2011; Tresilian, 2012; Diedrichsen and Kornysheva, 2015; Krakauer *et al.*, 2019).

This adaptive process, the transition from high cognitive demand to low, is orchestrated by highly integrated neural circuits (Banks, Mikell and McKhann, 2014). As a result, performers are able to produce movements with less motor planning or preparation time and end up with relatively autonomous performance at an execution level (through the formation of a new motor primitive in the brain) (Diedrichsen and Kornysheva, 2015). Thus, examining those instances where there is planning, but no movement, could be a useful avenue for probing learning.

Beyond planning, there may also be value in delineating motor execution (i.e. physical mechanics of movement) by measuring efficiency. Motor efficiency can be thought of as the conservation of time and motion and is often defined by the path length and, in dentistry, we can think of computing efficiency through tracking the total path followed by the tip of the handpiece instrument in dimensions x, y and z. There is a large body of research that has shown a gradual reduction in movement variability as the accuracy of an action improves

(Shumway-Cook and Woollacott, 2014; Komar, Seifert and Thouvarecq, 2016), with repeated movements correlated with shorter path lengths (Van Beers, 2009). In surgery, in specific tasks such as suturing, it has been shown that motor efficiency improves with each subsequent trial (D'Angelo, Rutherford, Ray, Mason, *et al.*, 2015). This measure of motor efficiency may also provide some insight into individual differences in performance (Sparrow, 1983; Ericsson, Krampe and Tesch-Römer, 1993).

As expert performance is characterised by smoothness (and driven by relatively automatic selection and execution of sensorimotor commands (Krakauer *et al.*, 2019)), we expect to find differences in both planning and efficiency of task execution as a function of surgical experience. In this study, we asked qualified dentists, and control participants to perform a simulated dental task on a high-fidelity virtual reality dental simulator in order to explore behavioural differences in these two metrics. We predicted that novices would produce a greater error percentage and require longer time to complete the task in comparison to experts. Second, we predicted that the experienced group would display a greater ability to make appropriate decisions with less idle time, as well as completing the task with less path length.

Finally, to examine whether these measures add any value above and beyond time and error, we correlated idle time and path length with error and time. We reasoned that weak relationships would indicate that performance in idle time and path length cannot be fully accounted for by error and time.

4.3 Materials and Methods

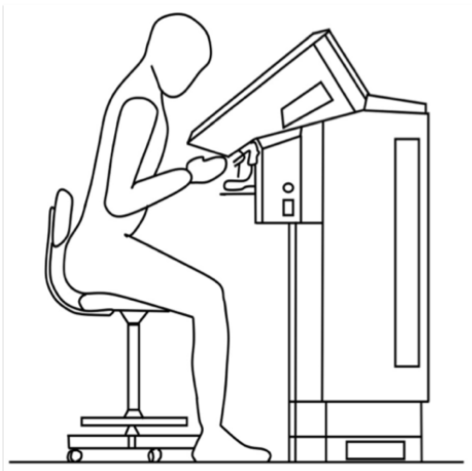
4.3.1 Participants

Experienced dental students (n=41, Male=16, Female=25, Average age= 30.45, SD=±3.77) enrolled on the final year of undergraduate study and in clinical postgraduate dentistry programme at the School of Dentistry and novice participants with no experience of using handpieces in clinical dentistry (n=38, Male=19, Female=19, Average age= 31.97, SD=±7.6) enrolled on different programmes (Including; Engineering, Psychology, Medicine) at the University of Leeds took part in this study. Data were recorded while they performed a drilling dental task on a high-fidelity virtual reality dental simulator. All participants expressed a preference to use their right hand for the task. Participants provided informed consent and were fully debriefed. The study was approved (Reference number: 17-0166; date approved: 30-May-2017) by the ethics committees based in School of Psychology at the University of Leeds, United Kingdom.

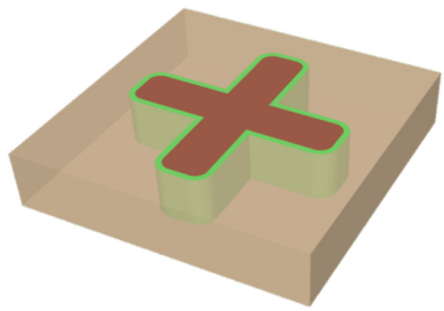
4.3.2 Experimental Protocol

Participant performed a drilling task from the manual dexterity module in The Simodont Haptic VR Dental Trainer®. A cross shape task was employed which consisted of three zones: a target zone which must be removed by the participant; Leeway zones (side and bottom) surrounding the target zone which the participants were instructed to avoid removing; and the container zones (sides and bottom) represented by a block that surrounds the abstract shape that participants were also told they must avoid during target removal (Figure 4-1).

The participants were required to complete drilling task which involved the use of a dental handpiece to remove the target ‘red zone’ in the middle of a block, whilst attempting to minimise removal of leeway zones (the ‘safe’ outer areas of the block). Real-time feedback on performance was presented on a computer monitor attached to the device throughout the task. The feedback information included target percentage, error percentage, and time in seconds. Participants were instructed that the aim of the task was to remove 99% of the target area “without touching the green and the beige zone as much as they could and as fast as possible”. Once this was achieved the participants were asked to stop drilling and the data were recorded.



(A)



(B)

Figure 4-1 Experimental setting and task. (A) Schematic drawing of the VR dental simulator and the experimental setting. (B) the cross-shape task, illustrating the location of the target area, the leeway and container.

4.4 Statistical Analysis

For statistical analysis, we measured performance on four outcome variables. Time (in seconds), total error (as a percentage), idle time (in seconds), and pathlength (in meters). The total error score was calculated as an average of leeway and container areas removed (side and bottom). The idle time score was calculated by subtracting drilling time from total time.

A One-way ANOVAs were conducted to compare the performance of participants on the dental task according to the participant group (novice and experienced) for each of the outcome variables (idle time, path length, total time, and error). The statistical significance threshold was set at $p < .05$ and we report generalised eta squared (η_G^2) as a measure of effect size. We considered η_G^2= 0.02 to be small, η_G^2=0.13 to be medium and η_G^2= 0.26 to be a large effect size. All statistical analyses were performed using RStudio Version 1.1.463 (R Foundation for Statistical Computing., 2018).

Correlation analyses were used to examine the relationship between idle time and path length against error rates and time to complete the task for each group using Pearson's correlation. We undertook these correlations to evaluate whether the variance in idle time and path length could be captured by more commonly used measures of performance (time and error). A comparison of the magnitude of correlations was performed where significant correlations were found using Hittner, May, and Silver's (2003) modification of Dunn and Clark's (1969) approach using a back-transformed average Fisher's Z procedure in cocor package in RStudio (Diedenhofen and Musch, 2015).

4.5 Result

As expected, we found a significant effect of group on error rates [F (1, 77) = 66.68 , p <0.0001, η_G^2= 0.464] (Figure 4-2, A), with percentage error being higher in the novice participants (M = 27.5%, 95% CI = [24.9, 30.2]) relative to the experienced participants (M = 12.7%, 95% CI = [10.1, 15.2]).

There was also a significant effect of time, [F (1, 77) = 51.15 , p <0.0001, η_G^2= 0.399] (Figure 4-2, B), with longer time to complete the task in the novice participants (M = 274 seconds, 95% CI = [251, 298]) relative to the experienced participants (M = 157 seconds, 95% CI = [134, 179]).

For the idle time, there was also a significant group effect, [F (1, 77) = 30.49 , p <0.0001, η_G^2= 0.283] (Figure 4-2, C), with idle time higher in the novice group (M = 94 seconds, 95% CI = [76.7, 111.2]) relative to the experienced group (M = 27.6 seconds, 95% CI = [11, 44.2]).

Finally, for the path length, we found the same pattern as our other behavioural measures, with a large effect of group, [F (1, 77) = 43.77 , p <0.0001, η_G^2= 0.362] (Figure 4-2, D) and longer path length for novices (M = 2.38 meters, 95% CI = [2.14, 2.63]) relative to experienced dentists (M = 1.27 meters, 95% CI = [1.03, 1.5]).

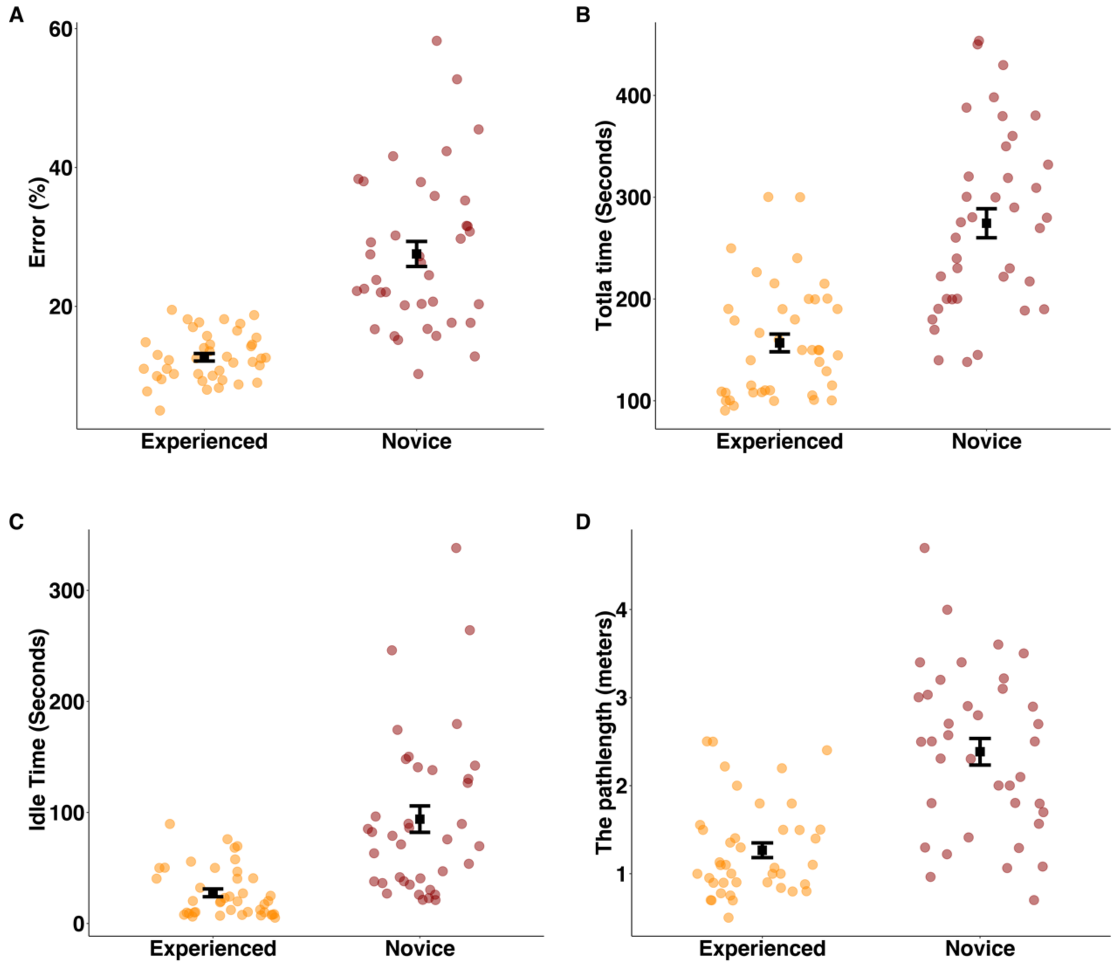


Figure 4-2 The differences in outcome variables between groups. The circles represent the participants performance, the squares are the mean and the error bars are the stander error. (A) A comparison between the performance of each group in the dental task represented the error measurement in which the lower the value the better the performance. Notably, the experienced group on average performed significantly better than novice group. (B) The time measurement (seconds) shows that the experienced group on average performed significantly faster than novice group. (C) The idle time graph shows that the experienced group on average complete the task with less periods of motor planning or decision making than novice group. (D). A comparison between the path length of each group in meter which shows that the experienced group on average complete the task with less than half of the novice average path length.

Next, we examined the extent to which idle time and path length correlated with time to complete the task and error rates. We found a significant positive correlation between path length and total time in both the novice group [$r = .76$, $n=39$, $p < 0.0001$], and expert group [$r = .7$, $n=41$, $p < 0.0001$] (Figure 4-3, C). Also a significant positive correlation between idle time and total time in both the novice group [$r = .74$, $n=39$, $p < 0.0001$], and expert group [$r = .75$, $n=41$, $p < 0.0001$] (Figure 4-3, D).

There was a significant positive correlation between path length and error rates in the novice group [$r = .35$, $n=39$, $p=.032$], but this correlation was weaker and not statistically significant in the experienced group [$r = .14$, $n=41$, $p=.38$] (Figure 4-3,A). We note, however, that in a comparison of the two correlations, we did not find a statistically significant difference [$z = .95$, $p = 0.337$]

Finally, there was a significant positive correlation between idle time and error rates in the experienced group whilst completing the task [$r = .34$, $n=41$, $p=.029$], but no relationship emerged for the novice group [$r = -.15$, $n=39$, $p=.38$] (Figure 4-3, B). When comparing the magnitude of the two correlations we found that these patterns were reliably different to one another [$z = -2.154$, $p = .031$].

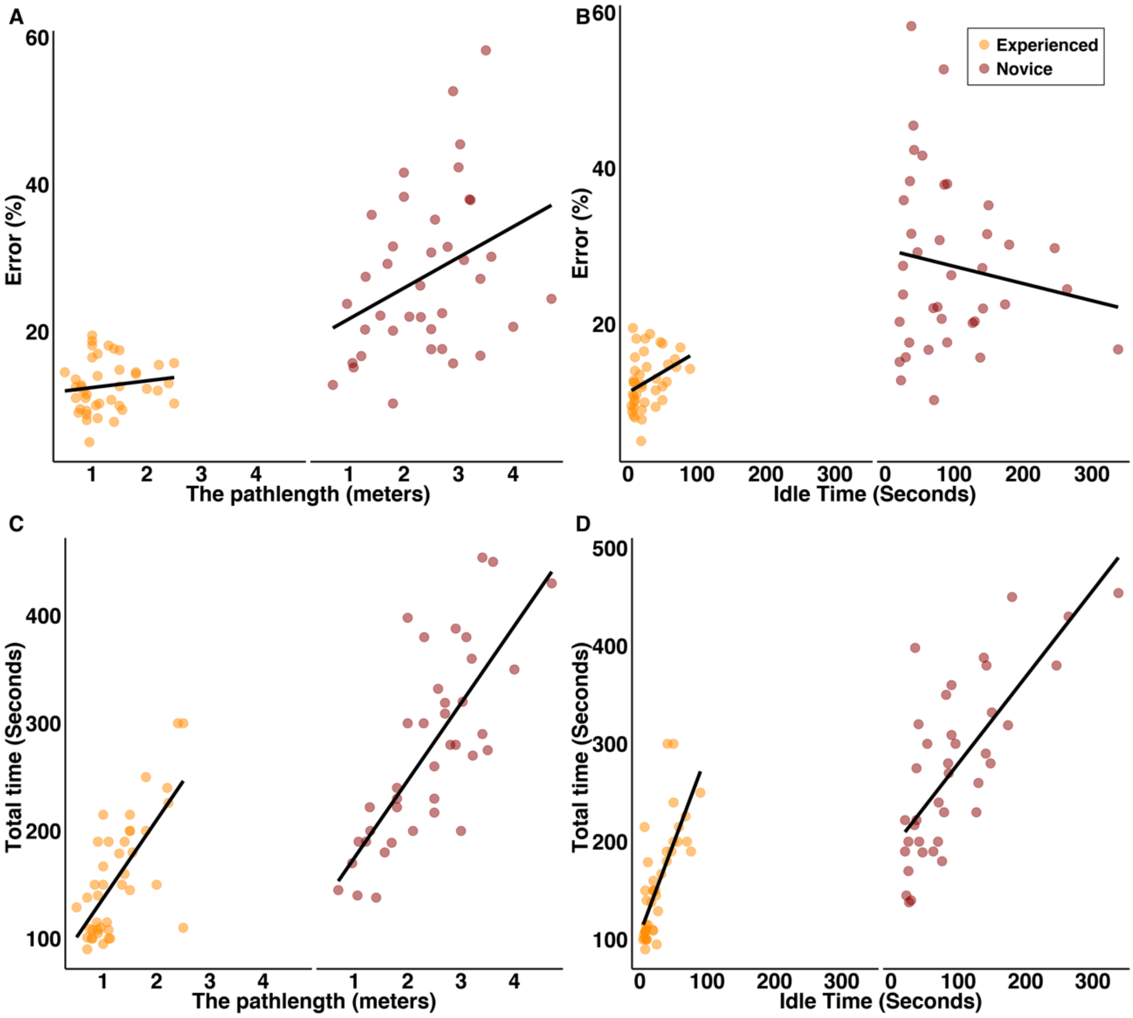


Figure 4-3 Scatterplots summarize the relationship between performance errors with both path length and idle time. (A) The correlation between path length and error in the novice group was statistically significant, longer path length was positively correlated to higher error percentage. (B) The correlation between idle time and error in the experienced group was statistically significant, decrease in the idle time was positively correlated to lower error. (C) The correlation between path length and total time in both groups was statistically significant, longer path length was positively correlated to longer time to complete the task. (D) The correlation between idle time and total time in both groups was statistically significant, longer total time to complete the task was positively correlated to longer idle time.

4.6 Discussion

The aim of this study was to identify differences between experienced and novice surgeons in the planning and efficiency of sensorimotor execution. We found significant differences between experienced individuals and novices in error percentage and time variables, with the expert group able to make faster and appropriate sensorimotor selection and execution. This by itself was not a surprising finding and merely replicates a large body of work on measuring differences in motor performance as a function of skill [see (Steinberg *et al.*, 2007; Ben-Gal *et al.*, 2011, 2013; Suebnukarn *et al.*, 2014; Mirghani *et al.*, 2016; Corrêa *et al.*, 2017)]. More interesting, however, was the observation of differences in idle time and path length of the hand movements.

We found that the experienced group had less idle time in comparison to novice participants and idle time positively correlated with error rates in the experienced group, but interestingly, there was no relationship with error in novice participants.

The weak correlation between idle time and error in the novice group is interesting for a number of reasons. First, it indicates a distinction between planning and error and thus, suggests that there may be value in using idle time as an adjunct performance measure. Second, it suggests that even though novice participants take longer in motor planning, they are unable to overcome a lack of skill necessary to effectively execute this task. These results are consistent with Chambers and Geissberger (Chambers *et al.*, 1997) who reported that, when beginners and competent students performed Class II cavity preparations, both groups achieved similar clinically acceptable preparations. However, beginners spent significantly more time in activities with no direct function to the cavity preparation. We propose that understanding what participants are doing while not moving their hands might be as important as the activities they undertake while drilling (D’Angelo, Rutherford, Ray, Laufer, *et al.*, 2015).

In contrast to idle time, path length has been used extensively in capturing surgical performance (Allen *et al.*, 2010; Chmarra *et al.*, 2010), but notably this measure is rarely reported in the dental literature (Suebnukarn *et al.*, 2014; Towers *et al.*, 2019). We found that the experienced group had smaller path lengths upon task completion relative to novices. Interestingly, we found that both groups showed a positive correlation between path length and error rates, however, this relationship reached statistical significance only in the novice group.

Our results are consistent with previous research showing that the total path length parameter is a reliable means of discriminating between different levels of experience (Suebnukarn *et al.*, 2014). There are a number of potential factors that may be driving these observed differences. One possibility is that the novice performer has more informational uncertainty that needs to be resolved and thus shows more exploratory behaviour, whereas the expert, through prior experience, is able to rapidly hone in on the most relevant sources of information and thus arrive at a solution much more rapidly and thus exert smoother control (Wolpert, Diedrichsen and Flanagan, 2011; Krakauer *et al.*, 2019).

Given that these measures of idle time and path length are (i) under-utilised in the examination of dental performance; and (ii) cannot be fully explained by the two widely used metrics of error and time, we suggest that the inclusion of these measures as adjuncts to time and error could be useful in undergraduate education programmes for accurate feedback and competency evaluations. We also note that due to duty hour restrictions, combined with the increasing complexity of clinical cases, surgeons are required to develop ever-more specialised skills in increasingly shorter amounts of time (Lewis and Klingensmith, 2012). An insight into the planning strategies being employed by a trainee, coupled with a measure of their motor efficiency, could help mentors better tailor feedback to help optimise learning for an individual.

4.7 Conclusion

We examine behavioural differences between novice and experienced performers in idle time and path length. The data show that expert performers produce movements that have superior economy and shorter planning times and that these performances measures are useful adjuncts to the oft-reported metrics of time and error. We propose that the addition of these measures in the dental education training program could be informative for mentors in providing more specific guidance to trainees to optimise the learning process.

Chapter 5 The Generalisability of Sensorimotor Skill in Healthcare Speciality

5.1 Abstract

Aim: One important aspect of a learned sensorimotor behaviour is generalisability of this skill to another related but unlearned behaviour. Sensorimotor theories play a fundamental role in the understanding of skill learning and the underlying flexibility and adaptability of the sensorimotor system. We aimed to examine the transfer of sensorimotor skills in healthcare specialty by exploring the performance of dentists and laparoscopic surgeons on simulated surgical and dental tasks and comparing them against surgically naïve, psychology group.

Methods: We used both a high-fidelity virtual reality dental simulator and a validated box laparoscopic simulator to record the performance of all participants to assess the transferability of skills between surgeons. Nineteen qualified dentists, surgeons and psychologists performed a dental drilling and a surgical thread transfer task.

Results: Both surgeons' and dentists' performances were superior to other groups in the domains that fall within their specialty and superior to controls in the domains that fall outside of their specialty.

Conclusion: The data confirm that there is a degree of generalisation - as both surgeons and dentists perform better at the opposite task than the psychologists. This transfer of highly skilled motor learning between healthcare specialty confirmed the importance of sensorimotor learning theories in understanding of the transfer of skill between specialty. Nevertheless, the two

specialised task-simulators detected performance differences across the three groups, suggesting that simulators could enable the identification of core skills that underpin surgical performance regardless of specialty. The identification of such core skills could improve the assessment of prospective surgeons, and lead to improved training provision prior to specialisation.

5.2 Introduction

If you were stuck between choosing a dentist and a psychologist to perform an emergency appendix removal, which one would you opt for? Similarly, for a tooth extraction, with no dentist available, would you call upon your social scientist friend or a surgeon specialising in Hepato-Pancreato-Biliary procedures? Whilst these far-fetched scenarios are unlikely to come into play in a health service near you (though a desperate shortage in dentists in the UK makes the latter an increasingly likely proposition), answers to these questions reveal much about our intuitions on the generalisability of skill.

From a surface level, it is clear that, although the subject specific knowledge required to perform surgery and dentistry are clearly unique, there seems to be an overlap in the sensorimotor demands of these professions. Both specialties require an integration of high-level medical knowledge to make appropriate evidence and experience-based decisions. But competency is also dependent on the sensorimotor skill that allows one to successfully execute these decisions. Whilst these sensorimotor abilities are trained in very different contexts, the extent to which these skills can be extrapolated to novel situations could provide an insight into the processes that underlie motor learning (Adams, 1987; Krakauer *et al.*, 2006).

Research in psychology has revealed that humans are able to rapidly learn tasks where the control parameters require simple input-output adjustments by some variable amount (e.g. adaptation to different degrees of sensorimotor rotation (Braun *et al.*, 2009)). In some scenarios we display a remarkable ability to extract general rules regarding tasks with a similar structure, by identifying invariants between different input–output mappings and can apply these rules to novel situations- a process referred to as motor structure learning (Braun *et al.*, 2009; Braun, Mehring and Wolpert, 2010; Pacheco and Newell, 2018). Indeed, structural learning is clearly manifest in a plethora of everyday activities. Perhaps the most popularly cited example of structural learning in action comes from Wolpert's bicycle analogy: 'When someone used to ride only one type of bike for example, racing bike and then asked to ride a mountain bike that will require changing in the control parameter from a racing bike to a mountain bike. Moreover, if the two bike models are close to each other in parameter space, then learning can be fast. Clearly, there is another possible way to speed up learning. If we have ridden many different types of bicycles, we might have extracted general rules for how the control parameters covary for different bicycles. Thus, when we are presented with a new task on the same structure, the search is restricted to a subspace of the full parameter space (e.g., the control subspace for the class of all bikes), thereby speeding up learning (Braun *et al.*, 2009; Braun, Mehring and Wolpert, 2010; Turnham, Braun and Wolpert, 2011).

However, highly skilled motor learning is also marked by a specialisation of function- with sensory motor commands developed and refined through interaction with, and tailored for, specific environments (Wolpert and Flanagan, 2010; Krakauer *et al.*, 2019). If a small change in the context is associated with a large alteration in the learning task, then generalising from prior leaning could also interfere with the new task and impair performance (Krakauer *et al.*, 2006). Here, we explore the extent of this transfer-interference phenomenon in the context of surgical performance. We asked nineteen qualified dentists, surgeons and psychologists (serving as a control group) to perform a series of simulated dental and surgical tasks on a validated box laparoscopic surgical simulator, and a high-fidelity virtual reality dental simulator.

5.3 Materials and Methods

5.3.1 Participants

The participants were clinical postgraduate dental students (N = 19, Male= 12, Female= 7, with an average of 4.7 ± 1.27 years of experience) enrolled from School of Dentistry at the University of Leeds, laparoscopic surgeons (N = 19 Male= 11, Female= 8 with an average of 5.05 ± 5.79 years of experience) enrolled from laparoscopy training course at the University of Leeds, and postgraduate psychology students with no experience in surgery and dentistry, serving as a control group (N = 19 Male= 8, Female= 11) enrolled from the School of Psychology at the University of Leeds. All participants had no previous experience of using these simulators. The study followed the tenets of the Declaration of Helsinki and was approved by the local Research Ethics Committee at the School of Dentistry and School of Psychology, University of Leeds, United Kingdom (Reference number: 17-0166; date approved: 30-May-2017).

5.3.2 Experimental Protocol

We used validated high-fidelity virtual reality simulator, The Simodont Haptic VR Dental Trainer® and a validated box laparoscopic simulator EoSim® described below to assess performance in each domain. For assessing the dental performance, a virtual cross geometric shape was employed in this experiment. A schematic example of the shapes is shown in (Figure 5-1). The task consisted of three zones: a target zone- which must be removed by the participant; Leeway zones (side and bottom) is adherently surrounding the target zone and the participants were instructed to avoid removing as possible; and the container zones (sides and bottom) represented by a block that surrounds the abstract shape and the participants were also told they must avoid during target removal. Furthermore, the participants were informed that the target removal percentage is 99%. For assessing surgical performance, a thread transfer task was employed to this experiment which involved inserting a thread in 5 pegs (holes) within a 10-minute using laparoscopic graspers on both hand (Figure 5-2).

5.3.3 EoSim®

The EoSim (developed by EoSurgical Ltd., Edinburgh, Scotland, United Kingdom) is a laparoscopic box trainer system that have been developed to train surgeons to acquired advanced skills such as suturing using laparoscopy. It is use innovative instrument tracking technology to measure the movement of the instruments as the surgeon perform each task. The tracking data is converted into performance metrics for each instrument. then it generates both kinematic and natural language feedback to help understand the metrics and highlight areas for improvement to refine the technique (Hennessey and Hewett, 2013; Retrosi *et al.*, 2015). It is consisting of a box and inside that box a high definition 1080p USB plug-and-play webcam built in with LED light strip built in to provide optimal illumination. From outside the box there is three keyholes opening for the laparoscopic instruments. The simulator connected to a computer by the USB plug and communicate through a software which contain the different practice tasks (Figure 5-2).

The kinematic performance measures provided by EoSim® are: time (second), distance (m): a measure of precision of control of the instrument in which expert are able to complete the task with lower instrument path distant, Speed (mm/s): the rate of movement of the instrument as the distance travelled in millimeter divided by the time in seconds, motion smoothness (mm/s^3): the continuality or non-intermittency of a movement, independent of its amplitude and duration. Intermittency in this context refers to movements that alternately decelerate and accelerate, and more intermittency corresponds to unsmooth movements (Hennessey and Hewett, 2013; Partridge *et al.*, 2014; Retrosi *et al.*, 2015).

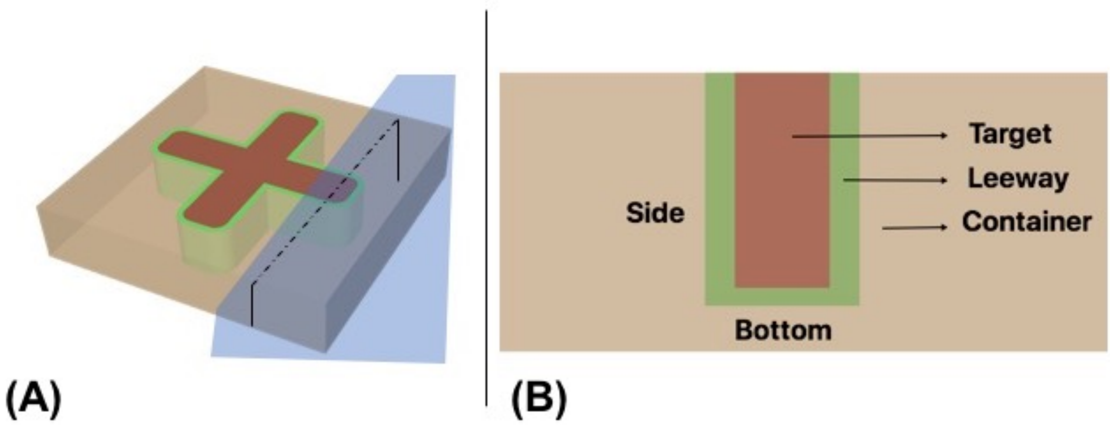


Figure 5-1 Schematic drawing of the cross-shape dental task (A), and (B) Cross-section showing the three layers of the dental task.

Figure 5-2 The set-up of the surgical simulator showing the performed surgical task which require passing the thread through the pegs.

5.4 Statistical Analysis

For performance assessment, due to the nature difference in the kinematic performance measures provided in both task , a z-score of a calculated composite measure that captured speed-accuracy trade-offs in performance for surgical (C_{surg}) and dental tasks (C_{dent}) is used to analyses the results. The dental composite measure was calculated by multiplying the total error by the time taken, such that lower scores indicate better performance (Equation 1). The surgical composite measure was calculated by multiplying the number of incomplete holes (n) plus one by the amount of time taken to complete the task within the maximum time period, such that lower scores indicate better performance (Equation 2).

$$C_{dent} = t_{dent} \times E_T \tag{1}$$

$$C_{surg} = t_{surg} \times (n + 1) \tag{2}$$

The variables were tested for normality to ensure the data met requirements for valid analysis of variance (ANOVA) by Q–Q plots and Shapiro–Wilk test ($P < 0.05$). A repeated measures ANOVA was conducted to compare the participants performance according to their group for the z-scored composite variables in dental and surgical simulators.

5.5 Results

A two by three analysis of variance conducted to compare the performance of participants on the task type (dental and surgical) according to the participants group (Surgeons, Dentists, and Controls) for the z-scored composite variables. This yielded a significant effect for the groups factor, [F (2, 54) = 18.23, p <0.001, $\eta_G^2 = 0.25$], and non-significant effect for the task type factor, [F (1, 54) = 0.00, p = 1.00, $\eta_G^2 = 0.00$]. However, the interaction effect was significant, F (2, 54) = 7.36, p < 0.001, $\eta_G^2 = 0.121$]. The descriptive analysis of each condition can be visualized in (Figure 5-3).

A series of one-way ANOVA then conducted to decompose the interaction by groups. For the controls' group, analysis of variance showed that the effect of task type was not significant, [F (1, 18) = 0.058, p =0.81, $\eta_G^2 = 0.0016$]. For the dentists' group, analysis of variance showed that the effect of task type was significant, [F (1, 18) = 24.52, p =0.0001, $\eta_G^2 = 0.348$]. For the surgeons' group, analysis of variance showed that the effect of task type was significant, [F (1, 18) = 9.05, p= 0.007, $\eta_G^2 = 0.189$].

Finally, we found that the performance of dentists and surgeons in the domains that fall within their specialty are not significantly difference, [F (1, 36) = 1.37, p= 0.24, $\eta_G^2 = 0.036$].

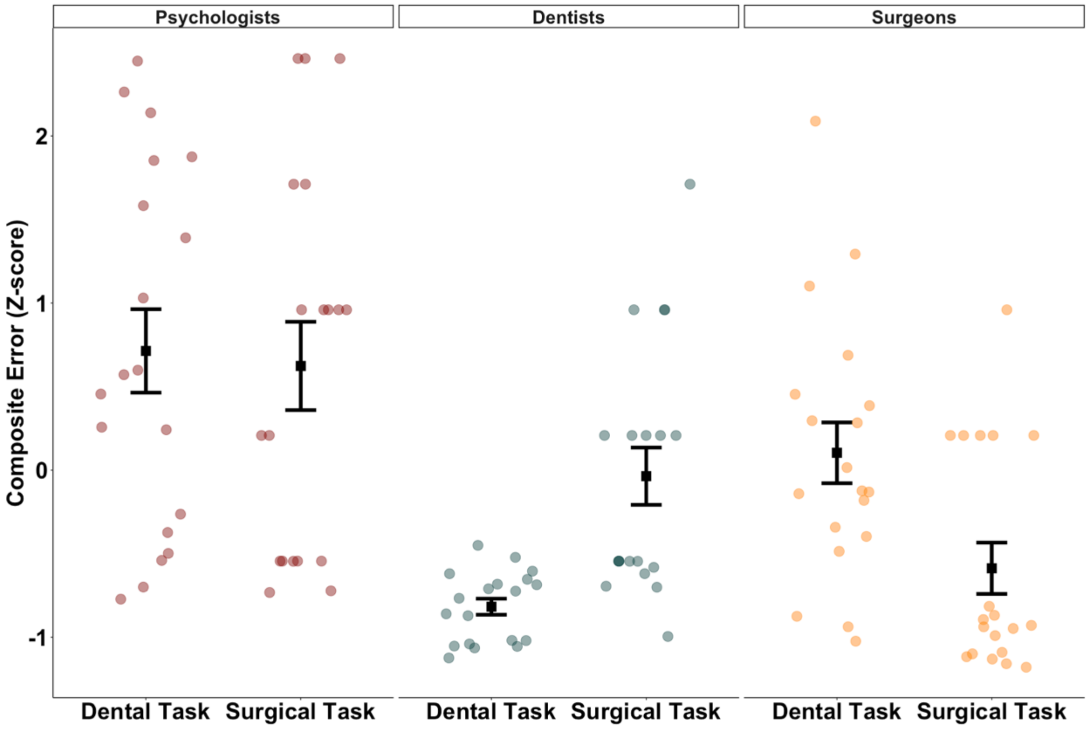


Figure 5-3 A comparison between the performance of each group in the dental and surgical tasks. The data is represented as a z-score of the composite error measures in which the lower the value the better the performance. The circles represent the participants performance, the squares are the mean and the error bars indicate the stander error. A statistically significant difference of composite error (negative indicates less composite error) was found across groups. Notably, the dentists' group on average performed better than controls group on the surgical task and performed better than controls and surgeons on the dental task. Similarly, the surgeons' group on average performed better than controls group on the dental task and performed better than controls and dentists on the surgical task.

5.6 Discussion

Surgical trainees face many challenges in their journey to learn the set of skills that are required for practice. Surgical simulation technology has the potential to support the learning of these skills and thereby improve the training and practise of surgeons. It is necessary to understand the acquisition of surgical skills in order to best use simulators and develop an evidence-based curriculum with more efficient training programmes.

In the present study, we identify evidence to support the intuitive notion that there are core surgical skills that transcend speciality (a conjecture that has support from the field of sensorimotor control and the phenomenon of structural learning). There would be great benefit in starting to understand the core skills that are common to all surgical specialties. First, this would allow better screening of prospective surgical trainees to identify individuals with neurological deficits that prevent the acquisition of surgical skills (estimates suggest that 5% of the general population have such deficits (Raw *et al.*, 2019)). Second, it would allow for more efficient training programmes within hospital and university settings. Third, it would improve our understanding of the fundamental processes that underpin the learning and practice of surgery. This would, in turn, allow us to take an evidence-based approach to accelerating the acquisition of surgical skills in trainees.

In sensorimotor research, speed and accuracy are two important aspects of evaluating performance but if these measurements analysed separately, sometimes lead to contradictory conclusions about the performance (Vandierendonck, 2017). To avoid such conflicts, a calculated measure that integrate speed and accuracy have been proposed, that collapses the duration in seconds and error together in a single kinematic measure with fast and accurate at one end of the continuum and slow and inaccurate at the other end. To make these composite measures of the dental and surgical task comparable, a z-score of each composite measures were calculated and then used as a performance metric for each task.

In the dental task, this score revealed that dentists' performance is better than controls and surgeons. In the surgical task, this score revealed that surgeons' performance is better than controls and dentists. Nevertheless, it shows that surgeons' performance is better than controls in the dental task and dentists' performance is better than controls in the surgical task. These data align well with the current understanding of some of the sensorimotor theories like, structural learning "Learning to learn" and hierarchical theories (selection and execution) which provide models for explaining how human can adapt and readapt to a sequence of similar tasks which lead to smoother and less jerky performance in the domains that fall outside of their specialty (Braun, Mehring and Wolpert, 2010; Gabriel, 2012; Diedrichsen and Kornysheva, 2015).

Certainly, a fundamental property of perceptual–sensorimotor skill is the precision and consistency of the spatial–temporal control of the arms. For example, in learning to drill into plastic teeth, the dental student must come to get their hand to the right place, at the right angle, and with the correct force. These factors can become highly precise and consistently repeatable (Wolpert, Diedrichsen and Flanagan, 2011; Rodger, Tang and White, 2016; Sadnicka *et al.*, 2017). Therefore, when practicing a new skill, even if we have been informed exactly how to do it, it is often required multiple practices to achieve proficient and fluid level of performance. The neural changes after practice allow us to accomplish a motor task better which observed as, increase in the speed of performance and reduction in the behavioral error (Adams, 1987;

Haith and Krakauer, 2018). However, our brain have the ability to do a decomposition to the structure of a learned skills leading to faster learning for problems sharing a similar structure (Seidler, 2004).

Here, we argue that this adaptation of surgeons (Dentists and Laparoscopic surgeons) to a new task without previous practice on them is recorded as reduction of systematic error induced by the perturbation, and this occurs through adjustment of an internal model that maps motor commands onto predicted sensory outcomes and result in smoother and faster execution of the movement with less amount of error (Shmuelof, Krakauer and Mazzoni, 2012). Explaining these data from a sensorimotor theory prospective give us a better understanding of what could be the difference between surgeons (Dentists and Laparoscopic surgeons) and controls in the performance of the tasks specially the task that fall outside of the surgeon's domain.

First, as discussed early in Chapter 1, the movements are generated through the interaction of different representational levels, ranging from movement goals (selection level) down to the specification of the actual muscle commands (execution level). The cognitive processing of task instructions occurs at the selection level then the most appropriate set of motor elements is mapped to task requirements at the execution level. In surgeons and dentists, the skill elements become encoded at an intermediate level within the dynamic neural network rather than at the selection level. As a result, the motor elements require little explicit or cognitive control to adapt to the new tasks (Sakai, Kitaguchi and Hikosaka, 2009; Diedrichsen and Kornysheva, 2015; Sadnicka *et al.*, 2017).

Second, this behavior could also be explained by structural learning theory which suggests that characteristic parameters are abstracted from a set of motor behaviours with the consequence that learning of similar tasks is facilitated. There are two distinct mechanisms by which the motor system might adjust its control parameters. First, fast learning could be a consequence of the similarity between the original and final settings of the control parameters. Second, as a result of structural learning, by extracting the common features and parameters of a learned sensorimotor skill and exploiting these parameters for efficient adaptation in novel tasks (Braun *et al.*, 2009; Braun, Mehring and Wolpert, 2010; Pacheco and Newell, 2018). For example, in adaptation to a new surgical task, dentists need to extract the common features and parameters of a learned skill (like suturing) and exploit that skill for efficient adaptation in the laparoscopic surgical task. Therefore, structural learning governs ‘learning to learn’ and transfer between tasks with the same task structure. (Krakauer, Ghez and Ghilardi, 2005; Krakauer *et al.*, 2006; Shmuelof, Krakauer and Mazzoni, 2012).

5.7 Conclusion

Our results demonstrate that both surgeons’ and dentists’ performances were superior in the domains that fall within their specialty and superior to controls in the domains that fall outside of their specialty. Furthermore, when surgeons and dentists perform a task in the domains that fall outside their specialty the motor control processes can extract the structure of the task and facilitate interference reduction due to their ability to execute a learned and fine sensorimotor skill. The data confirm that there is generalisation of skills between healthcare specialities in which surgeons and dentists had the ability to use their well-learned sensorimotor skill in a flexible manner to solve a new task.

Chapter 6 Frontal Theta Brain Activity Varies as a Function of Surgical Experience and Task Error

6.1 Abstract

Aim: Investigations into surgical expertise have almost exclusively focussed on overt behavioural characteristics with little consideration of the underlying neural processes. Recent advances in neuroimaging technologies e.g. scalp-recorded EEG, allow neural processes governing performance to be studied. We used EEG to examine whether surgical expertise and task performance could be differentiated according to an electrical brain activity signal known as frontal theta.

Methods: Behavioural and EEG data were acquired from dental surgery trainees with one (n = 25) and four years of experience (n = 20) while they performed low and high difficulty drilling tasks on a virtual reality surgical simulator. EEG power in the 4-7 Hz range in frontal electrodes (indexing frontal theta- an oscillatory brain activity signal and putative biomarker for cognitive control processes) was examined as a function of experience, task difficulty and error rate.

Results: Frontal theta activity was greater for novice participants (dental surgery trainees with one years of experience) relative to dental surgery trainees with four years of experience, expert participants (p = 0.001) but did not vary according to task difficulty (p = 0.15) and there was no experience X difficulty interaction (p = 0.87). Brain-behaviour correlations revealed a significant negative relationship between frontal theta and error in the

experienced group for the difficult task [r = -.594, p = .0058], but no such relationship emerged for novice participants.

Conclusion: We find frontal theta activity differentiates between surgical experience but correlates only with error rates for experienced surgeons whilst performing difficult tasks. These results provide a novel perspective on the relationship between expertise and surgical performance.

6.2 Introduction

Highly skilled surgeons have the ability to monitor and rapidly adapt to changes in the environment, appropriately tune into relevant information variables, select from a large repertoire of possible sensorimotor commands and execute with a smoothness that belies their many years of training (Hall, Ellis and Hamdorf, 2003; Sadideen *et al.*, 2013; Debarnot *et al.*, 2015; Krakauer *et al.*, 2019). Whilst the majority of research on surgical performance has examined the overt behavioural characteristics of such expertise (e.g. time to task completion (Wierinck *et al.*, 2007)) and subjective measures of mental workload (largely examined through post-hoc surveys (Byrne, 2011)), investigations into the underlying cognitive mechanisms that mediate the ability to carry out the complex sequences of action selection and execution required for surgical practice are rare.

In cognitive neuroscience, the processes involved in goal-directed attention, outcome monitoring, executing motor commands and suppressing irrelevant motor responses are clustered under the label of “cognitive control” (also referred to as “executive function (Mushtaq, Bland and Schaefer, 2011)”). One putative neural correlate of cognitive control is a pattern of oscillatory brain activity known as “frontal theta”- a signal that can be observed on the scalp through electroencephalography (EEG) recordings and quantified by calculating signal power in the 4-7 Hz range (Rabbi *et al.*, 2009).

Frontal theta is considered to be critical in performance monitoring (Alexander and Brown, 2011; Van Driel, Ridderinkhof and Cohen, 2012) and core to error detection (Ridderinkhof *et al.*, 2004; Van Driel, Ridderinkhof and Cohen, 2012)- the key to triggering selection and prioritization of information processing (Haith and Krakauer, 2018; Zavala *et al.*, 2018) and subsequent action (Cavanagh *et al.*, 2010; Cavanagh and Shackman, 2015). The recruitment of these “top down” control processes is heightened in scenarios where automatic processes are insufficient for successful adaptation to the current environment (Mushtaq, Bland and Schaefer, 2011), with the prefrontal cortex responsible for engaging a broad network of systems involved in goal-directed actions (Norman and Shallice, 1986; Miller and Cohen, 2001; Baddeley, 2003; Braver, Gray and Burgess, 2007; Parcutilo and Luna, 2016).

Extant theories of skill acquisition often describe a shift from deliberate to automatic action selection and execution, with requisite reductions in working memory requirements during the performance of a highly practiced action (Bassett *et al.*, 2015). A recent unifying framework for theories of cognition and action- known as the “Free Energy Principle" proposes that the neocortex (involved in higher order functions, such as sensory perception, motor commanded and spatial reasoning) constantly makes inferences about the world and learns from experiences through the violation of its predictions (Friston, 2010). Viewed in this framework, frontal theta activity could serve as both a teaching signal for the system to learn that it needs to refine its prediction for the future and simultaneously, trigger the cognitive resources required to produce adaptive control (Cavanagh and Frank, 2014). A more accurate world model would require fewer behavioural adjustments and thus, a reduction in the need to recruit cognitive control.

Predicated on this theory and evidence from neuroscience, we examined whether frontal theta activity could be used to distinguish between experienced and novice dental surgery trainees on a simulated drilling task carried out on a high-fidelity virtual reality simulator. We predicted that, overall, novice participants would exhibit greater task-related theta activity, reflecting greater top-down engagement of cognitive control processes relative to their more experienced counterparts. Secondly, given that behavioural adaptation following prediction error is a hallmark of learning, we expected a relationship between performance errors and frontal theta activity.

6.3 Materials and Methods

6.3.1 Participants

The data used for this study were obtained from undergraduate students of School of Dentistry, University of Leeds. The participants were assigned into two groups (as per their level of expertise). Forty-five participants were recruited with 25 first year dental students (the novice group = 16 females and 9 males, Age = 20.32 ± 2.54 years) and 20 fourth year dental students (the experienced group ;17 female and 3 male, Age = 23.7 ±0.58 years). All participants were right-handed, provided informed consent and were fully debriefed. The study was approved by the local Ethics Committee (REF:271016/IM/216) at the School of Dentistry, University of Leeds.

6.3.2 Experimental Protocol

Participants performed two tasks on The Simodont Haptic VR Dental Trainer® with two levels of difficulty. Specifically, participants were asked to drill two shapes, with difficulty operationalised as a function of geometric complexity. The “low difficulty” task involved drilling a simple straight shape, whilst the “high difficulty” task required participants to drill out a cross shaped object.

Each shape comprised three regions: (i) target - which participants were instructed must be removed; (ii) leeway area - which surrounded the target region on the sides and bottom (participants were instructed to avoid removing as much of this as possible); and (iii) a container area on the sides and bottom surrounding the leeway zone and represented as a brown coloured region that participants were instructed that they must avoid during drilling (see Figure 6-1, A).

The goal for participants was drill/cut 99% of the target region whilst minimising drilling in the leeway and avoiding the container regions as fast as they possibly could. To avoid potentially confounding order effects, we counterbalanced the presentation of shape across participants.

All data collection was carried out in the Dental Simulation suite at the University of Leeds. The total duration of the study for each participant was approximately 15-20 minutes. All Participants received a 5-10 minutes introduction to the simulator and the tasks to be performed. The EEG system was placed on participants' heads according to the manufacturer's instruction. Introduction to the simulator and the tasks to be performed took place after installation of the EEG headset and before the recording of any data (Figure 6-1). The EEG data were recorded continuously during the first dental task until the participants achieved the target performance level identified by the dental simulator. There was a two-minute break between the first and second task for all participants.

6.4 EEG Acquisition Device

EEG data were recorded using an Emotiv Epoc^{+} EEG wireless headset® (Emotiv Systems, Inc., San Francisco, CA). This system includes 14 active electrodes placed across the scalp according to the international 10-20 system (labelled AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1 and O2) with 2 reference electrodes placed on mastoids bone behind the ear (CMS-left mastoid)/driven right leg (DRL-right mastoid) ground (Figure 6-1, C). The signal from each electrode was converted to digital form via a 16-bit analogy to digital converter (ADC), with a sampling frequency of 128 Hz.

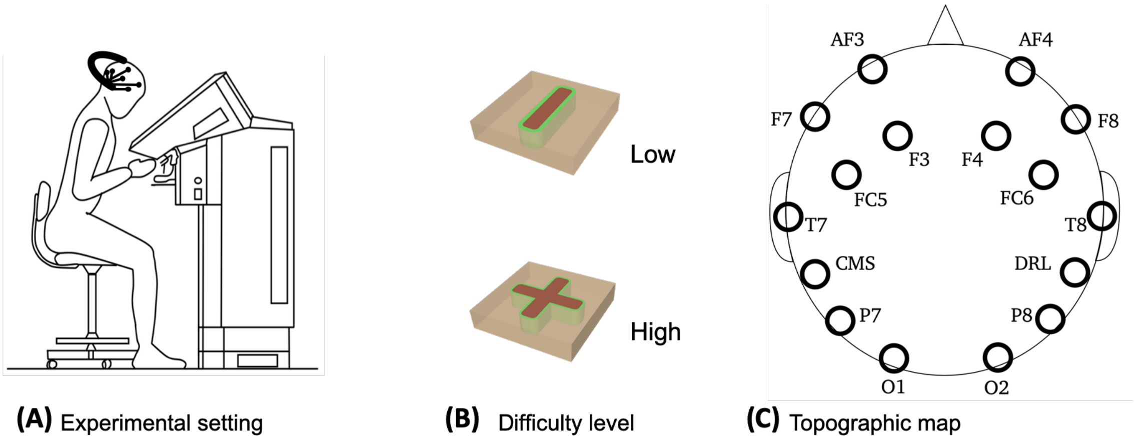


Figure 6-1 (A) Schematic drawing of the experimental setting (B) The drilling tasks: the straight shape task is defined as a low level of difficulty whilst the cross shape presents a high level of difficulty. (B) Location of the EEG electrodes relative to head position. Analysis focussed on channels F7, F8, AF3, AF4, F3, F4.

6.5 Data Analyses

6.5.1 EEG Data Analysis

As the EEG signals observed on the scalp are inherently noisy, we undertook a number of pre-processing steps before statistical analysis (Figure 6-2). Raw EEG data were pre-processed for artefact removal and band decomposition using EEG BESA® (Brain Electrical Source Analysis) (MEGIS Software GmbH, Gräfelfing, Germany). A linear Finite Impulse Response (FIR) filter was used to band pass the data between 1 and 20 Hz. Artefacts in the data were eliminated using an automatic artefact rejection routine implemented in BESA®.

Following this, we isolated theta band oscillations (4–7 Hz) from the channels in the frontocentral region of the scalp (FC; i.e. channels F7, F8, AF3, AF4, F3, F4) and band power was computed every quarter of a second using Welch's method, which estimates the power spectra based on a Fast Fourier Transform (FFT) (Welch, 1967; Shaker, 2007).

Finally, our measure of frontal theta was computed by averaging activity across the frontocentral region. This improved the signal-to-noise ratio by minimising the impact of any one single channel. One participant from the Novice Group had excessively noisy EEG data, with values more than three times the standard deviation of the mean in the high and low difficult tasks and was thus excluded from all analyses.

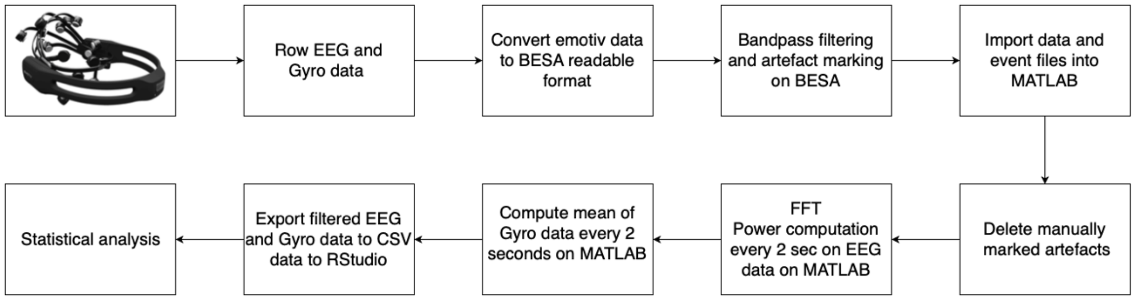


Figure 6-2 A flow chart for the pre-processing and filtering the EEG.

6.5.2 Statistical Analysis

For behaviour, we measured performance on the total error (quantified as percentages of drilling in the leeway regions). All measures were tested for normality to ensure the data met requirements for valid analysis of variance (ANOVA), by Q–Q plots and Shapiro-Wilk test. A 2 x 2 mixed ANOVA was conducted to compare performance across expertise (beginners vs. experts) X task difficulty (High vs. Low) for each dependent variable. Correlation analyses were used to examine the relationship between frontal theta and the amount of error in the behavioural data for both the low difficulty and high difficulty task for each group using a Pearson correlation. Comparisons of the magnitude of two correlations were performed where significant correlation were found using Hittner, May, and Silver's (2003) modification of Dunn and Clark's (1969) approach using a back-transformed average Fisher's Z procedure in cocor package in RStudio (Diedenhofen and Musch, 2015). The statistical significance threshold was set at $p < .05$ and we report generalised eta squared (η_G^2) as a measure of effect size and considered $\eta_G^2 = 0.02$ to be small, $\eta_G^2 = 0.13$ to be medium and $\eta_G^2 = 0.26$ to be a large effect size. All statistical analyses were performed using RStudio Version 1.1.463 (R Foundation for Statistical Computing., 2018).

6.6 Results

For behavioural performance (Figure 6-3, A), we found a significant effect of group [$F (1, 42) = 41.18$, $p < 0.001$, $\eta_G^2 = 0.41$]. As expected, (Mirghani et al., 2016), error rates were higher for novice participants (M = 16.51, 95% CI = [14.99, 18]) relative to the experienced participants (M = 9.68, 95% CI = [8.16, 11.2]). There was also a significant effect of task difficulty, [$F (1, 42) = 5.3$, $p = 0.03$, $\eta_G^2 = 0.04$], with more errors occurring in the high difficulty task (M = 13.9, 95% CI = [12.6, 15.1]) relative to the low difficulty variant (M = 12.3, 95% CI = [11, 13.6]). There was no interaction between task difficulty and group, [$F (1, 42) = 0.16$, $p = 0.69$, $\eta_G^2 = 0.001$].

For our EEG measure of cognitive control (Figure 6-3, B), we found a significant effect of group [$F (1, 42) = 12.05$, $p = 0.001$, $\eta_G^2 = 0.15$], with frontal theta activity greater for novice participants (M = 1.23, 95% CI = [1.02, 1.43]) relative to experienced participants (M = 0.72, 95% CI = [0.52, 0.93]), indicating an increase in the recruitment of cognitive control. However, there was no difference across tasks [$F (1, 42) = 2.11$, $p = 0.15$, $\eta_G^2 = 0.02$] and no interaction [$F (1, 42) = 0.03$, $p = 0.87$, $\eta_G^2 < 0.001$].

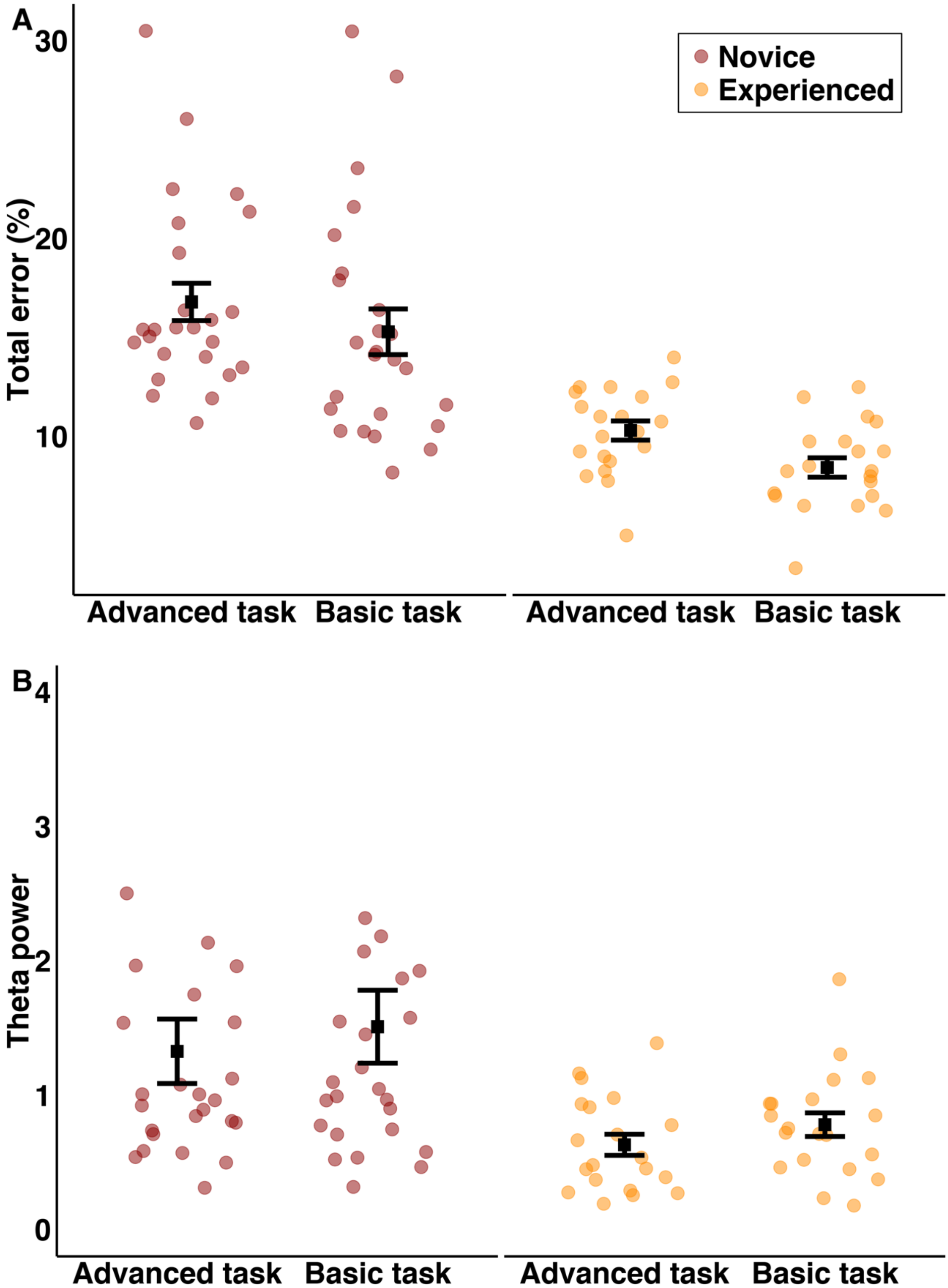


Figure 6-3 Group and task related differences in behaviour and neural activity. (A) Experienced participants made few errors relative to novice participants and participants made fewer errors in the low difficulty task relative to the high difficulty task. (B) On average, the experienced group showed lower theta activity relative to novice participants. There was no reliable difference in theta activity as a function of task difficulty.

Finally, we explored whether the amount of frontal theta activity exhibited by participants could be correlated with the amount of behavioural error within each of our groups and across the two tasks. We found a significant negative correlation between frontal theta activity and behavioural error rates in the experienced group whilst completing the high difficulty task (r = -.594, n = 20, p = .0058). In other words, smaller behavioural errors were associated with greater theta activity whilst higher error rates were correlated with lower theta activity in this high difficulty task for expert group. We found no other statistically significant relationships (r’s < 0.16; p’s > .46). To examine whether this observed relationship between neural activity and task error for our experienced group in the high difficulty task was significantly greater than the pattern found in the novice group (Figure 6-4), we compared the magnitude of the two correlations and confirmed that the patterns were reliably different from one another (z = 2.1779, p = 0.0294).

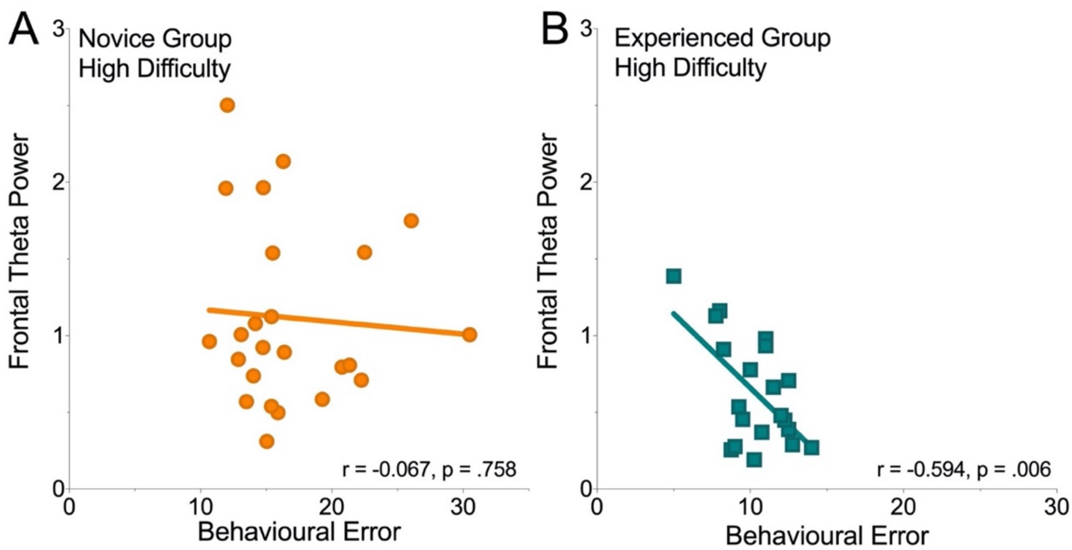


Figure 6-4 The relationship between behavioural performance and neural activity. Panel A shows no correlation between task error and frontal theta, but the experienced participants in the high difficulty task exhibited a strong negative correlation indicating that better performance was linked to greater theta activity (Panel B).

6.7 Discussion

Expert surgical performance is marked by seemingly effortless, flexible behaviour (Mylopoulos and Regehr, 2011) that typically manifests in smoother movements, shorter operating times and fewer errors (Hofstad *et al.*, 2013; Uemura *et al.*, 2014). This behaviour is the product of a distributed network of neural circuitry that takes a complex sequence of action selection and planned motor execution plan and refines over time to ensure smooth and seemingly automatic performance (Debarnot *et al.*, 2015; Diedrichsen and Kornysheva, 2015). However, there have been very few investigations into the neural processes linking brain and behaviour in the surgical domain (Bahrami *et al.*, 2014; Morris *et al.*, 2015; Lowe *et al.*, 2016; Kok *et al.*, 2018), with technological constraints limiting the ability to probe the neural underpinnings of surgical performance.

We took advantage of recent advances in wireless EEG technology to understand these processes in more detail. We reasoned that a neural signal, referred to as frontal theta, a putative marker of cognitive control would distinguish between experienced and novice dental students. We hypothesised that novice participants would require the recruitment of more cognitive resources to carry out the task relative to their more experienced counterparts and in line with this expectation, we found an increase in frontal theta activity for the novice.

We also found that the relationship between this signal and error correction was specific only to experts when performing difficult tasks. These results were not predicted a priori, and the result appears to indicate a reversal in the relationship between theta activity and performance and as such it is worth considering the nature of this relationship in detail. Here, in contrast to the global pattern in which theta activity was greater for the beginners relative to the experts, we found greater theta activity for experts who made fewer errors. To understand the processes underlying this phenomenon, consider the example of a learner driver stepping into the driving seat for the first time. We can imagine that our student is keen to learn and thus extremely attentive to the stimuli visible on the road ahead. However, it is also clear that there is a limit to the performance levels that could reasonably be expected of our student-irrespective of the amount of attentional resources that might be recruited for the task. In this context, making the driving conditions more hazardous is unlikely to modulate the relationship between cognitive control and performance to any reasonable degree, given that performance levels are low and attentional allocation is already high. Now contrast this with a more experienced driver, fewer attentional resources are required relative to the learner to exercise a higher level of performance. But if we heighten the task demands, say through poor weather conditions, we can reason that those who make fewer errors are also likely to be the individuals with increased allocation of attentional resources.

These results also speak to a more general issue of the relationship between expertise and performance. Expert surgeons are often, as in this study, considered a homogenous group and their performance benchmarked against trainee groups (Chandra *et al.*, 2010; Suebnukarn *et al.*, 2014). The present results indicate the existence of a more nuanced perspective on the relationship between experts and performance. Whilst on average, their performance may be better than trainees (on metrics relating to time and error), the performance of any one individual is likely to be modulated by a number of factors. Whilst the present data do not speak to causality, they do indicate a correlation between the amount of attentional allocation and performance in experienced participants and this may be a factor to consider in future comparisons- from experimental design (e.g. motivation and distraction) to measurement and statistical analysis (examining heterogeneity within expert groups).

Some limitations of the present work are worth noting: Whilst our sample size was comparable to the majority of previous research in this area, future research with larger sample sizes that have sufficient power to test the reliability of the brain-behaviour relationships identified here would be welcome. Our sample also comprised dental surgery trainees who differed in 3 years of experience. Exploring these relationships across different levels of experiences and specialities will be important in testing the generalisability of these findings.

Finally, given the increasingly lower costs of EEG technology and the ease in which these newer systems can be incorporated within surgical simulation, we suggest that the surgical education and performance community could benefit from integrating this measure into experiments in conjunction with behavioural measures of task performance.

6.8 Conclusion

In this study, we show that frontal theta, a putative marker of cognitive control, can differentiate between surgical experience. The data also indicate that frontal theta activity scales with the degree of behavioural adjustment following error commission only for expert surgeons performing difficult tasks. These results point towards a more nuanced interpretation of the relationship between expertise and performance- one that may be modulated by cognitive control.

Chapter 7 Discriminating Between Surgical Expertise by Manipulating Task Difficulty

7.1 Abstract

Aim: To be able to distinguish between different levels of skill requires that assessments are set at an appropriate level of difficulty. If the task is too hard the majority might struggle- a floor effect and if too easy, the majority of students might succeed- a ceiling effect. Here, we explore the impact of task difficulty as it relates to discriminating between different levels of dental training experience.

Methods: We recorded performance on a drilling task from undergraduate dental students (N=296) at the University of Leeds. We retrospectively collected data for an easy difficulty task (where participants were instructed to remove 60% of a target area; year 1 n = 40, year 3 n = 34, year 4 n = 37, and year 5 n = 39) and prospectively recorded data for a hard difficulty level (where the goal was to remove 96% of the target area; year 1 n = 61, year 3 n = 27, year 4 n = 27, and year 5 n = 30).

Results: We found a significant interaction between years of dental training and task difficulty [$F(3, 288) = 7.92$, $p < 0.0001$, $\eta_G^2 = 0.076$]. In the low difficulty condition, there was a smaller difference in comparison to the high difficulty condition.

Conclusion: The results demonstrate that manipulating difficulty level had an effect on distinguishing between years of dental expertise. These results highlight the importance of setting an optimal difficulty level for discriminatory assessment.

7.2 Introduction

Assessments are widely used in dental education to record the academic and clinical progress (Manogue, Brown and Foster, 2001; Shahriari-Rad, Cox and Woolford, 2017), sensorimotor skill development (Mirghani *et al.*, 2016), competency of the students (Evans, 2001) and ultimately determine whether individuals are ready to be independent dental practitioners (Patel *et al.*, 2018). As such, assessments must be carefully designed to allow for the effective evaluation of students against the learning outcomes that require the testing of relevant skills (Patel *et al.*, 2018).

When assessing whether students have acquired a specific preclinical dental skill (e.g. caries cavity preparation), a valid and reliable assessment is needed to ensure that students who pass this assessment are "competent" and can perform the dental procedures accurately and safely, and that students who fail need further skill development to reach competency. To achieve this level of discriminatory power, the assessment needs to be set at an optimal level of difficulty. If we sit the difficulty level of the task too high, the students might struggle in performing the task resulting in difficulty in capturing the different between them (often referred to as a floor effect (Lewis-Beck, Bryman and Futing Liao, 2004)). However, if the task is too easy, then we may suffer from the conceptually opposite problem (known as a ceiling effect), with all students able to pass the assessment, irrespective of their ability.

Preclinical dental skills are usually assessed using traditional phantom head simulators with typodont cast (Fugill, 2013). While there are standard predefined criteria on what might be an acceptable level of performance, this relies heavily on subjective evaluation through visual inspection (Manogue, Brown and Foster, 2001; Taylor, Grey and Satterthwaite, 2013; Huth *et al.*, 2017). Also, tailoring task difficulty according to skill level of the group is critical, time consuming and cumbersome process (e.g. study shows that the quality of root canal fillings performed by undergraduate students was adversely affected by case difficulty if they have less than two year of clinical experience (Alsulaimani *et al.*, 2015)).

In recent years, there has been a proliferation of haptic VR simulation in dental schools (Hollis, Darnell and Hottel, 2011; Cook *et al.*, 2015) to complement traditional training approaches (Quinn *et al.*, 2003; Urbankova and Engebretson, 2011; Shahriari-Rad, Cox and Woolford, 2017; Towers *et al.*, 2019). An oft-cited advantage of VR simulators is their ability to provide objective automated, standardised feedback (Perry *et al.*, 2015). A second, less cited, but potentially equally important feature is that one can automate task difficulty in a manner that is not possible with traditional teaching methods. However, the ease belies a potential pitfall- if implemented poorly, this can catastrophic for the validity of an assessment (manifesting in floor or ceiling effects).

To explore how manipulating the level of task difficulty can impact on our ability to be able to discriminate between different levels of performance, we asked undergraduate dental students with different level of expertise to perform a simulated dental task on a high-fidelity virtual reality dental simulator under two different levels of difficulty. Participants were exposed to a drilling task that required them to remove 96% of a target area (hard), or 60% of the target area (easy).

We expected to find an interaction between level of difficulty and level of expertise. Specifically, we expected the majority of students in the easy task would be able to perform the task, which would result in a ceiling effect and thus make it difficult to differentiate between different levels of expertise and discriminate between the level of skill within the same year group. But a harder difficulty level that is still obtainable, should result in the highest levels of discriminatory power within and between year groups.

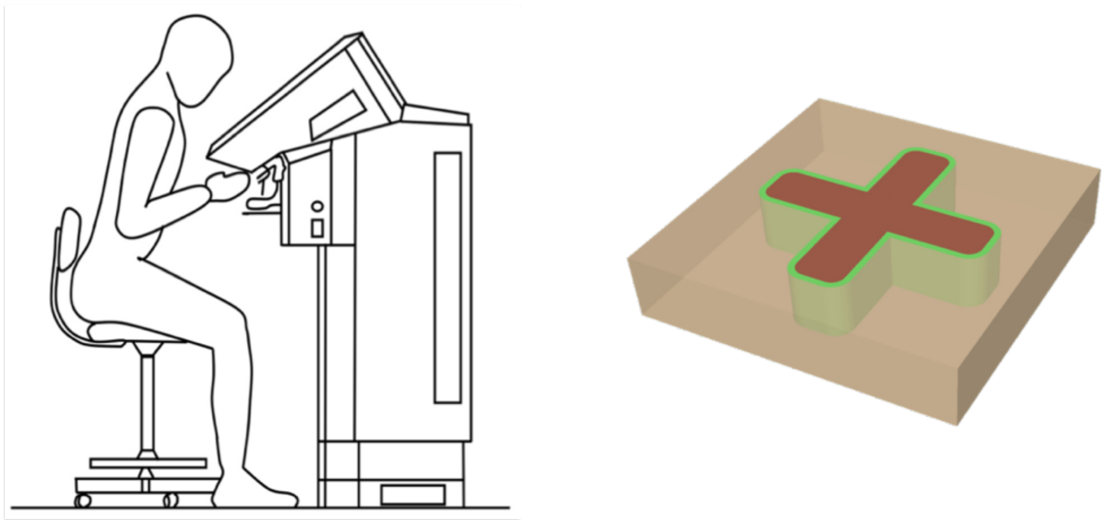
7.3 Materials and Methods

7.3.1 Participants

We recorded behavioural performance from undergraduate dental students (n=296) enrolled on the dentistry programme at the School of Dentistry at the University of Leeds. Data were prospectively recorded for the hard difficulty level, year 1 (n = 61), year 3 (n = 27), year 4 (n = 27), and year 5 (n= 30). For the easy difficulty level, the data were retrospectively extracted, year 1 (n = 40), year 3 (n= 34), year 4 (n = 37), and year 5 (n= 39). Participants provided informed consent and were fully debriefed. The study was approved (REF:271016/IM/216) by the ethics committees based in the School of Dentistry at the University of Leeds, United Kingdom.

7.3.2 Experimental Protocol

The students were required to complete a cross-shape task which involved the use of a dental handpiece to remove a target 'red zone' in the middle of a block, whilst attempting to minimise removal of leeway zones (the 'safe' outer areas of the block) as much as possible (Figure 7-1). Real-time feedback on performance was presented on a computer monitor attached to the device throughout the task. Participants were instructed that the aim of the task was to remove 96% in the hard difficulty level task, and 60% in the easy difficulty level task without touching the green and the beige zone as much as possible and as fast as possible.



(A) **(B)**

Figure 7-1 Experimental setting and task. (A) Schematic drawing of the VR dental simulator and the experimental setting. (B) the cross-shape task, illustrating the location of the different zones.

7.4 Statistical Analysis

For statistical analysis, we measured performance on two outcome variables, time (in seconds) and total error (as a percentage). The total error score was calculated as an average of leeway area removed from the side and bottom of the block. A 4 x 2 mixed ANOVA was conducted to compare performance across training year (1 vs. 3 vs. 4 vs. 5 years) X difficulty level (hard vs. easy) for each dependent variable (time and error). If an interaction was found, then a series of one-way ANOVAs were undertaken to decompose the interaction by difficulty level (hard vs. easy).

The statistical significance threshold was set at $p < .05$ and we report generalized eta squared (η_G^2) as a measure of effect size and considered $\eta_G^2 = 0.02$ to be small, $\eta_G^2 = 0.13$ to be medium and $\eta_G^2 = 0.26$ to be a large effect size (Olejnik and Algina, 2003; Bakeman, 2005). When a significant main effect was observed, Tukey's corrected post hoc comparisons were performed.

Finally, to further understand the relationship between motor performance and both difficulty level and training year we used the error and time scores and regressed this value against training year and difficulty level. This provided a model to predict the relationship between independent variables (time and error) and dependant variables (training year and difficulty level). All statistical analyses were performed using RStudio Version 1.1.463 (R Foundation for Statistical Computing., 2018).

7.5 Result

For error rate, we found a significant effect of training year [F (3, 288) = 18.20, $p < 0.0001$, $\eta_G^2 = 0.159$], with error rates significantly higher in year 1 (M = 21.3, 95% CI = [20.2, 22.5]) relative to year 3 (M = 17.4, 95% CI = [15.9, 18.9]), year 4 group (M = 16.4, 95% CI = [15, 17.9]), and year 5 (M = 15.4, 95% CI = [14, 16.8])

There was also a significant main effect of difficulty level [F (1, 288) = 1229.86, $p < 0.0001$, $\eta_G^2 = 0.810$], with more errors in the hard difficulty level task (M = 29.59, 95% CI = [28.61, 30.57]) relative to the easy difficulty (M = 5.68, 95% CI = [4.73, 6.62]) (Figure 7-2, A).

There was also an interaction between training year and difficulty level, [F (3, 288) = 7.92, $p < 0.0001$, $\eta_G^2 = 0.076$]. One-way ANOVAs were used to decompose the interaction by difficulty level.

For the easy difficulty level task, analysis of variance showed that the effect of training year was significant, [F (3, 147) = 4.89, $p = 0.002$, $\eta_G^2 = 0.091$], Tukey's post hoc indicate that year one group produced higher errors (M = 6.98, 95% CI = 6.13, 7.84) than only year 3 [(M = 4.93, 95% CI = 4, 5.85), p= .008], and year 5 [(M = 4.92, 95% CI = 4.06, 5.79), p=.005] with no other statistically significant results.

For the hard difficulty level, the effect of training year was significant and notably larger than the easy difficulty level, [F (3, 141) = 13.83, $p < 0.0001$, $\eta_G^2 = 0.227$], Tukey's post hoc indicate that year 1 group produced significantly higher errors (M = 35.7, 95% CI = 33.7, 37.7) than year 3 [(M = 29.9, 95% CI = 26.8, 32.9), p = 0.0097], year 4 [(M = 27, 95% CI = 24, 30), p < .0001], and year

5 [(M = 25.8, 95% CI = 22.9, 28.7, $p<.0001$] with no other statistically significant results.

For the time measurement, we found a significant effect of training year [$F_{(3, 228)} = 3.14$, $p = 0.03$, $\eta_G^2 = 0.03$], Tukey's post hoc indicate that year 1 group completing the task faster (M = 122, 95% CI = [109, 135]) relative to year 3 group [(M = 149, 95% CI = 133, 166),$p = .05$], year 4 group [(M = 149, 95% CI = 132, 165), $p = .05$], and year 5 group [(M = 139, 95% CI = 123, 154), $p = .36$].

We also observed a significant effect of difficulty level , [$F_{(1, 288)} = 589.14$, $p <0.0001$, $\eta_G^2 = 0.67$], with significantly longer time to complete the hard difficulty level task (M = 234.9, 95% CI = [223.6, 246.2]) relative to the easy (M = 44.3, 95% CI = [33.7, 54.8]), but no interaction between training year and difficulty level [$F_{(3, 228)} = 1.09$, $p = 0.35$, $\eta_G^2= 0.01$] (Figure 7-2, B).

Finally, multiple linear regression was carried out to investigate the relationship between training year, difficulty level and performance error. We found that both training year and difficulty level was a statistically significant predictor of performance [$F_{(2, 293)} = 682.3$, $p < 0.0001$, $R^2 = 0.823$]. For error, there was a 1.48% decrease in error for each extra year of training. For each extra percentage increase in the difficulty level, the performance on the error score increased by .73%. The adjusted R^2 value was 0.82 so 82% of the variation in performance error can be explained by the model: Error (y) = -35.28 - 1.48 * (Training year) + 0.728 * (Target area).

We also used the total time score and regressed this value against training year and difficulty level. We found that both training year and difficulty level was a statistically significant predictor of performance [$F_{(2, 293)} = 275.7$, $p < 0.0001$, $R^2 = 0.653$]. The regression analysis indicated that for every 1-unit increase in the training year, the performance on the total time score increased by 6.39 seconds and for every 1-percentage increase in the difficulty level, the performance on the total time score increased by 5.49 seconds. The adjusted R^2 value was 0.65 so 65% of the variation in time to complete the task can be explained by the model: Time (y) = - 320.58 + 6.39 * (Training year) + 5.49 * (Target area).

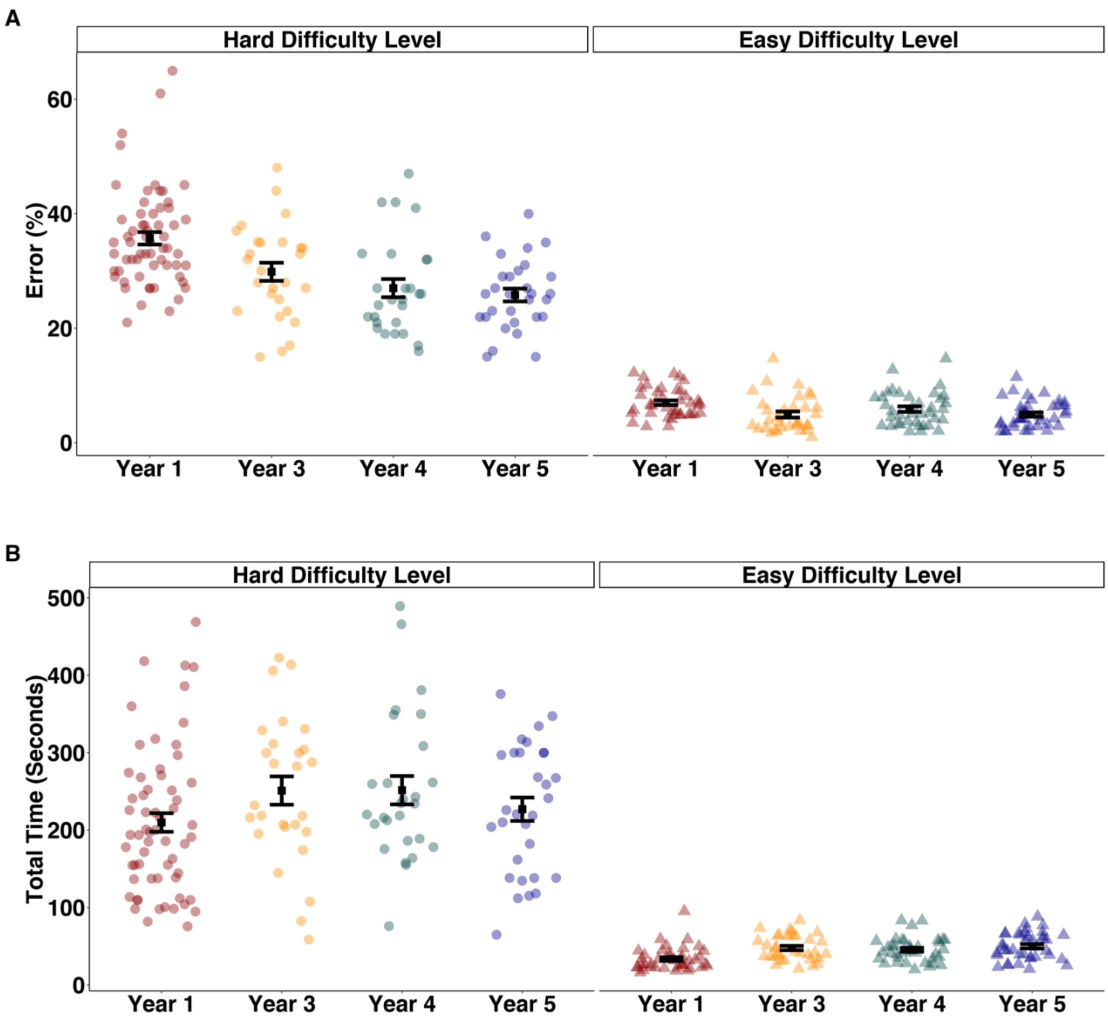


Figure 7-2 Performance outcomes across groups for hard and easy difficulty level. Black dots indicate mean and error bars indicate +/-1 S.E.M. Coloured dots indicate participants outcomes in the hard difficulty level task, coloured triangles indicate participants outcomes in the easy difficulty level task, (A) For the error measurement, both training year and difficulty level have a significant effect on the error score with significant interaction (B) For the time measurement, both training year and difficulty level have a significant effect on the time score.

7.6 Discussion

For preclinical operative assessment, preparation of typodont teeth is an almost universal competency assessment of students prior to allowing them to perform procedures on patient (Manogue, Brown and Foster, 2001; Taylor, Grey and Satterthwaite, 2013). Thus, there is possibly poor reliability and validity of such assessments. Until recently such skills have tended to be assessed in a subjective way by the assessors (Evans, 2001). As technology has advanced, more sophisticated simulators have been developed for training dental students. Haptic virtual reality simulators are an example of such advancement and these simulators are able to provide objective assessment of the students' performance. Moreover, task difficulty can, in comparison to phantom heads, be more readily adjusted.

In the present study we explored the effect of manipulating the level of difficulty on the ability to detect differences between different levels of skill. In order to address this issue, we recorded the performance of undergraduate dental students with varied level of expertise in virtual drilling task with two different difficulty levels.

We found an interaction between training year and difficulty level for error rates. When the task difficulty was set to a low threshold, all the year groups performed relatively well. On average the difference between year 5 group and year 1 was only 2% in error when completing the task. On average, the year 5 group produced less than 5% of error and year 1 produced less than 7% when completing the easy difficulty level task. In the hard difficulty level task, the differences between year groups become more apparent, with the average difference in error rates between year five and year one equal to ~10%. Even within the same year group, the hard difficulty task showed more discrimination than the easy one. For example, in the error variable, the standard deviation was smaller in the low difficulty task (year 1= ± 2.56, year 3 = ± 3.1, year 4 = ± 2.96, year 5 = ± 2.34) relative to the hard difficulty task (year 1= ± 8.46, year 3 = ± 8.23, year 4 = ± 8.27, year 5 = ± 6.14).

Together, these results point towards the importance of setting an optimum level of difficulty when assessing students on their learning and performance (Puryer and O'Sullivan, 2015). More broadly, they highlight the value of using VR simulation technology, given how readily task difficulty can be set by an instructor (either programmatically or by verbal instruction alone- as was the case with the present data), relative to phantom heads, where such task difficulty scaling is much more ambiguous (and assessment more subjective).

Looking forwards, having a benchmark to set case difficulty in preclinical dental simulation tasks would be invaluable in the process of creating valid and objective assessments. Such a benchmark would ensure students who pass their assessments can undertake the basic preclinical dental procedures safely and adequately, while at the same time highlighting students who need further practice (Taylor, Grey and Satterthwaite, 2013). However, we note that establishing a consensus on the appropriate level of difficulty, in view of the complexity in evaluating such an assessment, is not an easy task and would require a concerted effort amongst dental educators to reach agreement.

7.7 Conclusion

We aimed to test the influence of manipulating the task difficulty level on the performance and assessment of undergraduate dental students. The data show that, variations in the level of task difficulty may affect the usefulness of the task to capture the differences in performance. Our results show the importance of sitting an optimum goal difficulty level in order to have a valid and reliable assessment. The implications of this work relate to the evaluation of the use of VR dental simulators in the assessment of dental skills.

Chapter 8 More Than [a] Feeling: Using Haptics to Facilitate Motor Learning in a VR Surgical Simulator

8.1 Abstract

Introduction: Haptic virtual reality simulators are increasingly ubiquitous in the training of basic surgical motor skills, but the implementation of haptics seems to be limited only to delivering "realistic" representations of touch. We propose that haptics could also be used to manipulate a user's interactions with the environment by artificially increasing error information to accelerate motor learning. We tested whether this novel approach could benefit dental surgery training.

Methods: Sixty-one dental students performed tasks requiring either visual or haptic discrimination pre- and post-training on a haptic VR dentistry simulator. Participants were subsequently allocated to one of three training conditions: (i) an 'assistance' group- where participants were provided with haptic support that minimised error; (ii) a 'disruption' group - where participants trained with disruptive [forces that amplified error]; and (iii) a 'control' group where uniform "homogenous" haptic was provided without artificial haptic intervention. Learning (relative to baseline) was examined immediately after training and retention assessed 24 hours later.

Results: Learning was impaired in the assistance group relative to the disruption group ($p = .003$) and control group ($p = .02$) on the visual discrimination task, but there was no difference in retention rates ($P = 0.40$). On the haptic discrimination task, the assistance and disruption group outperformed the control group in assessments; learning was impaired in the control group

relative to the assistance group (p = .003), and retention was impaired in the control group relative to the disruptive group (p = .009).

Conclusion: We found the error amplification condition was best for learning in a visual discrimination task and that assistance and disruption were superior to the control group in the haptic discrimination task in both assessments of learning and retention. The results demonstrate that manipulating haptic information during training can improve learning and retention. More generally, the results show how haptics can be incorporated into VR simulators to provide more than a sense of touch- by actively manipulating user error to accelerate learning.

8.2 Introduction

Successful execution of the majority of dental procedures is only possible thanks to the remarkable sense of touch that humans are endowed with (Gardner, 2010). With a temporal resolution of 5 milliseconds and a fingertip spatial resolution that can be as low as 0.5 mm (Heller and Schiff, 1991), surgeons and dentists are able to use the rich data about the physical characteristics of objects contacted by the hand and instruments to perform millimetre precise procedures (Gallagher *et al.*, 2003; Perry *et al.*, 2015). It is the importance of this information source for clinical procedures that has led to the increasing implementation of haptic technologies in surgical simulation broadly (e.g. robotic surgery; (Culmer *et al.*, 2020)), and, more recently, dental training (Soumya and Ramachandra, 2011; Xia, Lopes and Restivo, 2013).

Haptic VR simulators are designed to present users with a ‘sense of touch’ and there is a growing body of evidence demonstrating that these simulators can be an effective means of training novice students on basic surgical tasks (Sturm *et al.*, 2008; Zendejas *et al.*, 2013; Al-Saud *et al.*, 2016). Indeed, recent work conducted by our group, and others, has shown that these simulators can display a degree of discriminant validity (being able to distinguish between different levels of real-world dental skill (Mirghani *et al.*, 2016; Al-Saud *et al.*, 2019) and predictive validity (i.e. performance on these measures indicates subsequent real-world clinical performance (Hung *et al.*, 2012; Al-Saud *et al.*, 2019).

The overarching design philosophy of these systems is that a more faithful rendition of the information available to the clinician in the real-world through simulation should have carry-over effects from simulation to the real-world (Cook *et al.*, 2011; Hamstra *et al.*, 2014). We propose that whilst this is indeed a critical part in engineering effective simulation that transfers to the real-world, it presents only a narrow perspective on the potential utility of haptics as it relates to the learning of skilled sensorimotor behaviours (Culmer *et al.*, 2020).

Research on sensorimotor learning through moving in forcefields generated by robotic haptic systems has shown that haptics can also be used to artificially manipulate interactions with the environment, and thus tailor the information available to the user to expedite the learning process (Emken and Reinkensmeyer, 2005; Reinkensmeyer and Patton, 2009).

Two popular approaches to using haptics to manipulate information involve using haptics to assist or disrupt task performance. Haptic assistance training involves providing artificial forces that guide a user to perform a task with little or no physical intervention, thus minimising performance error and providing a demonstration of the required movement (Cesqui *et al.*, 2008; Sigrist *et al.*, 2013). In contrast, disruption training involves the amplification of motor errors via force, either by perturbing movements spatially or temporally, or applying a force vector away from the desired trajectory.

The efficacy of these approaches for promoting skill learning is generally considered to vary as a function of expertise. For example, evidence indicates that individuals who are unable to perform the fundamental movements for a desired action (e.g. in the early stages of stroke recovery) benefit from assistance, but once a sufficient level of proficiency has been achieved constraining error during the practise of a motor task is ineffective in enhancing motor learning (Cesqui *et al.*, 2008; Sigrist *et al.*, 2013). Instead, augmenting error and active prospective control to minimise error is a key part of the learning process at this stage of recovery (Miall *et al.*, 1993; Wolpert, Ghahramani and Jordan, 1995; Wolpert and Kawato, 1998; Emken and

Reinkensmeyer, 2005; Reinkensmeyer and Patton, 2009; Van Beers, 2009; Wolpert and Flanagan, 2010; Marchal-Crespo *et al.*, 2014; Heuer and Lüttgen, 2015).

Predicated on this work, we propose that haptic engines in surgical simulators could do more than simply provide a representation of feeling as participants interact with the simulated environment. They could also be used to artificially increase and reduce error information to promote skill learning. This type of application is increasingly commonplace in motor rehabilitation (Wei *et al.*, 2005; Matsuoka, Brewer and Klatzky, 2007; Abdollahi *et al.*, 2011; Shirzad and Van Der Loos, 2012) but, to our knowledge, has not yet been applied to the training of dentists and surgeons.

In this study we set out to examine whether the provision of artificial haptic forces, above and beyond those used to represent objects in the environment, could support the learning of basis surgical drilling tasks. We compared performance change across three groups: one was provided with mild haptic assistance, which minimised training error, whilst another group experienced a disruptive force, which amplified error. We included a control group that carried out the same task but without any artificial haptic intervention.

We examined how these interventions influenced motor learning on two surgical tasks- one which required the ability to discriminate between different material densities in the absence of visual information and a second, which presented visual differences, but a homogenous density profile. We predicted that, for the visual discrimination task, the disruption group will be outperformed in the learning and retention assessments as a result of artificially increasing error information. Second, for the haptic discrimination task, we expected that learning and retention would be impaired in the control group relative to the disruption and assistance group as a result of practising in a homogenous density profile.

8.3 Materials and Methods:

8.3.1 Participants

Sixty-one first year undergraduate dental students were recruited to the study and randomly assigned into three groups. Twenty (12 females and 8 males, mean age = 19.1 years, SD = 2.31 years) were assigned as the ‘assistance group’ in which they experienced forces designed to constrain their movements to reduce error. Twenty (18 females and 2 males, mean age = 19 years, SD = 2.75 years) were assigned to a ‘disruption group’ in which they experienced forces that were designed to increase errors. Finally, twenty-one (20 females and 1 male, mean age = 22 years, SD = 5.8 years) were assigned to a control group in which they performed dental training tasks without any force intervention (a training as usual condition).

Three participants in the assistance group and two in the disruption group expressed a preference to use their left hand for the task, with all others used their right hand. All participants provided informed consent and were fully briefed. The study was approved by the local Ethics Committee at the School of Dentistry, University of Leeds (REF: 271016/IM/216).

8.3.2 Experimental Protocol

The experiment was conducted on two consecutive days at The Simodont Haptic VR Dental Trainer® skill Laboratory in the School of Dentistry, University of Leeds.

On the first day, participants were introduced to the dental simulator with a short overview, followed by a demonstration of the testing procedure. Each participant was allowed to try out the device as part of the introduction to familiarise themselves with the procedure and the required task. All the tasks were performed using a high-speed dental handpiece and diamond bur (FG 109010).

A ‘warm-up’ task consisted of a simple straight-line geometric shape with the aim of removing 93% of the target zone. All participants performed this task before starting the experiment to ensure a basic skill level was met.

The experiment on the first day lasted approximately 60-80 minutes in total to perform one attempt on the two assessment tasks (novel and distinct drilling tasks developed by the research team and uploaded to the Haptic VR Simulator, see Assessment Tasks section below) then to practise the training task (see training tasks section below) for 30 minutes, and then re-perform both assessment tasks.

We controlled for time rather than attempt to avoid confounding the results through trivial differences in training duration.

On the second day, participants were asked to once again perform the two assessment tasks to assess the 24-hour retention in learning. This design allowed for the assessment of not only the performance during training, but also the extent that the skills gained in training were transferred to the performance in the assessment tasks.

8.3.3 Assessment Tasks

Participant performance was assessed through two novel and distinct drilling tasks developed by the research team and uploaded to the Haptic VR Simulator (Figure 8-1). In one task, participants had to discriminate between different material densities in the absence of visual information to effectively drill through the object. We refer to this as the "haptic discrimination" task, because in the absence of visual information, participants could only rely on haptic cues to complete the task.

In a second task, participants were provided with visual information which indicated which areas of the object needed to be drilled, but there was no change in the haptic profile of the object i.e. it had a homogenous density profile. We refer to this as a "visual discrimination" task, as in the absence of any haptic information, participants could only use visual cues to successfully drill through the object.

All participants performed both tasks three times on both days. On day 1, before the training session, participants' performances on these tasks was recorded and used as the baseline performance measure and immediately after training, we recorded their post training scores to capture learning. Finally, on the second day, we asked participants to perform these tasks again to measure retention. We provide more details on the characteristics of these tasks in the following sections.

8.3.3.1 Haptic Discrimination Task

The haptic discrimination task consisted of two zones: a target zone, an amorphous region within the block with a low density (soft), and a container zone, surrounding the target with a high density (hard). Here, both zones were given the same colour (ensuring that task assesses a participant's ability to experience and act upon different levels of density they experience when drilling).

A single black dot was placed at the centre of the target area to act as a starting point. Participants were not provided with online feedback on the percentage of each zone removed and were instructed to drill only the soft structure without cutting into the hard zones, to drill as fast as they could and to stop when they believed they had removed all the soft area.

To avoid memorisation of the shape of the area, the haptic model was mirrored after training and participants were told the target shape would be different. The dimension of the haptic discrimination task block was 1.2 x 1.2 x 0.38 cm.

8.3.3.2 Visual Discrimination Task

The visual discrimination task consisted of three zones: (i) a target zone (coloured red) - which participants were instructed must be removed; (ii) leeway zones (coloured green) - which surrounded the target zone (on the sides and bottom) - participants were instructed to avoid removing this area as much as possible; and (iii) the container zones (coloured light brown) - a block that fully surrounded the cross that participants were told to avoid during target removal.

Importantly, this task had a uniform haptic profile (all three zones had the same density) to allow for the assessment of purely visuo-motor performance, i.e. control of the drill removing areas indicated by target colour alone.

Participants were given online feedback on the percentage of each zone removed and were instructed to remove the target area and avoid leeway and container zones as fast as they could and to stop when they removed 96% of the target area. The dimension of the visual discrimination task block was 1.2 x 1.2 x 0.26 cm.

8.3.4 Training Tasks

Three “doughnut” shaped tasks were developed for the simulator. The shape consisted of the target, leeway, and container zones as with the visual discrimination task. The rationale of choosing the different shape for the training tasks versus the assessment tasks was to ensure that the change in performance is not due to practice on the same shape, but a more general improvement in the ability of performing the tasks.

The models were designed to provide three different types of training (Assistance, Disruption and Control) and differed in their haptic profile, altering the forces on the drill as the participant attempted to remove the target zone. We describe the structure of these models in the following section.

8.3.4.1 Assistance

In the assistance model, the leeway and container zones were made of higher density material relative to the target. This had the effect of guiding the drill within the target area when moving closer to the edge, making mistakes potentially less impactful. It also served as a means of teaching the participant which zones should be removed by resisting movements away from the target.

8.3.4.2 Disruption

This model was conceptually opposite to the assistance condition. Here, the leeway and container zones were made with a lower density material relative to the target and the density of the target zone was modified with noise (generated from a 3D simplex noise function which includes local patches of high and low density). This had the effect of punishing mistakes and amplifying errors, as any movements close to the leeway zone would not be dampened with a force on the drill.

8.3.4.3 Control

The Control model set the leeway and container zone densities to the same as the target zone. This served as the experimental control condition (or training as usual) that could be used as a baseline for comparing performance changes against.

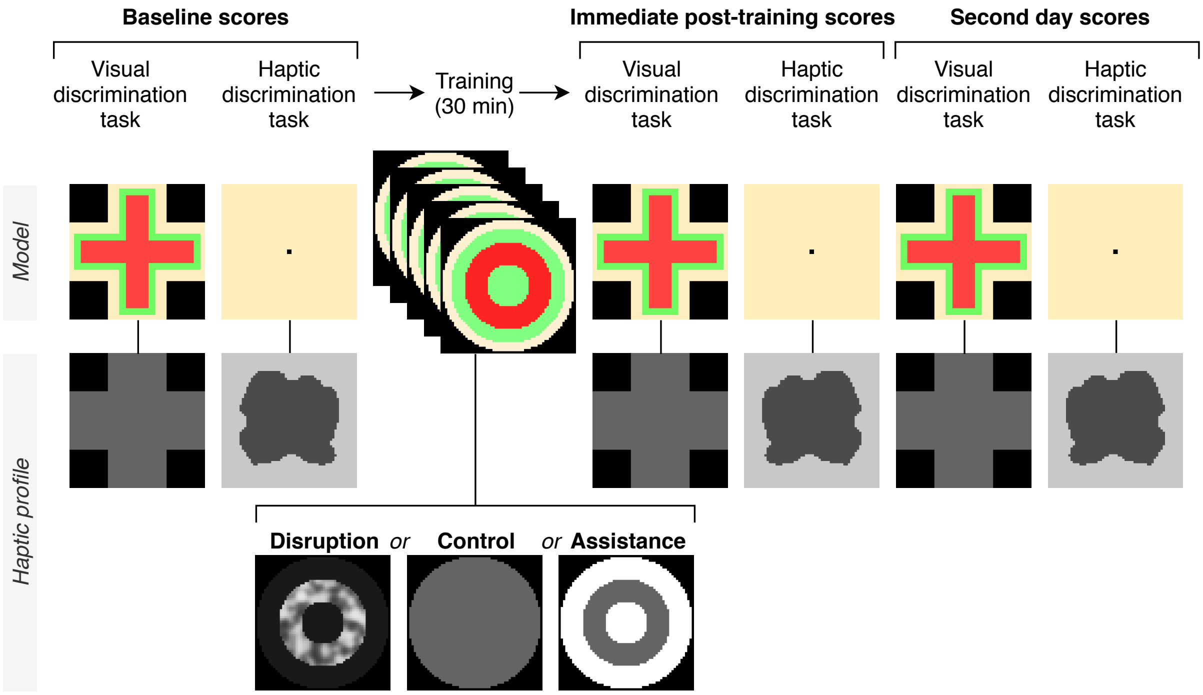


Figure 8-1 Experiment design. Participants performed two baseline assessment tasks (visual discrimination task and haptic discrimination task), 30 minutes of repeated training tasks (haptic profile assigned based on participant group assignment), then re-performed the same assessments (visual discrimination task and haptic discrimination task). Finally, both tasks were repeated on the second day to evaluate the retention. The haptic profiles are represented by images (bottom two thirds) with the brightness of the pixels indicating the density of the material in the model (black = softest, white = hardest).

8.4 Statistical Analysis

To capture performance in the visual discrimination task, we calculated "weighted error" as the percentage of leeway area removed plus two times the percentage of the container area removed, reasoning that there was a greater total volume of container, and the instruction given to completely avoid this area. In addition to this, we calculated time taken to reach 96% target area removed (in line with the instructed goal).

In the haptic discrimination task, the error was computed as the remaining percentage of target area to be removed. Finally, the time taken to complete the haptic assessment was used (the assessment ended when participants believed they had removed all the target area).

For the assessment of learning and retention we used analysis of covariance (ANCOVA; see (Vickers, 2001) for evidence of its superior statistical power in examining differences in learning c.f. ANOVA). Specifically, for learning we compared performance differences in the immediate post training scores and controlled for initial differences between groups by baseline scores as a covariate.

For retention, we compared the groups' performance differences on the second day scores and controlled for post-training differences between groups by using scores at the end of day 1 (the immediate post training scores) as a covariate. Tukey-corrected post hoc comparisons were performed where significant main effects were found.

8.5 Results

8.5.1 Learning Assessment

For the error measurement in the visual discrimination task (Figure 8-2, a), a one-way ANCOVA [between-subjects factor: group (control, assistance, disruptive); covariate: baseline scores], indicated that, after controlling for the baseline error, the training group showed a significant difference in the immediate post-training error [$F(2, 57) = 6.57$, $p = .002$, $\eta p^2 = .18$]. The assistance group performed significantly worse (Adj M = 50.9, SE = 2.96) than the disruption group (Adj M = 37.0, SE = 2.98), p = .005 and significantly worse than control group (Adj M = 38.6, SE = 2.89), p = .01 without significant difference between the control and disruption groups (p = .93).

For the time measurement in the visual discrimination task, group had a significant effect on the post-training time [$F(2, 57) = 5.15$, $p = .008$, $\eta p^2 = .15$]. The assistance group (Adj M = 110, SE = 15.2) performed the task significantly faster than the disruption group (Adj M = 176, SE = 15.2), p = .008 but this was not significantly different to the control group (Adj M = 158, SE = 14.8), p = .06) with non-significant differences between control and disruption groups p = .67 (Figure 8-2, b).

For the error measurement in the haptic discrimination task (Figure 8-2, c), the ANCOVA revealed a main effect of group [$F(2, 57) = 5.99$, $p = .004$, $\eta p^2 = .17$]. The assistance group performed significantly better (Adj M = 21.8, SE = 4.5) than the control group (Adj M = 43.6, SE = 4.4), p = .003 but non-significantly to disruption group (Adj M = 35.3, SE = 4.48), p = .09 with non-significant differences between the control and disruption group (p = .39).

For the time measurement in the haptic discrimination task (Figure 8-2d) the expected main effect of groups was non-significant [$F(2, 57) = 2.40$, $p = .1$, $\eta p^2 = .07$] (Figure 8-2, d).

8.5.2 Retention Assessment

For the retention assessment, we found a non-significant effect of the group on the visual discrimination task measurements for error [$F(2,57) = .92$, $P = .40$, $\eta p^2 = .03$] (Figure 8-2a) and time [$F(2,57) = .87$, $P = .42$, $\eta p^2 = .03$] (Figure 8-2, b).

For the haptic discrimination task (Figure 8-2, c), there was a significant effect of group on error [$F(2, 57) = 4.63$, $p = .01$, $\eta p^2 = .14$]. The disruption group (Adj $M = 32.7$, $SE = 4.10$) performed significantly better than the control group (Adj $M = 41.4$, $SE = 4.20$), $p = .009$), but there was no significant difference between the disruption and assistance groups (Adj $M = 32$, $SE = 4.34$), $p = .34$). There was also no significant difference between the control and assistance groups ($p = .30$).

For time taken to complete the haptic discrimination task (Figure 8-2, d), there was no significant effect of group [$F(2,57) = 1.26$, $P = .29$, $\eta p^2 = .04$]

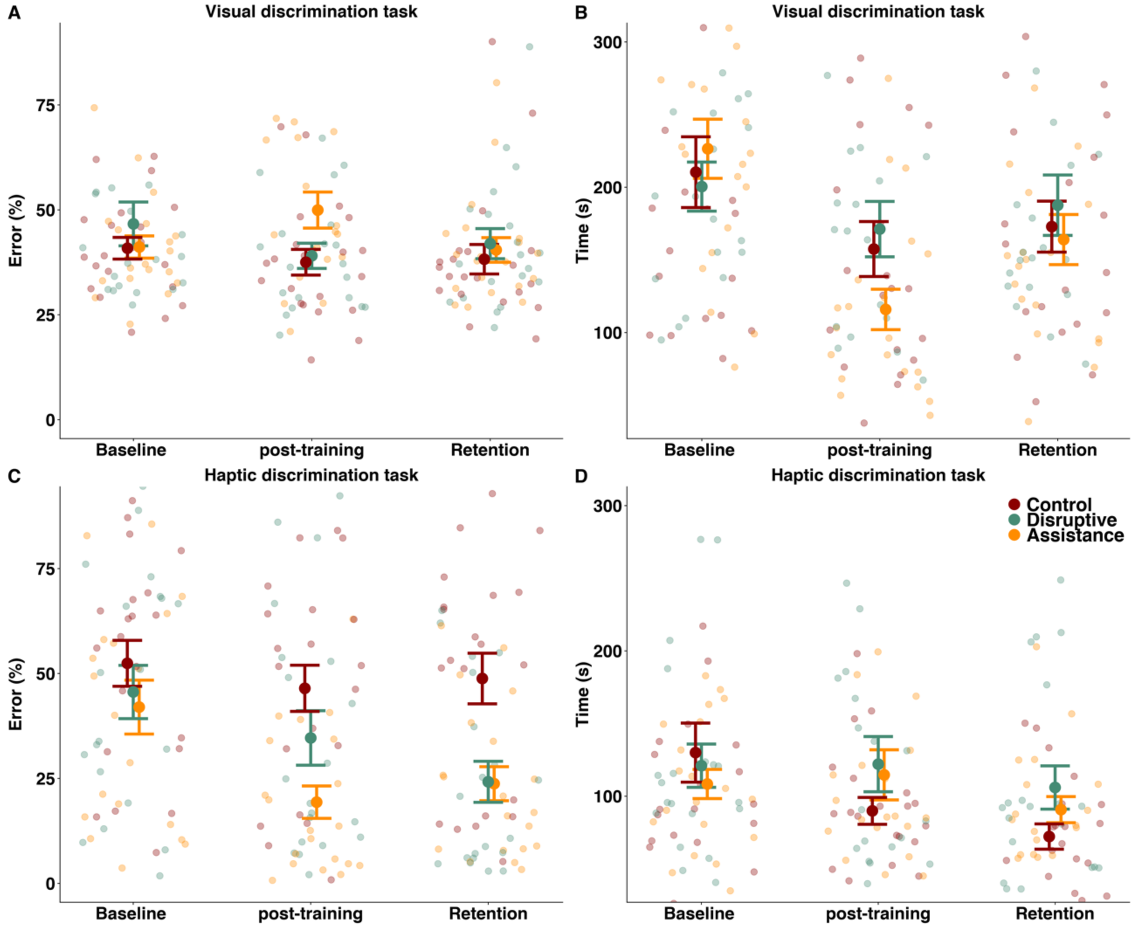


Figure 8-2 Performance outcomes across groups. Small dots indicate each participants performance, large dots indicate mean, and error bars indicate +/-1 S.E.M. (a) For the error measurement in the visual discrimination task, the assistance group on average increased their error on the learning assessment and interestingly decreased on the retention assessment. (b) For the time measurement in the visual discrimination task, the assistance group reduced the amount of time taken to complete the task on the learning assessment by over 100 seconds on average and significantly increased on the retention assessment. (c) For the error change in the haptic discrimination task, a decrease on average of error for all groups on the learning assessment and continue to decreased only in the disruption group on the retention assessment.(d) For both the learning and retention assessment, the amount of time taken to complete the haptic discrimination task was not found to be statistically significant.

8.6 Discussion

In recent years there has been a proliferation of haptic simulators in dentistry. These systems are primarily focused on replicating the 'feel' of performing procedures, but this does not necessarily translate to efficient training. Haptic technology potentially has substantial utility in promoting learning by directly manipulating movement (think of a physiotherapist guiding a patient's arm), but, thus far, no simulators have exploited this potential to accelerating learning. In existing systems, the tasks, quality and quantity of haptic feedback are typically generic and at most have graded levels of difficulty to be completed in sequential order on the basis of subjective (self or teacher) imposed timelines. Recent research suggests that learning processes can be accelerated through tailored delivery of haptic feedback. Specifically, evidence suggests that once a certain level of proficiency is obtained, augmenting the learning process through haptic assistance (i.e. forces that help minimise task-related error) hinders learning (Lee and Choi, 2010), whilst paradoxically, forces that push participants away from the goal and increase error, help acquisition (Emken and Reinkensmeyer, 2005; Reinkensmeyer and Patton, 2009).

Predicated on these ideas, we sought to examine the potential contributions of manipulating task error through haptic assistance and disruption on learning in undergraduate dental trainees. We asked participants to complete two novel dental drilling tasks – one requiring haptic discrimination, with no visual cues and a second requiring visual discrimination, with no haptic information.

For our learning assessment, we found that in the visual discrimination task, assistance hindered participants’ ability to reduce the percentage of error relative to the control and disruption group. In fact, disruptive forces led to significantly better performance in comparison to the assistance forces. These findings are consistent with previous results suggesting that disruptive forces might be beneficial for motor learning in comparison with assistance (Emken and Reinkensmeyer, 2005; Huang and Shadmehr, 2007; Cesqui *et al.*, 2008; Lee and Choi, 2010). This pattern may be explained by considering that during training, in order to develop a reliable and adaptable action, the learner needs to be able to explore the possible space of actions and outcomes. Making errors during practice allows the learner to experience a broad range of perceptual motor variables which eventually lead to successful action (Rodger, Tang and White, 2016), which may have manifested in the increased learning shown in this condition.

Notably, there was no significant difference between disruption and control groups in this task. A potential explanation for this lack of difference may come from the fact that the doughnut practice task for the control group and the visual discrimination task had uniform density.

In the haptic discrimination task, we found that participants who experienced more than one haptic profile “more than one level of density” had a better ability to distinguish between the hidden soft and hard structures. However, the assistance group was the only group that was able to significantly remove a higher percentage of the soft structure in the haptic task compared to the control group. Once again, we can consider the characteristics of the target condition as a potential explanation for lack of difference, but this time between the doughnut practice task for the assistance group and the haptic discrimination task.

With regards to retention, we found a non-significant effect of group on the visual discrimination task. We believe that this lack of difference may be driven by behavioural de-adaptation of the assistance group. Assistance haptic intervention seems to facilitate a strategy of moving fast and making more errors (Williams, Tremblay and Carnahan, 2016; Clamann and Kaber, 2018). But in our experiment, this seems to be only a short-term effect, as they de-adapt to become slower and more accurate in the retention performance change. Thus, it seems that any improvements at post-test may be accounted for, by an adjustment of strategy, i.e. moving more carefully or more recklessly, rather than an overall improvement in both speed and accuracy measurements. Here, it is also worth considering the clinical implications of this finding. The pre-clinical task of training students to remove caries safely and effectively is often difficult. Our data indicate that for a task that requires high precision, with a material that is generally uniform in density (e.g. crown preparation), 30 minutes of training may be insufficient to deliver measurable differences to drilling accuracy when only visual cues are available.

In the haptic task, the control group struggled in differentiating between the soft and hard structures in comparison to the assistance and disruption groups. Interestingly, this time, it was only the disruption group that was able to significantly remove a higher percentage of the soft structure in the haptic task compared to control group during the retention test, thus suggesting that this condition promoted consolidation of learning after 24 hours.

Work on examining whether the haptic technology present in these systems can be used to manipulate movement by providing assistance and/or disruptive forces to accelerate motor skill learning is still in its infancy. The mechanisms that contribute to the learning benefits facilitated by haptic interventions are not fully understood and more fundamental motor learning experiments are required. Future research is also required to address the feasibility of integrating multimodal simulation to examine the effect of combined the assistance and the disruptive model in skill acquisition.

8.7 Conclusion

In summary, these data indicate that the acquisition and retention of basic dental motor skills in novice trainees varies depending on the training method. In a visual discrimination task, haptic assistance on its own hinders the acquisition of skill in comparison to disruption and control groups. In a haptic discrimination task, training on uniform haptic density hinders the acquisition of skill in comparison to disruption and assistance groups. These results demonstrate the potential utility of using the haptic interface present in VR dental simulation to provide more than a sense of touch and use this feature to help accelerate learning.

Chapter 9 General Discussion and Conclusion

9.1 Thesis Context

Before we delve into the specific outcomes from this thesis, it is worth reminding ourselves of the context in which this work was undertaken and the overall aims and objectives of this project. Central to this thesis is the idea that a multi-modal approach is required to identify the fundamental differences between experts. We set out to examine this question because we reasoned that, in providing insights into the numerous different characterises of experts, we may be able to provide trainees and mentors with more tools in their armoury that can be used to help students transition towards expertise.

A second important theme in this thesis was the idea that simulation, and specifically Haptic VR simulation, could be an effective means of capturing potential individual differences and that these systems could also be used to support training and assessment (LeBlanc *et al.*, 2004; Issenberg *et al.*, 2005). While simulation in dentistry has been around for almost as long as the profession itself, the idea of using virtual reality and haptics is relatively new and only made possible in recent years thanks to advances in computing power and novel design engineering (Wang *et al.*, 2003; Kim *et al.*, 2005; Yau, Tsou and Tsai, 2006; Cao *et al.*, 2007; Rhienmora *et al.*, 2008; Konukseven *et al.*, 2010; Tse *et al.*, 2010; Yoshida *et al.*, 2011) along with experimental research demonstrating the validity of these systems (Imber *et al.*, 2003; Steinberg *et al.*, 2007; Ben-Gal *et al.*, 2011; Urbankova and Engebretson, 2011; Ben-Gal *et al.*, 2013; Shahriari-Rad, 2013; Suebnukarn *et al.*, 2014; Mirghani *et al.*, 2016, 2019; Corrêa *et al.*, 2017; Shahriari-Rad, Cox and Woolford, 2017; Al-Saud *et al.*, 2019).

Despite these important developments, there has been little work on how these devices can be used to (i) understand the development of expertise and (ii) use such knowledge to accelerate learning. Without such work (and then application into real world contexts), we might ask ourselves what the value of such endeavours are beyond traditional methods of learning. Do we want students to practice more often? Do we want them to practise alone? What makes these devices different to phantom heads? A particularly difficult question that will need to be navigated over the coming years for almost all dental schools across the globe (if they haven’t already tackled it) is whether such systems add sufficient value for dental schools to purchase and implement them in their curricula.

9.2 General Conclusions

Within the context presented above, the work in this thesis sought to identify the fundamental differences between experts and novices in sensorimotor skill at a behavioural and cognitive level. We recap some of the main findings from each of the experimental chapters next and summarise the main conclusions.

In Chapter 3, we found that learning over a considerable period of time appears to be quantitatively similar for an expert and novice. Interestingly, despite extended extensive practice (over 600 attempts across 7 months), both our expert and novice continued to make improvements, but error rate reductions tailed off. These data re-emphasise the point that acquiring dental drilling skills can take many training hours, but also highlight that expertise is a moving feast and improvements remain possible long after most research studies on the topic stop examining performance.

Building on these findings, we examined whether there were qualitative differences in the strategy employed by experienced and novice dentists. In Chapter 4, we found expert participants showed significantly less idle time and moved their hands shorter distances to complete drilling tasks relative to novices. These data indicate that experienced performers have shorter planning times and superior movement economy. We propose that these performance metrics could be valuable adjuncts to the commonly used measures of error and time in the dental education programme and be useful for mentors/trainers in providing tailored feedback to novice students to optimise their learning and performance.

In Chapter 5, we probed the nature of expertise in more detail by asking the extent to which expert skill in one domain can be generalised to another. We found a degree of generalisation- with surgeons performing better at dental tasks and dentists performing better at surgical tasks than a control group. This transfer of highly skilled motor learning between the two healthcare specialties is an important finding for understanding the aetiology of skill learning and highlights the existence of core skills that underpin surgical performance regardless of specialty. The identification of such core skills could improve the assessment of prospective surgeons and lead to improved training provision prior to specialisation.

In Chapter 6, we took advantage of recent advances in wireless EEG technology to examine neural differences as a function of expertise. We found that a putative marker of cognitive control, a signal known as frontal theta, could differentiate between skilled performance. We established that novice undergraduate dental students require the recruitment of more cognitive resources to carry out the same VR drilling task relative to more experienced dental students and this could be informative for trainers. Secondly, we found that individual differences in frontal theta scaled with performance errors in our expert group, which presents a novel explanation for heterogeneity in experts.

The aforementioned chapters relied extensively on haptic VR simulation and it is clear from this work that these technologies have the potential to add considerable value above and beyond traditional phantom head simulation. Thus, in the final two experimental chapters we sought to formalise these ideas by examining ways in which VR dental simulators could be used for assessment and training. In Chapter 7, we examined how artificially manipulating task difficulty (fundamentally difficult to employ with phantom heads) could impact on the performance and assessment of undergraduate dental students. The data show that variations in the level of difficulty had an effect on the discriminant validity of assessment and that the ability to tailor difficulty according to skill level could be valuable in supporting student learning.

Finally, In Chapter 8, we sought to examine whether we could take advantage of the haptic technology included within dental simulation systems to artificially increase error (haptic disruption) or reduce error (haptic assistance) information to accelerate motor learning. We found that error amplification through haptic disruption was better for learning in a visual discrimination task and that both disruption and assistance led to faster learning relative to a control group (who followed a standard training protocol). To our knowledge, this is the first study in the dental education literature that proposes a means of accelerating skill acquisition by manipulating the ways in which the learner interacts with a given task. This work provides a novel contribution to a growing body of evidence on the value of VR haptic dental simulators in dental education.

9.3 Limitations and Future Work

There are some limitations of the current thesis and numerous opportunities to build on the present work in future research that must be acknowledged.

First, it is important to note that the majority of participants in the studies reported in this thesis came from a single dental school in the UK and the remaining participants were recruited from students enrolled on other courses at the University of Leeds. Whilst this sample is likely typical of a UK student population, a key theme that has emerged from the work above (particularly strong from our work on tailoring assessment difficulty) is that tailoring sessions that examine training and performance to an individual's ability could yield more dividends than treating all novices (and indeed experts, as highlighted from our EEG analysis) as a homogenous group. This may be particularly important when groups vary more extensively e.g. in regions where selection for an undergraduate programme follows a less prescriptive approach.

Perhaps more important for the generalisability of the findings of this work is the equipment employed throughout the experimental work in this thesis. We used one specific haptic VR simulator for all training and assessment that has undergone various examinations to establish its construct and predictive validity. Whilst haptic simulators are becoming more commonplace, the extent to which these findings can be applied will be dependent on various factors including the technical capabilities of these systems and a careful analysis of the similarities of tasks and training protocols.

Finally, it is important to consider some of the experimental design choices made in this thesis. We note that the 5 of the six studies in this thesis employed a cross-sectional design whilst the 6th study used a longitudinal case study approach and that most often the "control" group in our experiments included participants with no experience of dental training. Whilst this was particularly useful for eliminating potential confounds such as previous experience and task knowledge that could have contaminated the results of our experimental manipulations, we were also bound by administrative restrictions and ethical considerations that precluded intervention randomized controlled trials with our target population. Certainly, providing training to students that may benefit or hinder them compared to others brings considerable ethical concerns and these issues emerge in all forms of educational research. The longer-term hope is that the work presented here can be strengthened with replications and demonstrations across different contexts to establish an evidence base that can inform the implementation of training that can benefit all students.

In closing, this thesis has borrowed extensively from experimental psychology and existing theories of sensorimotor learning and applied this to understanding the nature of expertise in the domain of dental training. We have shown some ways in which these ideas may be translated to accelerating learning. Coupled with simulation technology, we believe there is great promise to build on this work and incorporate principles from sensorimotor learning into dental education curriculums across the globe.

References

AAHE (2002) *Quality Assurance Agency for Higher Education. Subject Benchmarking Statements - Dentistry.*

Abdollahi, F. *et al.* (2011) 'Arm control recovery enhanced by error augmentation', in *IEEE International Conference on Rehabilitation Robotics*. doi: 10.1109/ICORR.2011.5975504.

Abe, M. *et al.* (2011) 'Reward improves long-term retention of a motor memory through induction of offline memory gains', *Current Biology*, (12;21(7):557-62). doi: 10.1016/j.cub.2011.02.030.

Adams, J. A. (1987) 'Historical Review and Appraisal of Research on the Learning, Retention, and Transfer of Human Motor Skills', *Psychological Bulletin*, (101(1):41–74). doi: 10.1037/0033-2909.101.1.41.

Al-Saud, L. *et al.* (2016) 'Feedback and motor skill acquisition using a haptic dental simulator', *European Journal of Dental Education*, (11(4):1–8). doi: 10.1111/eje.12214.

Al-Saud, L. *et al.* (2017) 'Drilling into the functional significance of stereopsis: the impact of stereoscopic information on surgical performance', *Ophthalmic and Physiological Optics*, 37(4). doi: 10.1111/opo.12393.

Al-Saud, L. *et al.* (2019) 'Early assessment with a virtual reality haptic simulator predicts performance in clinical practice.', *BMJ Simulation & Technology Enhanced Learning. ISSN 2056-6697 ,In Press.* doi: 10.1136/bmjstel-2018-000420.

Alexander, W. H. and Brown, J. W. (2011) 'Medial prefrontal cortex as an action-outcome predictor', *Nature Neuroscience*, (14(10):1338-44). doi: 10.1038/nn.2921.

Allen, B. *et al.* (2010) 'Support vector machines improve the accuracy of evaluation for the performance of laparoscopic training tasks', *Surgical Endoscopy*, (24(1):170–8). doi: 10.1007/s00464-009-0556-6.

Alsulaimani, R. S. *et al.* (2015) 'Effects of clinical training and case difficulty on the radiographic quality of root canal fillings performed by dental students in Saudi Arabia', *Iranian Endodontic Journal*, (10(4):268-73). doi: 10.7508/iej.2015.04.012.

Altbach, P. G., Reisberg, L. and Rumbley, L. E. (2009) *Trends in Global Higher Education : Tracking an Academic Revolution*, *UNESCO 2009 World Conference on Higher Education*. doi: 10.1016/j.bse.2004.04.006.

Anderson, J. R. (1982) 'Acquisition of cognitive skill', *Psychological Review*, (89(4), 369–406). doi: 10.1037/0033-295X.89.4.369.

Aune, T. K., Ingvaldsen, R. P. and Ettema, G. J. C. (2008) 'Effect of physical fatigue on motor control at different skill levels', *Perceptual and Motor Skills*, (106(2):371–86). doi: 10.2466/PMS.106.2.371-386.

Baddeley, A. (2003) 'Working memory: looking back and looking forward', *Nature Reviews Neuroscience*, (4(10):829-839). doi: 10.1038/nrn1201.

Bahrami, P. *et al.* (2014) ‘Neuroanatomical correlates of laparoscopic surgery training’, *Surgical Endoscopy*, (28(7):2189-98). doi: 10.1007/s00464-014-3452-7.

Bakeman, R. (2005) ‘Recommended effect size statistics for repeated measures designs’, *Behavior research methods*, (37(3):379-384). doi: 10.3758/BF03192707.

Bakker, D. *et al.* (2010) ‘Transfer of manual dexterity skills acquired in the Simodont, a dental haptic trainer with a virtual environment, to reality: a pilot study’, *Bio-Algorithms and Med-Systems*, (6(11):21–2).

Banks, G. P., Mikell, C. B. and McKhann, G. M. (2014) ‘Practice makes efficient: The low metabolic cost of expertise’, *Neurosurgery*, (74(4):N12–4). doi: 10.1227/01.neu.0000445334.50327.11.

Bassett, D. S. *et al.* (2015) ‘Learning-induced autonomy of sensorimotor systems’, *Nature Neuroscience*, (18(5):744-51). doi: 10.1038/nn.3993.

Bayne, S. (2015) ‘What’s the matter with “technology-enhanced learning”?’, *Learning, Media and Technology*, (40(1):5–20). doi: 10.1080/17439884.2014.915851.

Van Beers, R. J. (2009) ‘Motor Learning Is Optimally Tuned to the Properties of Motor Noise’, *Neuron*, (63(3):407-417). doi: 10.1016/j.neuron.2009.06.025.

Ben-Gal, G. *et al.* (2011) ‘Preliminary Assessment of Faculty and Student Perception of a Haptic Virtual Reality Simulator for Training Dental Manual Dexterity’, *Journal of Dental Education*, (75(4):496-504).

Ben-Gal, G. *et al.* (2013) ‘Testing manual dexterity using a virtual reality simulator: Reliability and validity’, *European Journal of Dental Education*, (17(3):138-142). doi: 10.1111/eje.12023.

Ben-Gal, G. *et al.* (2017) ‘Testing Motor Learning Curves Among Dental Students.’, *Journal of dental education*, (81(10):1171-1178). doi: 10.21815/JDE.017.076.

Berka, C. *et al.* (2007) ‘EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks’, *Aviation Space and Environmental Medicine*, (78(5 Suppl):B231–44).

Blandin, Y., Lhuisset, L. and Proteau, L. (1999) ‘Cognitive Processes Underlying Observational Learning of Motor Skills’, *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, (52(4):957-979). doi: 10.1080/713755856.

De Boer, I. R. *et al.* (2015) ‘Evaluation of the appreciation of virtual teeth with and without pathology’, *European Journal of Dental Education*, (19(2):87-94). doi: 10.1111/eje.12108.

De Boer, I. R., Wesselink, P. R. and Vervoorn, J. M. (2013) ‘The creation of virtual teeth with and without tooth pathology for a virtual learning environment in dental education’, *European Journal of Dental Education*, (17(4):191–7). doi: 10.1111/eje.12027.

Boring, E. G. (1954) ‘The nature and history of experimental control.’, *The American journal of psychology*, (67(4):573-589). doi: 10.2307/1418483.

Bradley, P. (2006) ‘The history of simulation in medical education and possible future directions.’, *Medical Education*. England, (40(3):254-262). doi: 10.1111/j.1365-2929.2006.02394.x.

Braun, D. A. *et al.* (2009) ‘Motor Task Variation Induces Structural Learning’, *Current Biology*, (19(4):352-357). doi: 10.1016/j.cub.2009.01.036.

Braun, D. A. *et al.* (2010) ‘Structure learning in a sensorimotor association task’, *PLoS ONE*. doi: 10.1371/journal.pone.0008973.

Braun, D. A., Mehring, C. and Wolpert, D. M. (2010) ‘Structure learning in action’, *Behavioural Brain Research*, (206(2):157–165). doi: 10.1016/j.bbr.2009.08.031.

Braver, T. S., Gray, J. R. and Burgess, G. C. (2007) *Explaining the Many Varieties of Working Memory Variation: Dual Mechanisms of Cognitive Control*, *Variation in Working Memory*. Oxford University Press. doi: 10.1093/acprof:oso/9780195168648.003.0004.

Bryan, W. L. and Harter, N. (1897) ‘Studies in the physiology and psychology of the telegraphic language’, *Psychological Review*, (4(1):27–53). doi: 10.1037/h0073806.

Buccino, G., Binkofski, F. and Riggio, L. (2004) ‘The mirror neuron system and action recognition’, *Brain and Language*, (89(2): 370-376). doi: 10.1016/S0093-934X(03)00356-0.

Buchanan, J. (2001) ‘Use of simulation technology in dental education’, *Journal of Dental Education*, (65(11):1225-1231).

Buchanan, J. (2004) ‘Experience with virtual reality-based technology in teaching restorative dental procedures.’, *Journal of Dental Education ducation*, (68(12):1258-1265).

Burgman, M. A. *et al.* (2011) ‘Expert status and performance’, *PLoS ONE*, (6(7): e22998). doi: 10.1371/journal.pone.0022998.

Buschkuehl, M., Jaeggi, S. M. and Jonides, J. (2012) ‘Neuronal effects following working memory training’, *Developmental Cognitive Neuroscience*, (2(1): S167-S179). doi: 10.1016/j.dcn.2011.10.001.

Byrne, A. (2011) ‘Measurement of mental workload in clinical medicine: a review study’, *Anesthesiology and pain medicine*, (2(1):90-94). doi: 10.5812/kowsar.22287523.2045.

Cao, Y. *et al.* (2007) ‘Design and analysis of a new six-DOF haptic device for dental training’, in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*. doi: 10.1109/ROMAN.2007.4415091.

Cattaneo, L. and Rizzolatti, G. (2009) ‘The mirror neuron system’, *Archives of Neurology*, (27:169-92). doi: 10.1001/archneurol.2009.41.

Cavanagh, J. F. *et al.* (2010) ‘Frontal theta links prediction errors to behavioral adaptation in reinforcement learning’, *NeuroImage*, (49(4):3198-3209). doi: 10.1016/j.neuroimage.2009.11.080.

Cavanagh, J. F. and Frank, M. J. (2014) ‘Frontal theta as a mechanism for cognitive control’, *Trends in Cognitive Sciences*, (18(8):414-21). doi: 10.1016/j.tics.2014.04.012.

Cavanagh, J. F. and Shackman, A. J. (2015) ‘Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence', *Journal of Physiology Paris*, (109(1-3):3-15). doi: 10.1016/j.jphysparis.2014.04.003.

Cesqui, B. *et al.* (2008) 'On the use of divergent force fields in robot-mediated neurorehabilitation', *Proceedings of the 2nd Biennial IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, BioRob 2008*. IEEE, pp. 854–861. doi: 10.1109/BIOROB.2008.4762927.

Chambers, D. W. *et al.* (1997) 'Toward a competency analysis of operative dentistry technique skills.', *Journal of Dental Education*, (61(10):795-803).

Chandra, V. *et al.* (2010) 'A comparison of laparoscopic and robotic assisted suturing performance by experts and novices', *Surgery*, (147(6):830–9). doi: 10.1016/j.surg.2009.11.002.

Cheng, A. *et al.* (2016) 'Reporting guidelines for health care simulation research: Extensions to the CONSORT and STROBE statements', *BMJ Simulation and Technology Enhanced Learning*, (2(3):51-60). doi: 10.1136/bmjstel-2016-000124.

Chmarra, M. K. *et al.* (2010) 'Objective classification of residents based on their psychomotor laparoscopic skills', *Surgical Endoscopy*, (24(5): 1031–1039). doi: 10.1007/s00464-009-0721-y.

Churchland, M. M. *et al.* (2012) 'Neural population dynamics during reaching', *Nature*, (7405(487):51–56). doi: 10.1038/nature11129.

Clamann, M. and Kaber, D. B. (2018) 'Augmenting fine motor skill training with haptic error amplification', *Proceedings of the Human Factors and Ergonomics Society*, (62(1):1547–1551). doi: 10.1177/1541931218621350.

Cohen, H. S., Bloomberg, J. J. and Mulavara, A. P. (2005) 'Obstacle avoidance in novel visual environments improved by variable practice training', *Perceptual and Motor Skills*, (101(3):853–61). doi: 10.2466/pms.101.3.853-861.

Cohen, M. X. (2017) 'Where Does EEG Come From and What Does It Mean?', *Trends in Neurosciences*, (40(4):208-218). doi: 10.1016/j.tins.2017.02.004.

Collins, A. and Halverson, R. (2010) 'The second educational revolution: Rethinking education in the age of technology', *Journal of Computer Assisted Learning*, (26(1):18-27). doi: 10.1111/j.1365-2729.2009.00339.x.

Cook, D. *et al.* (2011) 'Technology-enhanced simulation for health professions education: a systematic review and meta-analysis.', *JAMA*, (306(9):978–88). doi: 10.1001/jama.2011.1234.

Cook, D. A. *et al.* (2015) 'A contemporary approach to validity arguments : a practical guide to Kane ' s framework', *Medical Education*, (49(6):560-575). doi: 10.1111/medu.12678.

Corrêa, C. G. *et al.* (2017) 'Virtual Reality simulator for dental anesthesia training in the inferior alveolar nerve block.', *Journal of applied oral science : revista FOB*, (25(4):357-366). doi: 10.1590/1678-7757-2016-0386.

Costa, R. M. (2011) 'A selectionist account of de novo action learning', *Current Opinion in Neurobiology*, (21(4):579-86). doi: 10.1016/j.conb.2011.05.004.

Cox, M. *et al.* (2015) 'A Collaborative Review of the Field of Simulation-

Enhanced Dental Education’, *ADEE Annual Meeting : Communication and Interaction in Dentistry, At Szeged, Hungary*. doi: 10.13140/RG.2.1.2347.7202.

Culmer, P. *et al.* (2012) ‘Reviewing the technological challenges associated with the development of a laparoscopic palpation device’, *International Journal of Medical Robotics and Computer Assisted Surgery*. Robotic Publications, (8(2):146-159). doi: 10.1002/rcs.1421.

Culmer, P. *et al.* (2020) *Haptics in Surgical Robots*, *Handbook of Robotic and Image-Guided Surgery*. doi: 10.1016/b978-0-12-814245-5.00015-3.

D’Angelo, A. L. D., Rutherford, D. N., Ray, R. D., Laufer, S., *et al.* (2015) ‘Idle time: An underdeveloped performance metric for assessing surgical skill’, *American Journal of Surgery*, (209(4):645–51). doi: 10.1016/j.amjsurg.2014.12.013.

D’Angelo, A. L. D., Rutherford, D. N., Ray, R. D., Mason, A., *et al.* (2015) ‘Operative skill: Quantifying surgeon’s response to tissue properties’, *Journal of Surgical Research*, (198(2):294–8). doi: 10.1016/j.jss.2015.04.078.

Dahiya, R. S. *et al.* (2010) ‘Tactile Sensing—From Humans to Humanoids’, *IEEE Transactions on Robotics*, (26(1):1–20). doi: 10.1109/TRO.2009.2033627.

Dargahi, J. and Najarian, S. (2004) ‘Human tactile perception as a standard for artificial tactile sensing-a review.’, *The international journal of medical robotics and computer assisted surgery : MRCAS*. Robotic Publications, (1(1):23-35). doi: 10.1002/rcs.3.

Debarnot, U. *et al.* (2015) ‘Corrigendum: Experts bodies, experts minds: how physical and mental training shape the brain’, *Frontiers in Human Neuroscience*. doi: 10.3389/fnhum.2015.00047.

Dental Education at the Crossroads: Challenges and Change (1995) *National Academy Press*. doi: 10.1017/CBO9781107415324.004.

Department of Health and Social Care (2011) *A Framework for Technology Enhanced Learning.*

Diedenhofen, B. and Musch, J. (2015) ‘Cocor: A comprehensive solution for the statistical comparison of correlations’, *PLoS ONE*, (10(6): e0131499). doi: 10.1371/journal.pone.0121945.

Diedrichsen, J. *et al.* (2010) ‘Use-dependent and error-based learning of motor behaviors’, *Journal of Neuroscience*, (30(15):5159-66). doi: 10.1523/JNEUROSCI.5406-09.2010.

Diedrichsen, J. and Kornysheva, K. (2015) ‘Motor skill learning between selection and execution’, *Trends in Cognitive Sciences*, (19(4):227-233). doi: 10.1016/j.tics.2015.02.003.

Diego, J. P. S. *et al.* (2012) ‘Researching haptics in higher education: The complexity of developing haptics virtual learning systems and evaluating its impact on students’ learning’, *Computers and Education*, (59(1):156-166). doi: 10.1016/j.compedu.2011.11.009.

Dimitrijevic, T. *et al.* (2011) ‘Depth and distance perception of dentists and dental students’, *Operative Dentistry*, (36(5):467-77). doi: 10.2341/10-290-L.

Van Driel, J., Ridderinkhof, K. R. and Cohen, M. X. (2012) ‘Not All Errors

Are Alike: Theta and Alpha EEG Dynamics Relate to Differences in Error-Processing Dynamics’, *Journal of Neuroscience*, (32(47):16795-806). doi: 10.1523/jneurosci.0802-12.2012.

Dubrowski, A. (2005) ‘Performance vs learning curves: What is motor learning and how is it measured?’, *Surgical Endoscopy and Other Interventional Techniques*, (19(9):1290). doi: 10.1007/s00464-004-8261-y.

Duong, J. K. T., Gardner, K. and Rucker, L. M. (2010) ‘Development and retention of fine psychomotor skills: implications for the aging dentist.’, *Journal of Canadian Dental Association*.

Durand, R. (2008) ‘Rethinking Expertise’, *Contemporary Sociology: A Journal of Reviews*. doi: 10.1177/009430610803700656.

Emken, J. L. and Reinkensmeyer, D. J. (2005) ‘Robot-enhanced motor learning: accelerating internal model formation during locomotion by transient dynamic amplification.’, *IEEE transactions on neural systems and rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society*, (13(1):33–9). doi: 10.1109/TNSRE.2004.843173.

Ericsson, K. A. and Charness, N. (1994) ‘Expert performance: Its structure and acquisition.’, *American Psychologist*, (49(8):725–747). doi: 10.1037/0003-066x.49.8.725.

Ericsson, K. A., Krampe, R. T. and Tesch-Römer, C. (1993) ‘The role of deliberate practice in the acquisition of expert performance.’, *Psychological Review*, (100(3):363–406). doi: 10.1037/0033-295x.100.3.363.

Ericsson, K. A. and Lehman, A. C. (1996) ‘Expert and exceptional performance: evidence of maximal adaptation to task constraints’, *Annual Review of Psychology*, (47:273-305). doi: 10.1146/annurev.psych.47.1.273.

Evans, A. W. (2001) ‘Assessing competence in surgical dentistry’, *British Dental Journal*, (190(7):343–6). doi: 10.1038/sj.bdj.4800969a.

Fitts, P. M. and Posner, M. I. (1967) *Human Performance*.

Flanagan, J. R. and Johansson, R. S. (2003) ‘Action plans used in action observation’, *Nature*, (424:769–771). doi: 10.1038/nature01861.

Friston, K. (2010) ‘The free-energy principle: a unified brain theory?’, *Nature Reviews Neuroscience*, (11(2):127-38). doi: 10.1038/nrn2787.

Fugill, M. (2013) ‘Defining the purpose of phantom head’, *European Journal of Dental Education*, (17(1):e1–4). doi: 10.1111/eje.12008.

Gaba, D. M. (2004) ‘The future vision of simulation in health care’, *Quality & safety in health care*, (13(1):i2–10). doi: 10.1136/qhc.13.suppl_1.i2.

Gabriel, A. (2012) *An introduction to Structural Learning-A new approach in Reinforcement Learning*.

Gallagher, A. G. *et al.* (2003) ‘Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures’, *Journal of the American College of Surgeons*, (197(3):479–88). doi: 10.1016/S1072-7515(03)00535-0.

Gardner, E. P. (2010) *Touch*, *Encyclopedia of Life Sciences .* doi: 10.1002/9780470015902.a0000219.pub2.

Gheorghe, D. A., Panouillères, M. T. N. and Walsh, N. D. (2018) 'Psychosocial stress affects the acquisition of cerebellar-dependent sensorimotor adaptation', *Psychoneuroendocrinology*, (92:41–49). doi: 10.1016/j.psyneuen.2018.03.013.

Gobet, F. (2013) 'Expertise vs. talent', *Talent Development and Excellence*, (5(1):75-86).

Gobet, F. (2015) *Understanding Expertise: A Multi-Disciplinary Approach*.

Gottlieb, R., Vervoorn, J. M. and Buchanan, J. (2013) 'Simulation in Dentistry and Oral Health', in *The Comprehensive Textbook of Healthcare Simulation*. Springer, pp. 329–340.

Guida, A. *et al.* (2012) 'How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: A two-stage framework', *Brain and Cognition*, (79(3):221–44). doi: 10.1016/j.bandc.2012.01.010.

Gwilliam, J. C. *et al.* (2010) 'Human vs. robotic tactile sensing: Detecting lumps in soft tissue', in *2010 IEEE Haptics Symposium*, pp. 21–28. doi: 10.1109/HAPTIC.2010.5444685.

Hackshaw, A. (2009) *A Concise Guide to Clinical Trials*. doi: 10.1002/9781444311723.

Hagen, M. E. *et al.* (2008) 'Visual clues act as a substitute for haptic feedback in robotic surgery', *Surgical Endoscopy*, (22(6):1505-1508). doi: 10.1007/s00464-007-9683-0.

Haith, A. M. and Krakauer, J. W. (2018) 'The multiple effects of practice: skill, habit and reduced cognitive load', *Current Opinion in Behavioral Sciences*, (20:196-201). doi: 10.1016/j.cobeha.2018.01.015.

Hall, J. C., Ellis, C. and Hamdorf, J. (2003) 'Surgeons and cognitive processes', *British Journal of Surgery*, (90(1):10–6). doi: 10.1002/bjs.4020.

Hambrick, D. Z. *et al.* (2014) 'Deliberate practice: Is that all it takes to become an expert?', *Intelligence*, (45:34-45). doi: 10.1016/j.intell.2013.04.001.

Hamstra, S. J. *et al.* (2014) 'Reconsidering fidelity in simulation-based training', *Academic Medicine*, (89(3):387-92). doi: 10.1097/ACM.0000000000000130.

Harris, D. J. *et al.* (2018) 'Action observation for sensorimotor learning in surgery', *British Journal of Surgery*, (105(13):1713-1720). doi: 10.1002/bjs.10991.

Harrysson, I. J. *et al.* (2014) 'Systematic review of learning curves for minimally invasive abdominal surgery: A review of the methodology of data collection, depiction of outcomes, and statistical analysis', *Annals of Surgery*, (260(1):37–45). doi: 10.1097/SLA.0000000000000596.

Hauser, A. M. and Bowen, M. D. (2009) 'Primer on Preclinical Instruction and Evaluation', (73(3):390–8).

HEA (2009) *Transforming Higher Education Through Technology-Enhanced Learning*.

Heathcote, A., Brown, S. and Mewhort, D. J. K. (2000) 'The power law repealed: The case for an exponential law of practice', *Psychonomic Bulletin* *and Review*, (7:185–207). doi: 10.3758/BF03212979.

Heller, M. and Schiff, W. (1991) *The psychology of touch.* Edited by M. A. Heller and William Schiff.

Hendee, W. R. (2001) ‘To Err is Human: Building a Safer Health System’, *Journal of Vascular and Interventional Radiology*. doi: 10.1016/s1051-0443(01)70072-3.

Hennessey, I. a M. and Hewett, P. (2013) ‘Construct, concurrent, and content validity of the eoSim laparoscopic simulator.’, *Journal of laparoendoscopic & advanced surgical techniques*, (23(10):855-60). doi: 10.1089/lap.2013.0229.

Heuer, H. and Lüttgen, J. (2015) ‘Robot assistance of motor learning: A neuro-cognitive perspective’, *Neuroscience and Biobehavioral Reviews*. Elsevier Ltd, (56:222–40). doi: 10.1016/j.neubiorev.2015.07.005.

Heyes, C. M. and Foster, C. L. (2002) ‘Motor learning by observation: Evidence from a serial reaction time task’, *Quarterly Journal of Experimental Psychology*, (55(2):593-607). doi: 10.1080/02724980143000389.

Hick, W. E. (1952) ‘On the Rate of Gain of Information’, *Quarterly Journal of Experimental Psychology*, (4(1):11–26). doi: 10.1080/17470215208416600.

Higher Education Funding Council for England (2009) ‘Enhancing learning and teaching through the use of technology: A revised approach to HEFCE’s strategy for e-learning’.

Hofstad, E. F. *et al.* (2013) ‘A study of psychomotor skills in minimally invasive surgery: What differentiates expert and nonexpert performance’, *Surgical Endoscopy*, (27(3):854-63). doi: 10.1007/s00464-012-2524-9.

Hollis, W., Darnell, L. A. and Hottel, T. L. (2011) ‘Computer assisted learning: a new paradigm in dental education.’, *The Journal of the Tennessee Dental Association*, (91(4):14–8).

Holm, A. *et al.* (2009) ‘Estimating brain load from the EEG’, *The Scientific World Journal*, (9:639-651). doi: 10.1100/tsw.2009.83.

Huang, V. S. *et al.* (2011) ‘Rethinking Motor Learning and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines with Internal Models’, *Neuron*, (70(4):787-801). doi: 10.1016/j.neuron.2011.04.012.

Huang, V. S. and Shadmehr, R. (2007) ‘Evolution of motor memory during the seconds after observation of motor error.’, *Journal of neurophysiology*, (97(6):3976–85). doi: 10.1152/jn.01281.2006.

Hung, A. J. *et al.* (2012) ‘Concurrent and predictive validation of a novel robotic surgery simulator: A prospective, randomized study’, *Journal of Urology*, (187(2):630–7). doi: 10.1016/j.juro.2011.09.154.

Huth, K. C. *et al.* (2017) ‘Assessment of practical tasks in the Phantom course of Conservative Dentistry by pre-defined criteria: a comparison between self-assessment by students and assessment by instructors’, *European Journal of Dental Education*, (21(1):37-45). doi: 10.1111/eje.12176.

Imber, S. *et al.* (2003) ‘A virtual reality dental simulator predicts performance in an operative dentistry manikin course’, *European Journal of*

Dental Education, (7:160-163). doi: 10.1111/eje.12176.

Issenberg, B. and Scalese, R. (2007) 'Best evidence on high-fidelity simulation: what clinical teachers need to know', *The Clinical Teacher*, (4(2):73-77). doi: 10.1111/j.1743-498X.2007.00161.x.

Issenberg, S. *et al.* (2005) 'Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review.', *Medical teacher*, (27(1):10-28). doi: 10.1080/01421590500046924.

Jäncke, L., Shah, N. J. and Peters, M. (2000) 'Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists', *Cognitive Brain Research*, (10(1-2):177-83). doi: 10.1016/S0926-6410(00)00028-8.

Jasinevicius, T. R. *et al.* (2004) 'An evaluation of two dental simulation systems: virtual reality versus contemporary non-computer-assisted', *Journal of Dental Education*, (68(11):1151).

Johnson, K. O. and Hsiao, S. S. (1992) 'Neural Mechanisms of Tactual form and Texture Perception', *Annual Review of Neuroscience*, (15:227-50). doi: 10.1146/annurev.ne.15.030192.001303.

Kim, L. *et al.* (2005) 'Dental Training System using Multi-modal Interface', *Computer-Aided Design and Applications*, (2(5):591-598). doi: 10.1080/16864360.2005.10738323.

Kirkwood, A. and Price, L. (2014) 'Technology-enhanced learning and teaching in higher education: what is "enhanced" and how do we know? A critical literature review', *Learning, Media and Technology*, (39(1):6–36). doi: 10.1080/17439884.2013.770404.

Kok, E. *et al.* (2018) 'The neural implementation of surgical expertise within the mirror-neuron system: An fMRI study', *Frontiers in Human Neuroscience*, (12: 291). doi: 10.3389/fnhum.2018.00291.

Komar, J., Seifert, L. and Thouvarecq, R. (2016) 'What variability tells us about motor expertise: Measurements and perspectives from a complex system approach', *Movement and Sports Sciences - Science et Motricite*. doi: 10.1051/sm/2015020.

Konukseven, E. I. *et al.* (2010) 'Development of a visio-haptic integrated dental training simulation system', *Journal of Dental Education*, (74(8):880-91).

Krakauer, J. W. *et al.* (2006) 'Generalization of motor learning depends on the history of prior action', *PLoS Biology*, 4(10), pp. 1798–1808. doi: 10.1371/journal.pbio.0040316.

Krakauer, J. W. *et al.* (2019) 'Motor Learning', *Comprehensive Physiology*, (9(2):613-663). doi: doi:10.1002/cphy.c170043.

Krakauer, J. W., Ghez, C. and Ghilardi, M. F. (2005) 'Adaptation to Visuomotor Transformations: Consolidation, Interference, and Forgetting', *Journal of Neuroscience*, 25(25(2):473-478). doi: 10.1523/JNEUROSCI.4218-04.2005.

Lashley, K. (1951) 'The problem of serial order in behavior', *Cerebral Mechanisms in Behavior*.

LeBlanc, V. *et al.* (2004) 'A preliminary study in using virtual reality to train dental students', *Journal of Dental Education*, (68(3):378-383).

Lee, J. and Choi, S. (2010) 'Effects of haptic guidance and disturbance on motor learning: Potential advantage of haptic disturbance', *2010 IEEE Haptics Symposium, HAPTICS 2010*. IEEE, pp. 335–342. doi: 10.1109/HAPTIC.2010.5444635.

Leinster, S. (2009) 'Learning in the clinical environment', *Medical teacher*, (31(2):79-81). doi: 10.1080/01421590902744936.

Levine, A. et al (2013) *The Comprehensive Textbook of Healthcare Simulationle*, *New York, NY: Springer.*

Lewis-Beck, M., Bryman, A. and Futing Liao, T. (2004) 'Floor Effect'. Thousand Oaks, California: The SAGE Encyclopedia of Social Science Research Methods. doi: 10.4135/9781412950589.

Lewis, F. R. and Klingensmith, M. E. (2012) 'Issues in general surgery residency training', *Annals of Surgery*, (256(4):553–9). doi: 10.1097/SLA.0b013e31826bf98c.

Littlewood, K. (2011) 'High fidelity simulation as a research tool.', *Best practice & research. Clinical anaesthesiology*, (25(4):473-87). doi: 10.1016/j.bpa.2011.08.001.

Longo, M. R. and Haggard, P. (2011) 'Weber's illusion and body shape: Anisotropy of tactile size perception on the hand', *Journal of Experimental Psychology: Human Perception and Performance*, (37(3):720–6). doi: 10.1037/a0021921.

Lotze, M. *et al.* (2003) 'The musician's brain: functional imaging of amateurs and professionals during performance and imagery.', *NeuroImage*, (20(3):1817-29). doi: 10.1016/j.neuroimage.2003.07.018.

Lowe, D. J. *et al.* (2016) 'Feasibility of EEG to monitor cognitive performance during venous cannulation: EEG Distracted Intravenous Access (E-DIVA)', *BMJ Simulation and Technology Enhanced Learning*, (2:68-72).

Macey, R. *et al.* (2018) 'Tests to detect and inform the diagnosis of caries', *Cochrane Database of Systematic Reviews*. doi: 10.1002/14651858.CD013215.

Maia, T. V. (2009) 'Reinforcement learning, conditioning, and the brain: Successes and challenges', *Cognitive, Affective and Behavioral Neuroscience*, (9(4):343-64). doi: 10.3758/CABN.9.4.343.

Manogue, M., Brown, G. and Foster, H. (2001) 'Clinical assessment of dental students: Values and practices of teachers in restorative dentistry', *Medical Education*, (35(4):364-70). doi: 10.1046/j.1365-2923.2001.00733.x.

Marchal-Crespo, L. *et al.* (2014) 'Learning a locomotor task: with or without errors?', *Journal of Neuroengineering and Rehabilitation*, (11:25). doi: 10.1186/1743-0003-11-25.

Mason, D. (2005) 'Oswald Fergus , a pioneer of simulated dental clinical practice.', *The Henry Noble History of Dentistry Research Group Newsletter*, pp. 1–4.

Matsuoka, Y., Brewer, B. R. and Klatzky, R. L. (2007) 'Using visual feedback distortion to alter coordinated pinching patterns for robotic rehabilitation', *Journal of NeuroEngineering and Rehabilitation*. doi:

10.1186/1743-0003-4-17.

Mattar, A. A. G. and Gribble, P. L. (2005) ‘Motor learning by observing’, *Neuron*, (46(1):153-160). doi: 10.1016/j.neuron.2005.02.009.

McDougle, S. D. *et al.* (2016) ‘Credit assignment in movement-dependent reinforcement learning’, *Proceedings of the National Academy of Sciences of the United States of America*, (13(24):6797-802). doi: 10.1073/pnas.1523669113.

McDougle, S. D. *et al.* (2019) ‘Neural Signatures of Prediction Errors in a Decision-Making Task Are Modulated by Action Execution Failures’, *Current Biology*, (29(10):1606-1613.e5). doi: 10.1016/j.cub.2019.04.011.

Miall, R. C. *et al.* (1993) ‘Is the cerebellum a smith predictor?’, *Journal of motor behavior*, 25(25(3):203-216). doi: 10.1080/00222895.1993.9942050.

Miall, R. C. and Wolpert, D. M. (1996) ‘Forward models for physiological motor control’, *Neural Networks*, (9(8):1265-1279). doi: 10.1016/S0893-6080(96)00035-4.

Mihelj, M. and Podobnik, J. (2012) *Haptics for Virtual Reality and Teleoperation*.

Miller, E. K. and Cohen, J. D. (2001) ‘An integrative theory of prefrontal cortex function.’, *Annual Review of Neuroscience*, (24:167-202). doi: 10.1146/annurev.neuro.24.1.167.

Minogue, J. and Jones, M. G. (2006) ‘Haptics in Education: Exploring an Untapped Sensory Modality’, *Review of Educational Research*, (76(3):317-348). doi: 10.3102/00346543076003317.

Mirghani, I. *et al.* (2016) ‘Capturing differences in dental training using a virtual reality simulator’, *European Journal of Dental Education*, (22(1):67-71). doi: 10.1111/eje.12245.

Mirghani, I. *et al.* (2019) ‘The factors that count in selecting future dentists: sensorimotor and soft skills’, *British Dental Journal*, (226(6):417-421). doi: 10.1038/s41415-019-0030-3.

Morris, M. C. *et al.* (2015) ‘Assessment of competence in surgical skills using functional magnetic resonance imaging: A feasibility study’, *Journal of Surgical Education*, (72(2):198-204). doi: 10.1016/j.jsurg.2014.09.007.

Münte, T. F., Altenmüller, E. and Jäncke, L. (2002) ‘The musician’s brain as a model of neuroplasticity’, *Nature Reviews Neuroscience*, (3:473–478). doi: 10.1038/nrn843.

Mushtaq, F. *et al.* (2016) ‘Randomised prior feedback modulates neural signals of outcome monitoring’, *NeuroImage*, (125:868-879). doi: 10.1016/j.neuroimage.2015.10.046.

Mushtaq, F., Bland, A. R. and Schaefer, A. (2011) ‘Uncertainty and cognitive control’, *Frontiers in Psychology*, (3(2):249). doi: 10.3389/fpsyg.2011.00249.

Mylopoulos, M. and Regehr, G. (2011) ‘Putting the expert together again’, *Medical Education*, (45(9):920–6). doi: 10.1111/j.1365-2923.2011.04032.x.

Newell, A. and Rosenbloom, P. S. (1981) ‘Mechanisms of skill acquisition and the law of practice’, *Cognitive skills and their acquisition*.

Newell, K. M. (1991) ‘Motor skill acquisition’, *Physical therapy*, (42:213–37). doi: 10.1146/annurev.ps.42.020191.001241.

Newell, K. M., Liu, Y. T. and Mayer-Kress, G. (2001) ‘Time scales in motor learning and development’, *Psychological Review*, (108(1):57–82). doi: 10.1037/0033-295X.108.1.57.

Norman, D. and Shallice, T. (1986) ‘Attention to action: willed and automatic control of behavior’, in *Consciousness and Self-Regulation: Advances in Research and Theory IV*. doi: 10.1007/978-1-4757-0629-1_1.

Okuda, Y. *et al.* (2009) ‘The utility of simulation in medical education: What is the evidence?’, *Mount Sinai Journal of Medicine*, (76:330-343). doi: 10.1002/msj.20127.

Olejnik, S. and Algina, J. (2003) ‘Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs’, *Psychological Methods*, (8(4):434-447). doi: 10.1037/1082-989X.8.4.434.

Oropesa, I. *et al.* (2011) ‘Methods and tools for objective assessment of psychomotor skills in laparoscopic surgery’, *Journal of Surgical Research*, (171(1):e81–95). doi: 10.1016/j.jss.2011.06.034.

Oropesa, I. *et al.* (2013) ‘Relevance of motion-related assessment metrics in laparoscopic surgery’, *Surgical Innovation*, (20(3):299-312). doi: 10.1177/1553350612459808.

Owen, H. (2016) *Simulation in healthcare education: An exclusive history*.

Oxford English Dictionary (2017) *Oxford English Dictionary Online*, *Oxford English Dictionary*.

Pacheco, M. M. and Newell, K. M. (2018) ‘Transfer of a learned coordination function: Specific, individual and generalizable’, *Human movement science*, (59:66-80). doi: 10.1016/j.humov.2018.03.019.

Parcutilo and Luna (2016) ‘Working memory’, *American Journal of Educational Research*, (2(2):166-168). doi: 10.12691/education-4-18-2.

Partridge, R. W. *et al.* (2014) ‘Accessible laparoscopic instrument tracking (“InsTrac”): construct validity in a take-home box simulator.’, *Journal of laparoendoscopic & advanced surgical techniques*, (24(8):578-83). doi: 10.1089/lap.2014.0015.

Patel, U. S. *et al.* (2018) ‘Assessment in a global context: An international perspective on dental education’, *European Journal of Dental Education*, (22(S1):21-27). doi: 10.1111/eje.12343.

Perry, S. *et al.* (2015) ‘A review of the use of simulation in dental education’, *Simulation in healthcare : journal of the Society for Simulation in Healthcare*, (10(1):31-37). doi: 10.1097/SIH.0000000000000059.

Peterson, D. and Robertson, C. (2013) ‘Utilizing Virtual Reality in Teaching and Training : Progress Toward Virtual Surgical Simulations.’, in *7th International Association of Technology, Education and Development Volume: 7*.

Puryer, J. and O’Sullivan, D. (2015) ‘An introduction to standard setting methods in dentistry’, *British Dental Journal*, (219(7):355–8). doi: 10.1038/sj.bdj.2015.755.

Pusic, M., Pecaric, M. and Boutis, K. (2011) ‘How Much Practice Is Enough? Using Learning Curves to Assess the Deliberate Practice of Radiograph Interpretation’, *Academic Medicine*, (86(6):731-736). doi: 10.1097/ACM.0b013e3182178c3c.

Quinn, F. *et al.* (2003) ‘A study comparing the effectiveness of conventional training and virtual reality simulation in the skills acquisition of junior dental students.’, *European Journal of Dental Education*, (7(4):164-169). doi: 10.1034/j.1600-0579.2003.00309.x.

R Foundation for Statistical Computing. (2018) *R: a Language and Environment for Statistical Computing.*, *http://www.R-project.org/*.

Rabbi, A. F. *et al.* (2009) ‘Human performance evaluation based on EEG signal analysis: A prospective review’, in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1879–1882. doi: 10.1109/IEMBS.2009.5333877.

Ramsay, C. R. *et al.* (2002) ‘Assessing the learning curve effect in health technologies. Lessons from the nonclinical literature.’, *International journal of technology assessment in health care*, (18(1):1–10).

Raw, R. K. *et al.* (2019) ‘Skill acquisition as a function of age, hand and task difficulty: Interactions between cognition and action’, *PLoS ONE*, (14(2): e0211706). doi: 10.1371/journal.pone.0211706.

Reinkensmeyer, D. J. and Patton, J. L. (2009) ‘Can robots help the learning of skilled actions?’, *Exercise and sport sciences reviews*, (37(1):43-51). doi: 10.1097/JES.0b013e3181912108.

Retrosi, G. *et al.* (2015) ‘Motion Analysis-Based Skills Training and Assessment in Pediatric Laparoscopy: Construct, Concurrent, and Content Validity for the eoSim Simulator.’, *Journal of laparoendoscopic & advanced surgical techniques*, (25(11):944-950). doi: 10.1089/lap.2015.0069.

Rhienmora, P. *et al.* (2008) ‘Development of a dental skills training simulator using virtual reality and haptic device’, *National Electronics and Computer Technology Center Technical Journal*.

Ridderinkhof, K. R. *et al.* (2004) ‘The role of the medial frontal cortex in cognitive control’, *Science*, (306(5695): 443-447). doi: 10.1126/science.1100301.

Roberts, J. W. *et al.* (2014) ‘Top-down and bottom-up processes during observation: Implications for motor learning’, *European Journal of Sport Science*, (14(1):S250–6). doi: 10.1080/17461391.2012.686063.

Robles-De-La-Torre, G. (2010) ‘Virtual Reality: Touch/Haptics’, *Encyclopedia of Perception*. doi: 10.4135/9781412972000.

Rodger, M., Tang, T. and White, C. (2016) ‘Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills’, *The Obstetrician & Gynaecologist*, (18(1):53-63). doi: 10.1111/tog.12255.

Ross, M. (2004) ‘The increasing demand and necessity for a team approach’, *British Dental Journal*, (196:181). doi: 10.1038/sj.bdj.4810946.

Roy, E., Bakr, M. and George, R. (2017) ‘The need for virtual reality simulators in dental education: A review’, *The Saudi Dental Journal*, (29(2):41-47). doi: 10.1016/j.sdentj.2017.02.001.

Sadideen, H. *et al.* (2013) ‘Surgical experts: Born or made?’, *International Journal of Surgery*, (11(9):773–8). doi: 10.1016/j.ijsu.2013.07.001.

Sadnicka, A. *et al.* (2017) ‘A unifying motor control framework for task-specific dystonia.’, *Nature reviews. Neurology*, (14(2):116-124). doi: 10.1038/nrneurol.2017.146.

Sakai, K., Kitaguchi, K. and Hikosaka, O. (2009) ‘Chunking during human visuomotor sequence learning.’, *Experimental brain research*. Springer, (152(2):229-242). doi: 10.1007/s00221-003-1548-8.

Sands, W. A. (2017) ‘Practice’, in *The Science of Gymnastics: Advanced Concepts: Second Edition*. doi: 10.4324/9781315203805.

Schmidt, R. A. R. and Wrisberg, C. A. (2000) *Motor Learning and Performance*, *Human Kinetics*.

Schmidt, R. A. and Young, D. E. (1986) ‘Transfer of movement control in motor skill learning’. doi: 10.1016/B978-0-12-188950-0.50009-6.

Schmitz, C. C. *et al.* (2014) ‘Development and verification of a taxonomy of assessment metrics for surgical technical skills’, *Academic Medicine*, (89(1):153–61). doi: 10.1097/ACM.0000000000000056.

Seidler, R. D. (2004) ‘Multiple Motor Learning Experiences Enhance Motor Adaptability’, *Journal of Cognitive Neuroscience*, (6(1):65-73). doi: 10.1162/089892904322755566.

Seidler, R. D. *et al.* (2013) ‘Neurocognitive mechanisms of error-based motor learning’, *Advances in Experimental Medicine and Biology*, (782:39–60). doi: 10.1007/978-1-4614-5465-6_3.

Sevdalis, N. *et al.* (2016) ‘Joint leap into a future of high-quality simulation research: standardising the reporting of simulation science’, *BMJ Simulation and Technology Enhanced Learning*, (2(3):49-50). doi: 10.1136/bmjstel-2016-000127.

Shahriari-Rad, A. (2013) ‘The Extent of Virtual Reality Simulation Enhancing Assessment of Psychomotor Clinical Skills’, in *Proceedings of E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*.

Shahriari-Rad, A., Cox, M. and Woolford, M. (2017) ‘Clinical Skills Acquisition: Rethinking Assessment Using a Virtual Haptic Simulator’, *Technology, Knowledge and Learning*, (22:185–197). doi: 10.1007/s10758-017-9308-1.

Shaker, M. M. (2007) ‘EEG Waves Classifier using Wavelet Transform and Fourier Transform’, *International Journal of Bioengineering and Life Sciences*.

Shirzad, N. and Van Der Loos, H. F. M. H. (2012) ‘Error amplification to promote motor learning and motivation in therapy robotics’, in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. doi: 10.1109/EMBC.2012.6346821.

Shmuelof, L., Krakauer, J. W. and Mazzoni, P. (2012) ‘How is a motor skill learned? Change and invariance at the levels of task success and trajectory control.’, *Journal of Neurophysiology*, (108(2):578-594). doi: 10.1152/jn.00856.2011.

Shumway-Cook, A. and Woollacott, M. H. (2014) *Motor control: Translating research into clinical practice: Fourth edition*, *Motor Control: Translating Research into Clinical Practice: Fourth Edition*.

Sigrist, R. *et al.* (2013) 'Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review', *Psychonomic Bulletin & Review*, (20(1):21-53). doi: 10.3758/s13423-012-0333-8.

da Silva, F. L. (2013) 'EEG and MEG: Relevance to Neuroscience', *Neuron*, (80(5):1112-1128). doi: https://doi.org/10.1016/j.neuron.2013.10.017.

Smith, P. L. and Little, D. R. (2018) 'Small is beautiful: In defense of the small-N design', *Psychonomic Bulletin and Review*, (25(6):2083-2101). doi: 10.3758/s13423-018-1451-8.

Soumya, B. G. and Ramachandra, S. S. (2011) 'Haptics application in dentistry: Is the time poised yet?', *Dental Hypotheses*. doi: 10.5436/j.dehy.2011.2.00019.

Sparrow, W. A. (1983) 'The efficiency of skilled performance', *Journal of Motor Behavior*, (15:3:237-261). doi: 10.1080/00222895.1983.10735299.

Stein, C. D. *et al.* (2014) 'What dental educators need to understand about emerging technologies to incorporate them effectively into the educational process', *Journal of Dental Education*, (78(4):520–9).

Steinberg, A. D. *et al.* (2007) 'Assessment of faculty perception of content validity of PerioSim, a haptic-3D virtual reality dental training simulator.', *Journal of Dental Education*, (71(12):1574-1582).

Sturm, L. P. *et al.* (2008) 'A systematic review of skills transfer after surgical simulation training', *Annals of Surgery*, (248(2):166–79). doi: 10.1097/SLA.0b013e318176bf24.

Suebnukarn, S. *et al.* (2009) 'Process and outcome measures of expert/novice performance on a haptic virtual reality system', *Journal of Dentistry*, (37(9):658-665). doi: 10.1016/j.jdent.2009.04.008.

Suebnukarn, S. *et al.* (2014) 'Construct validity and expert benchmarking of the haptic virtual reality dental simulator', *Journal of Dental Education*, (78(10):1442–50).

Tamè, L., Azañón, E. and Longo, M. R. (2019) 'A conceptual model of tactile processing across body features of size, shape, side, and spatial location', *Frontiers in Psychology*, (10: 291). doi: 10.3389/fpsyg.2019.00291.

Taylor, C. L., Grey, N. and Satterthwaite, J. D. (2013) 'Assessing the Clinical Skills of Dental Students: A Review of the Literature', *Journal of Education and Learning*, (2(1): 20-31). doi: 10.5539/jel.v2n1p20.

Thimbleby, H. (2013) 'Technology and the future of healthcare', *Journal of Public Health Research*, (2(3): e28). doi: 10.4081/jphr.2013.e28.

Tiwana, M. I., Redmond, S. J. and Lovell, N. H. (2012) 'A review of tactile sensing technologies with applications in biomedical engineering', *Sensors and Actuators A: Physical*, (179: 17–31). doi: https://doi.org/10.1016/j.sna.2012.02.051.

Todorov, E. and Jordan, M. I. (2002) 'Optimal feedback control as a theory of motor coordination', *Nature Neuroscience*, 5(5(11):1226-1235). doi:

10.1038/nn963.

Towers, A. *et al.* (2019) ‘A scoping review of the use and application of virtual reality in pre-clinical dental education’, *British dental journal*, (226(5):358). doi: 10.1038/s41415-019-0041-0.

Tresilian, J. (2012) *Sensorimotor control and learning: An introduction to the behavioral neuroscience of action.*

Tse, B. *et al.* (2010) ‘Design and development of a haptic dental training system - HapTEL’, in *Haptics: Generating and Perceiving Tangible Sensations. EuroHaptics 2010.* doi: 10.1007/978-3-642-14075-4_15.

Turnham, E. J. A., Braun, D. A. and Wolpert, D. M. (2011) ‘Facilitation of learning induced by both random and gradual visuomotor task variation’, *Journal of Neurophysiology*, (107(4):1111–22). doi: 10.1152/jn.00635.2011.

Uemura, M. *et al.* (2014) ‘Analysis of hand motion differentiates expert and novice surgeons’, *Journal of Surgical Research*, (188(1):8–13). doi: 10.1016/j.jss.2013.12.009.

Urbankova and Engebretson (2011) ‘Computer-assisted dental simulation as a predictor of preclinical operative dentistry performance.’, *Journal of Dental Education*, (75(9):1249–55).

Vandierendonck, A. (2017) ‘A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure’, *Behavior Research Methods*, (49(2):653-673). doi: 10.3758/s13428-016-0721-5.

Vervoorn, J. *et al.* (2015) ‘Evidence from the Dental Education Communities: review of the field.’, in *Symposium on Technology Enhanced Learning to Enrich Dental Education at the International Association for Dental Research Conference*. Boston.

Vickers, A. J. (2001) ‘The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: A simulation study’, *BMC Medical Research Methodology*. doi: 10.1186/1471-2288-1-6.

Virtual Plant Cell Evaluation (2019). Available at: https://culturecounts.cc/r/plant-energy-biology/2019/#0.

Wang, D. X. *et al.* (2003) ‘Development of dental training system with haptic display’, in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*. doi: 10.1109/ROMAN.2003.1251836.

Weeks, D. L. and Anderson, L. P. (2000) ‘The interaction of observational learning with overt practice: Effects on motor skill learning’, *Acta Psychologica*, (104(2):259–71). doi: 10.1016/S0001-6918(00)00039-1.

Wei, Y. *et al.* (2005) ‘Visual error augmentation for enhancing motor learning and rehabilitative relearning’, in *Proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics*. doi: 10.1109/ICORR.2005.1501152.

Welch, P. D. (1967) ‘The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms’, *IEEE Transactions on Audio and Electroacoustics*. doi: 10.1109/TAU.1967.1161901.

Wielinga, B. J., Bredeweg, B. and Breuker, J. A. (1988) 'Knowledge acquisition for expert systems', in. doi: 10.1007/3-540-50676-4_11.

Wierinck, E. *et al.* (2007) 'Expert performance on a virtual reality simulation system', *Journal of Dental Education*, (71(6):759-766).

Williams, C. K., Tremblay, L. and Carnahan, H. (2016) 'It pays to go off-track: Practicing with error-augmenting haptic feedback facilitates learning of a curve-tracing task', *Frontiers in Psychology*. doi: 10.3389/fpsyg.2016.02010.

Wolpert, D. M., Diedrichsen, J. and Flanagan, J. R. (2011) 'Principles of sensorimotor learning', *Nature Reviews Neuroscience*, (12(12):739-51). doi: 10.1038/nrn3112.

Wolpert, D. M. and Flanagan, J. R. (2010) 'Motor learning', *Current Biology*, (20(11):467-472). doi: 10.1016/j.cub.2010.04.035.

Wolpert, D. M., Ghahramani, Z. and Jordan, M. (1995) 'An Internal Model for Sensorimotor Integration', *Science*, (269(5232):1880–2). doi: 10.1126/science.7569931.

Wolpert, D. M. and Kawato, M. (1998) 'Multiple paired forward and inverse models for motor control', *Neural Networks*, (11(7-8):1317-1329). doi: 10.1016/S0893-6080(98)00066-5.

Wright, T. P. (1936) 'Factors Affecting the Cost of Airplanes', *Journal of the Aeronautical Sciences*. doi: 10.2514/8.155.

Wu, T. *et al.* (2016) 'The Capacity of Cognitive Control Estimated from a Perceptual Decision Making Task', *Scientific Reports*, (6:34025). doi: 10.1038/srep34025.

Wulf, G. and Lewthwaite, R. (2016) 'Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning', *Psychonomic Bulletin and Review*, (23(5):1382-1414). doi: 10.3758/s13423-015-0999-9.

Wulf, G., Shea, C. and Lewthwaite, R. (2010) 'Motor skill learning and performance: A review of influential factors', *Medical Education*, (44(1):75-84). doi: 10.1111/j.1365-2923.2009.03421.x.

Xia, P., Lopes, A. M. and Restivo, M. T. (2013) 'Virtual reality and haptics for dental surgery: A personal review', *Visual Computer*, (29(5):433-447). doi: 10.1007/s00371-012-0748-2.

Yau, H. T., Tsou, L. S. and Tsai, M. J. (2006) 'Octree-based Virtual Dental Training System with a Haptic Device', *Computer-Aided Design and Applications*, (3(1-4):415-424). doi: 10.1080/16864360.2006.10738480.

Yoshida, Y. *et al.* (2011) 'Development of a multi-layered virtual tooth model for the haptic dental training system', *Dental Materials Journal*, (30(1):1–6). doi: 10.4012/dmj.2010-082.

Zavala, B. *et al.* (2018) 'Cognitive control involves theta power within trials and beta power across trials in the prefrontal-subthalamic network', *Brain*, (141(12): 3361–3376). doi: 10.1093/brain/awy266.

Zendejas, B. *et al.* (2013) 'State of the evidence on simulation-based training for laparoscopic surgery: A systematic review', *Annals of Surgery*, (257(4):586–93). doi: 10.1097/SLA.0b013e318288c40b.