The efficacy of virtual reality in professional soccer

being a thesis submitted in fulfilment of the
requirements for the degree of
Doctor of

Philosophy in Sport, Health, and Exercise Science

in the University of Hull

by

Benjamin Thomas Greenhough, MSc, BSc
March 2023

Acknowledgements

First and foremost, I must thank my three supervisors, Steve Barrett, Chris Towlson, and Grant Abt. Applied and academic knowledge was required for this thesis, to which this supervisory team provided the perfect blend. Steve, thank you for challenging me, both in the PhD and as a practitioner. Our daily interactions (and your terrible banter) at the football club were a welcomed break from the thesis, and even after you left you would never be short of time to guide me and provide valuable advice. Chris, thank you for always bringing your positive persona into our conversations. Every interaction left me feeling revitalised to keep moving forward. Your support during the survey chapter was invaluable and set the benchmark for the standard of work required in future studies. Lastly, Grant, I cannot thank you enough for getting me to this stage. You constantly challenged my critical thinking and developed me into a better researcher. Even in my current role, you have provided guidance and mentoring, and got the very best out of me, to which I'm hugely grateful. Our '30 minute' meetings would leave me with more questions than answers, and it's fair to say that since our first encounter, I'm only a little closer to understanding half of the concepts you come up with. But with time, that will come, and I'm looking forward to learning so much more from you.

Next, I must thank the funders of this thesis, Hull City Football Club for offering this PhD studentship. From the lows of relegation to the highs of promotion, it was a wonderful experience made even more special by the many colleagues I got to spend with. A special thank you to Matthew Busby, for his guidance and endless "character building". It certainly developed me into a stronger person. Thank you to Aimee Evans, for supporting me as we started this journey together, and to Callum MacMaster and Warren Bradley, for their humour and friendly distraction at work. A thank you also to Alan Peacham, who's daily "did you do a bit last night" in reference to the PhD was a welcomed acknowledgement to the work being put into the thesis, and much needed at times. Lastly, I would like to thank all the talented team at Rezzil for their technical expertise needed to complete this thesis, with a special thank you to Andy Etches who had faith in me since our first introduction. Before, during, and after my time at Hull City, you have provided me with the greatest opportunities, to develop and grow in my career, and I appreciate everything you have done for me.

A huge thanks my family and friends for their ongoing love and support throughout the last four years. To my parents, thank you for always believing in me, and pushing me and my brothers to follow our aspirations. To my brother, Liam, thank you for your silent but ever encouraging words of wisdom, and to Luke, for engaging in many stories as we discussed our own joys and struggles of doing a PhD. I'll be forever reminded to the fact you submitted your thesis first, by a narrow margin of five days. To my friend Jack Shreeve, thank you for always providing the

words of encouragement, through the many PhD applications I applied to before I finally got my opportunity. And to Neil Wilyman, thank you for providing comfort during the most difficult period of the PhD. Lastly, to my beautiful fiancé Maria Cepeda, thank you for being so patient, supportive, and loving throughout this journey. Even at the most challenging of times, you’ve always known the right things to say to make me smile. I look forward to sharing many more journeys with you.

Publications and Conferences

Greenhough, B., Barrett, S., Towlson, C., & Abt, G. (2021). Perceptions of professional soccer coaches, support staff and players toward virtual reality and the factors that modify their intention to use it. PLOS ONE, 16(12), e0261378. https://doi.org/10.1371/journal.pone.0261378

Abstract

Professional soccer clubs have taken an interest to virtual reality, however, only a paucity of evidence exists to support its use in the soccer training ground environment. Further, several soccer virtual reality companies have begun providing solutions to teams, claiming to test specific characteristics of players, yet supportive evidence for certain measurement properties remain absent from the literature. The aims of this thesis were to explore the efficacy of virtual reality being used in the professional football training ground environment. To do so, this thesis looked to explore the fundamental measurement properties of soccer specific virtual reality tests, along with the perceptions of professional coaches, backroom staff, and players that could use virtual reality.

The first research study (Chapter 3) aimed to quantify the learning effect during familiarisation trials of a soccer-specific virtual reality task. Thirty-four professional soccer players age, stature, and body mass: mean (SD) 20 (3.4) years; 180 (7) cm; 79 (8) kg, participated in six trials of a virtual reality soccer passing task. The task required participants to receive and pass 30 virtual soccer balls into highlighted mini-goals that surrounded the participant. The number of successful passes were recorded in each trial. The one-sided Bayesian paired samples t-test indicated very strong evidence in favour of the alternative hypothesis (H_1)(BF_{10} = 46.5, *d* = 0.56 [95% CI = 0.2 to 0.92]) for improvements in total goals scored between trial 1: 13.6 (3.3) and trial 2: 16 (3.3). Further, the Bayesian paired-samples equivalence t-tests indicated strong evidence in favour of H_1 (BF_{10} = 10.2, *d* = 0.24 [95% CI = -0.09 to 0.57]) for equivalence between trial 4: 16.7 (3.7) and trial 5: 18.2 (4.7); extreme evidence in favour of H_1 (BF_{10} = 132, *d* = -0.02 [95% CI = -0.34 to 0.30]) for equivalence between trials 5 and 6: 18.1 (3.5); and moderate evidence in favour of H_1 (BF_{10} = 8.4, *d* = 0.26 [95% CI = -0.08 to 0.59]) for equivalence between trials 4 and 6. Sufficient evidence indicated that a learning effect took place between the first two trials, and that up to five trials might be necessary for performance to plateau in a specific virtual reality soccer passing task.

The second research study (Chapter 4) aimed to assess the validity of a soccer passing task by comparing passing ability between virtual reality and real-world conditions. A previously validated soccer passing test was replicated into a virtual reality environment. Twenty-nine soccer players participated in the study which required them to complete as many passes as possible between two rebound boards within 45 s. Counterbalancing determined the condition order, and then for each condition, participants completed four familiarisation trials and two recorded trials, with the best score being used for analysis. Sense of presence and fidelity were also assessed via questionnaires to understand how representative the virtual environments were compared to the real-world. Results showed that between conditions a difference was observed (EMM = -3.9, 95% HDI = -5.1 to -2.7) with the number of passes being greater in the real-world (EMM = 19.7, 95% HDI = 18.6 to 20.7) than in virtual reality (EMM = 15.7, 95% HDI = 14.7 to 16.8). Further, several subjective differences for fidelity between the two conditions were reported, notably the ability to control the ball in virtual reality which was suggested to have been more difficult than in the real-world.

The last research study (Chapter 5) aimed to compare and quantify the perceptions of virtual reality use in soccer, and to model behavioural intentions to use this technology. This study surveyed the perceptions of coaches, support staff, and players in relation to their knowledge, expectations, influences, and barriers of using virtual reality via an internet-based questionnaire. To model behavioural intention, modified questions and constructs from the Unified Theory of Acceptance and Use of Technology were used, and the model was analysed through partial least squares structural equation modelling. Respondents represented coaches and support staff (n = 134) and players (n = 64). All respondents generally agreed that virtual reality should be used to improve tactical awareness and cognition, with its use primarily in performance analysis and rehabilitation settings. Generally, coaches and support staff agreed that monetary cost, coach buy-in and limited evidence base were barriers towards its use. In a sub-sample of coaches and support staff without access to virtual reality (n = 123), performance expectancy was the strongest construct in explaining behavioural intention to use virtual reality, followed by facilitating conditions (i.e., barriers) construct which had a negative association with behavioural intention.

This thesis aimed to explore the measurement properties of soccer specific virtual reality tests, and the perceptions of staff and players who might use the technology. The key findings from exploring the measurement properties were (1) evidence of a learning curve, suggesting the need for multiple familiarisation trials before collecting data, and (2) a lack of evidence to support the validity of a virtual reality soccer passing test as evident by a lack of agreement to a real-world equivalent. This finding raises questions on the suitability for virtual reality being used to measure passing skill related performance. The key findings from investigating the perceptions of users included, using the technology to improve cognition and tactical awareness, and using it in rehabilitation and performance analysis settings. Future intention to use was generally positive, and driven by performance related factors, yet several barriers exist that may prevent its widespread use.

In Chapter 7 of the thesis, a reflective account is presented for the reader, detailing some of the interactions made with coaches, support staff and players in relation to the personal, moral, and

ethical challenges faced as a practitioner-researcher, working and studying, respectively, in a professional soccer club.

Contents

List of Figures

List of Tables

Abbreviations, Acronyms and Symbols

3D	Three dimensional
Bca	Bias-corrected and accelerated
BFs	Bayes factors
CB-SEM	Covariance based structural equation modelling
CI	Confidence interval
d	Cohen’s d effect size
EMM	Estimated marginal means
F-MARC	FIFA Medical Assessment and Research Centre Battery
GPS	Global position systems
HDI	Highest density interval
ICC	Intra-class correlation coefficients
IMU	Inertial measurement units
ITC-SOPI	International Television Commission Sense of Presence Inventory
LOO	Leave-One-Out Cross-Validation
LSPT	Loughborough Soccer Passing Test
PLS-SEM	Partial least squares structural equation modelling
PSIS	Pareto Smoothed Importance Sampling
r^2	Explained variance
ROPE	Region of practical equivalence
SD	Standard deviation
SIM-TLX	Simulator task load index

TRI 2.0	Technology readiness index 2.0
UQFSA	University of Queensland Football Skill Assessment
UTAUT	Unified Theory of Acceptance and Use of Technology
VIF	Variance inflation factor
VO_2	Oxygen consumption
VR	Virtual reality
β	Beta path coefficient

External funding

I would like to acknowledge and thank Hull City Football Club for providing the stipend to support these studies.

Chapter 1 General Introduction

1.1 Introduction

Virtual reality (VR) has been defined as a computer-simulated environment that aims to induce a sense of being mentally and/or physically present in another place (Baños et al., 2000; Neumann et al., 2018). The technology uses both hardware and software components, and usually incorporates a head-mounted display that obstructs the user's vision of the real world (Slater & Sanchez-Vives, 2016). Inside the head mounted displays, digital images of a virtual environment are presented to the user, and sensors encased in the headset track head movements and alter the virtual environment accordingly (Bird, 2020). While these components are an essential part of VR and contribute to the identification and understanding of VR, its definition is mostly influenced by the experience it provides. Further, VR allows for real time interaction with the computer-simulated environment through an individual's own natural senses and skills (Mérienne, 2017; Miles et al., 2012). For these reasons, VR has been applied in industries such as medical surgery (Marescaux et al., 1998; Satava, 1993) and aviation (Neretin et al., 2021) since the 1990s to provide a controlled and safe environment to train personnel in complex situations. However, during that time, such systems were expensive (up to $200,000 USD), difficult to operate, and presented poor quality graphics, but were justifiable based on the benefits outweighing the costs (e.g., patient life) (Berkman, 2018; Marescaux et al., 1998). As such, VR remained a technology inaccessible to many organisations. It was only in 2016 that VR began to see exponential growth due to the new generation of devices from the likes of Oculus (now rebranded and owned by Meta) and HTC being released that year (Bird, 2020). In comparison to their predecessors, these devices were more powerful, cheaper to manufacture, more intuitive, and with improved graphics. Further, these devices had become smaller, and therefore more ergonomic, which is part explained by the observed doubling of component numbers per chip every 1 – 2 years, which allows smaller microchips to be produced that have the same processing capacity as their predecessors (Burg & Ausubel, 2021). This observation is typically referred to as "Moore's Law" (Moore, 2006). All of these improvements led to a rapid increase in sales of VR devices, that are projected to be in excess of 21.6 million as of 2021, and an estimated 65.9 million devices by 2026 (Ballhaus et al., 2022). With improved accessibility came opportunity for increased revenue across multiple industries. In the entertainment and media sector alone, VR sales increased by 36.5% in 2021 from 2020, to reach a value of US$2.6bn, and is expected to reach US$7.6bn by 2026 (Ballhaus et al., 2022). With increased accessibility and growth, new industries such as sport have begun to explore the use of VR for training.

Although VR has only recently been adopted by professional soccer teams, coaches and sport scientists have been pondering its use for much longer. In 2013, The Guardian published an article titled ‘Virtual reality simulators could end England’s penalty shootout woe’ (Riach, 2013). In the article, it was announced that British Aerospace systems and UK Sport had partnered to develop VR technology for Great British athletes at the 2016 Rio de Janeiro Olympics. Discussing how the technology could be implemented, Dr Scott Drawer, former Head of Research and Innovation at UK Sport, suggested that VR could replicate the stressful environment of penalty shoot-outs at international competitive events such as the football World Cup, allowing players to better prepare themselves for the realities of competition (Riach, 2013). Players could kick a ‘virtual’ ball and experience some degree of stress or anxiety of a large crowd, a feat that is hard to achieve in the training ground environment. While no evidence exists for VR being used for that specific purpose, the statement holds some truth as research has demonstrated that VR sports environments can be manipulated to induce symptoms of acute anxiety (Kelly et al., 2022). To that end, VR allows environments and scenarios to be created that are impractical, or simply not possible to replicate in the real world. For instance, goalkeepers can be tested in free-kick situations where the ball’s end-placement, trajectory, and speed are fully controlled, allowing multiple players to be tested under the same condition (Dessing & Craig, 2010). Similarly, outfield players could be assessed on their decision-making ability (Wood et al., 2021), or characteristics that are closely associated with decision making such as their visual exploratory activity (Wirth et al., 2021). In such a case, VR would allow for soccer scenarios to be controlled so that team-mates and opposition player movement patterns during a pattern of play are replicated precisely between tests.

Coinciding with professional soccer clubs taking an interest in VR, several companies specialising in VR solutions for professional soccer have entered the market (Thatcher et al., 2021). Notably, VR companies are also promoting their capabilities at testing the characteristics of soccer players. For instance companies such as Rezzil (Rezzil, 2020) and BeYourBest (Be Your Best, 2020) claim to measure and benchmark decision making and scanning abilities of outfield players, whereas Incisiv (Incisiv, 2020) suggest they can profile the perceptual abilities of goalkeepers (Thatcher et al., 2021). However, evidence to support the validity of such claims remains absent. For clubs looking to use VR for these kinds of applications, it is important that coaches and backroom staff are confident in the quality of the data so that the results can be used to inform decision making. In assessing the quality of the data, researchers tend to explore the measurement properties of a test, and it is recommended that this is done before the test is introduced into practice (Mokkink et al., 2010). One measurement property that is often assessed (and then implemented) before a test is the assessment of familiarity. Familiarity is a state of being accustomed to an action or use of a device such that any substantial learning effect has been removed prior to subsequent use (Hopkins, 2000), and is typically achieved through practice (also referred to as familiarisation) sessions (Robertson et al., 2017). Reducing or removing these learning or reactivity effects is important if VR is going to be used in repeated trials (e.g., as part of an intervention) as it ensures that any changes in the characteristic being measured is a result of the intervention and not learning how to complete the test better or more efficiently (Hopkins, 2000; Robertson et al., 2017). Familiarity is also unique in VR, as players must be accustomed not only to the specifications of the test, but also the novelty of the VR device (e.g., wearing a headset, interacting with the environment). From a pragmatic view, practitioners working in professional soccer are also restricted by the time available to them. While ensuring the quality of the data is important, it could come at the cost of time if lengthy familiarisation periods are required. At present, no study has explored the familiarisation requirements of a soccer VR test.

Another of these measurement properties is the assessment of validity, defined as the degree to which an assessment measures what it intends to measure (Nortwick et al., 2009). Ultimately, practitioners need to be confident that the characteristic that is being measured by the VR test is an accurate representation of that characteristic in the real world. This is especially important if the test is to be used to inform decision making, as decisions could be made based on the test result that led to consequences for the player, such as under or over allocation of training, or missed opportunities in selection/talent-identification purposes. To validate VR in sports such as soccer, researchers have been influenced by methods used to validate VR surgical training applications, by applying the known-group difference techniques (Wood et al., 2021). This method of construct validation aims to demonstrate validity by assessing two or more groups of participants on the same test, with the expectation that differences will emerge in the variable of interest (Impellizzeri & Marcora, 2009). For instance, because experienced surgeons are expected to perform better in a surgical situation than a junior surgeon in the real world, a valid VR system should demonstrate differences in the variable of interest during a VR surgical task. However, although this method of validation is necessary, it is not sufficient for evaluation of performance, and is only applicable for discriminative purposes (Impellizzeri & Marcora, 2009). For instance, it does not allow evaluation of performance between the same group of players (e.g., professional players – where this technology is being marketed). To which, a criterion method such as concurrent validity could be more appropriate, where real-world performance is measured against performance in VR. The reason for its underutilisation in surgical VR applications is unknown but is possibly due to the difficulty in comparing performance in VR to that of the real-world where it would have been difficult to control the real-world environment (i.e., due to variation between patient anatomy, complexity of procedure), and unethical due to the risk of patient health, especially given VR surgical testing is aimed at novice and inexperienced surgeons. To date, no study has explored the concurrent validity of VR in soccer.

Understanding the measurement properties of a test such as familiarity and validity are important and contribute to the trustworthiness of the data. However, irrespective of these properties, a technology might fail in the soccer environment if implementation problems occur (Windt et al., 2020). The diffusion of innovation theory suggests that the rate of technology adoption depends more on the end-users' subjective perceptions of usability, complexity, and observability of the technology rather than objective evidence of the technology's efficacy (Donaldson et al., 2012; Rogers, 1995). Even technologies that adhere to the gold standards in data quality might struggle in the soccer environment. For instance, portable metabolic gas analysis systems used to measure physiological variables of aerobic endurance such as oxygen consumption are rarely used in professional soccer despite being capable of measuring a player's physiological response to training (Stølen et al., 2005). Instead, intermittent running tests such as the Yo-Yo intermittent recovery test (Stølen et al., 2005) or the 30-15 (Buchheit, 2008) are used as practical alternatives for testing large groups at the same time, are inexpensive, and require less expertise to implement and analyse (Grgic et al., 2021). However, the latter might also be seen as a limitation, as it could mean that those implementing the test are not aware of its limitations when compared to the gold standard, or worse, do not understand why it's being deployed in the first place. In contrast, some technologies might succeed in the soccer environment, even when the evidence for efficacy is not compelling. For example, the use of cryotherapy chambers for recovery has contrasting evidence for its benefit and costs a significant amount of money to install and operate, yet these chambers are used in professional soccer (Adnan et al., 2018). The reasons for soccer clubs adopting VR could be multifaceted, such as gaining a competitive advantage, or searching for 'marginal gains' in performance. Equally, social factors such as technology being used as a tool to win favour over upper management, by encouraging enthusiasm, hope and belief in players and stakeholders (Windt et al., 2020), could be a reason for adopting VR. Further, soccer clubs might not be fully aware of the power of the placebo, meaning they are excited about buying something new without full consideration of potential limitations that the technology could bring (Windt et al., 2020). However, these factors are just speculative, as it is not clear what mechanistic factors are influencing VR use (or desire for use) in soccer. Exploring these factors through the perceptions of key stakeholders involved in the use of VR in soccer is required, which might not only improve our understanding of why VR is being adopted but will allow teams to invest in VR for the appropriate reasons (i.e., on the belief it might improve aspects of soccer performance, versus socially or emotionally orientated influences). Further, understanding the perceptions of VR in soccer might allow for key stakeholders responsible for implementing VR into the soccer club to identify appropriate areas for its use, whereas strategies can be implemented to alleviate barriers towards its adoption. To date only two studies have explored the perceptions of VR use in soccer. In the most recent study, Mascret, Montagne, Devrièse-Sence, Vu, and Kulpa, (2022), surveyed a large sample of athletes from multiple sports and competitive levels, including soccer, on the acceptance of VR to aid them in improving sport performance. The results revealed athletes perceived VR to be useful (except for recreational athletes) and had an intention to use VR in future. Recreational athletes not finding VR useful might be explained by the primary reasons they are taking part in sport (e.g., social, physical, and mental benefits) which differ to competitive athletes, who are focused on the desire to compete, and win competition. However, because the study used a variety of sports, and a variety of competitive levels, the results cannot be generalised to professional soccer (Mascret et al., 2022). Further, because sports 'performance' is a multifactorial construct, consisting of multiple components (e.g., physical, cognitive, and tactical), it is unclear what component(s) are perceived as being useful to improve through VR, and which components are not. In the second study, Thatcher et al. (2021) did explore the perceptions of those working in professional soccer, and the areas in which the technology was perceived as being most useful (e.g., player rehabilitation, and performance analysis). However, the results were based on a small sample (n = 6) which is problematic in perceptual based studies, as sources of bias might exist from small samples due to them not representing the target population. Furthermore, the sample only reflected opinions of technical coaches and performance analysis staff, not the multidisciplinary team (i.e., soccer players, sport scientists, rehabilitation, and medical staff) that professional soccer clubs now employ.

In summary, VR has taken an exponential rise over the last decade across a multitude of industries, including professional sport. Virtual reality's capability to produce controlled and repeatable scenarios, often impractical or impossible in the real world, has resulted in numerous use-cases for it to be used when testing soccer players. However, for VR to effectively assess soccer players and inform decision making, it is important that the measurement properties of these systems are adequately assessed, such as familiarisation and validity. Furthermore, if VR is to be used in professional soccer, it is important that the perceptions of those working in the industry are explored to identify key areas of use, in and adoption barriers.

1.2 Aims of the thesis

The general aim of this thesis was to increase the understanding of VR being used in the professional soccer training ground environment. To do so, this thesis examined the fundamental measurement properties of soccer-specific VR tests as well as the perceptions of VR use in professional soccer. The specific aims of this thesis were:

- To examine the familiarisation properties of a soccer-specific VR passing test.

- To examine the concurrent validity of a soccer-specific VR passing test compared to the real-world version of the test.
- To explore the perceptions and behavioural intentions to use VR in professional soccer coaches, support staff, and players.

Chapter 2 **Literature review**

2.1 Aims

The aim of this chapter was to provide background information, and where appropriate, a critical analysis of the research that forms the key components of this thesis, being: familiarisation properties and the validation of VR tests in soccer, and the user perceptions of VR in professional soccer. This chapter also provides a broader perspective of VR use outside of soccer, to include other sports and professions.

2.2 Demands of Soccer

Coaches and sport scientists require a thorough knowledge of the demands that a soccer player faces during competition, and how these demands relate to the outcomes of a match. This is important if coaches and sport scientists aim to optimise the training, testing, and monitoring of soccer players as to influence match competition. For example, during a typical 90 minute game, the professional soccer player will cover approximately 10-12 km, with a mean intensity close to the anaerobic threshold, and actions that consist of accelerations and decelerations, changes in direction, sprinting and jumping (Stølen et al., 2005). Knowledge of these physical demands allow for training strategies to be implemented, such as altered training programs that aim to evoke a specific adaptation (e.g., small-sided games to improve intensive period of play) (Stølen et al., 2005). Alternatively, knowledge of the physical demands might result in a reduced training volume for certain players, particularly for those that have experienced a spike in volume and intensity (e.g., following a congested fixture period) that unacknowledged could lead to injury (Bowen et al., 2019). Physical demands are not the only component the sport scientist and coach must consider if they are to influence competition outcomes. Players must perform a range of technical skills such as dribbling, passing, and shooting (Russell et al., 2013), that are all developed at a young age and refined throughout a players career. Cognitive demands are also present throughout soccer competition, as players are required to make instantaneous and continuous decisions throughout a game based on the information and resources that are made available to them (Pruna & Bahdur, 2016). Finally, players must also learn and implement individual and team tactics based on the methodology of the club or manager, which is frequently adapted depending on the different phases or state of the game (e.g., winning or losing), the tactics of the opposition team, and player availability (Rein & Memmert, 2016). Historically, coaches and sport scientists have relied on notational analysis, previous experiences, and discussions with peers to understand these demands (Barris & Button, 2008). For instance, the number of passes completed during competition might rely on individual coding by performance analysis staff, whereas previous experience from coaches that once played at a professional level, might provide an understanding of what it ‘felt like’ physically to compete at such a level. However, notational analysis is time consuming for the practitioner, and the reliability can vary depending on the number of observers, their experience and their viewing perspective, leading to a lack of precision (Barris & Button, 2008; Windt et al., 2020). Further, while prior experience (or intuition) contributes towards the understanding of these demands, they are subjective and might be influenced by bias. For these reasons, amongst others, coaches and sport scientists have broadened their technology scope, to embrace a wider array of technologies.

2.2.1 Technology in soccer

Technology, in accordance with the Cambridge dictionary, is defined as the practical, especially industrial, use of scientific discoveries (Cambridge Dictionary, n.d.). In reality, the definition of technology is far more complex as evidenced in detail by Carroll (2017), who also devised a framework on defining technology. The definition of technology by Carroll (2017) includes something that is always inherently intelligent enough to function, to be imbued with, or be interpreted as having a function, that only intelligent beings (human or otherwise) can comprehend. Further, technology is defined as something devised, designed, or discovered that serves a particular purpose, and provides (or is intended to provide) some form of benefit (Carroll, 2017). Digital technology, therefore, takes on the definition that is previously outlined, but applied solely towards electronical devices, tools, and systems, that generate, store, and/or process digital data. For this thesis, ‘digital technology’ is simply referred to as ‘technology’, throughout.

In an attempt to improve the physical, technical, cognitive and tactical components of soccer, teams and researchers have used many technologies in competition and training, for numerous reasons (Almulla et al., 2020). First, compared to human observation and manual data entry (e.g., the initial stage of notational analysis), technology can provide data that is more reliable, and made available much faster or in real-time (Barris & Button, 2008). For the practitioner this means more confidence in the data, along with more time availability (if used wisely) to delve into the data and inform practice more thoroughly (Windt et al., 2020). Examples of technologies that have allowed this transition include, global positioning systems (GPS); inertial measurement units (IMU), and optical tracking systems that capture the physical responses during competition and training (e.g., total distance covered, time spent in different speed zones) which are then used in the monitoring of the acute and chronic volumes and intensities (Theodoropoulos et al., 2020) for injury prevention (Bowen et al., 2019), optimising the return to play (Taberner, Van Dyk, et al., 2019), and benchmarking players (Goto et al., 2015). Technology has also enabled insights to be made, that are simply not possible from observations alone. For instance, technologies such as force plates (Brownlee et al., 2018) and heart rate tracking (Alexandre et al., 2012) have been used to measure the recovery status of players, with the collected data used to provide recovery interventions or monitoring strategies (e.g., by comparing data to baseline conditions). Technology has also become prevalent for measuring technical characteristics of soccer players, where teams have often relied on semi-automated multiple camera tracking systems to quantify technical actions such as number of passes, crosses, and shots (Russell et al., 2013). However, such systems require complex and expensive infrastructure, limiting their use to teams with the financial capabilities to invest in the technology. Even then, limitations exist on their transferability to the training ground environment, where multiple pitches are likely to be used at any given time. As such, foot mounted IMU technology has addressed this gap, providing a transferrable and inexpensive solution for measuring technical actions in soccer (Marris et al., 2022; Myhill et al., 2023). That said, in contrast to the forementioned camera-based system, these foot-mounted IMUs can not provide potentially important information such as if the action (e.g., pass, cross) was the best decision at the time, based on information that was available to the player (e.g., location of team-mates, phase of the game). More recently, technologies that purport to track and improve cognitive demands, such as vision, memory, attention, and decision making have become more common in soccer (Renshaw et al., 2019). Examples include visual-motor reaction training which require players to quickly make motor responses to visual information, usually in the form of an illuminated button (Appelbaum & Erickson, 2018), and stroboscopic glasses, that work by temporarily disrupting the vision of the player, forcing them to more effectively use the limited visual input that is available to them (Appelbaum & Erickson, 2018). However, the efficacy of cognitive technologies has recently been criticised. For instance, three dimensional (3D) multiple object tracking technologies that require athletes to fixate on a single object and track the movement of other objects through peripheral vision, have been shown to have weak evidence for far transfer, meaning, that although performance improves in the 3D multiple object tracking task, performance is unlikely to improve in the sporting task (Vater et al., 2021). Further, studies that have shown evidence for efficacy have been questioned based on methodological concerns such as small sample sizes, inappropriate statistical analyses, and frequent publications from authors connected directly to the technology companies (Harris et al., 2018; Vater et al., 2021). The latter, while not a methodological concern in itself, might lead to researcher bias, which could be minimised if open science protocols such as open access data and pre-registrations were used, which to date no study has done in this field (Vater et al., 2021). Finally, technology has also shaped how team and player tactics are understood and implemented. For instance, technologies that manage and process big data and machine learning are become more prevalent in soccer, providing a more comprehensive and efficient understanding of the game (Rein & Memmert, 2016). Therefore, although a technology might have numerous benefits, there are often obstacles that the coach and sport scientist must consider before its implementation, which is a challenge especially given the rate of new technologies being introduced into the soccer market. One of these technologies is VR.

2.3 Virtual reality

Virtual reality was officially termed in the early 1990s, with definitions provided by several authors. For example, Bishop and Fuchs (1992) define VR as real-time interactive graphics with three-dimensional models, when combined with display technology that give users immersion in the model world and direct manipulation. Gigante (1993) defined VR as the illusion of participation in a synthetic environment rather than observation of such environment. Virtual reality relies on 3D, stereoscopic head-tracking displays, hand/body tracking, and binaural sound. Virtual reality is an immersive, multi-sensory experience. Elsewhere, Baños et al. (2000) provided a more simple definition of VR, emphasising it more as an experience as opposed to a technology: a computer-simulated environment that aims to induce a sense of being mentally and/or physically present in another place. In essence, all these definitions of VR are true to some degree, as it is generally agreed that VR consists of a selection of the following: an ability to interact and manipulate a digital environment, sensory inputs (e.g., sight, touch, sound), and reproduction of perceptual information and behavioural constraints found in the real-world environment (Slater & Sanchez-Vives, 2016). The latter is especially important, as it allows VR simulations to be used for training purposes in high-risk industries, such as surgery, aviation, and military, at a reduced risk of injury and cost (Buckley et al., 2012; Le Noury et al., 2022). Virtual reality can be presented through different hardware solutions, including display screens, Cave Automatic Virtual Environments, often referred to simply as a ‘cave’ system (Cruz-Neira et al., 1992), and head-mounted displays. However, the type of hardware used is likely to affect the level of immersion and presence, features that are discussed in the next section.

2.3.1 Features of virtual reality

Virtual reality consists of hardware (type of display, auditory and haptic system etc), and software (computer graphics, real-time computing, etc) used to immerse the user in a virtual environment (Mérienne, 2017; Slater & Sanchez-Vives, 2016). The type of hardware and software used will largely determine the level of immersion, which is an objective level of sensory fidelity a VR system provides (Bowman & McMahan, 2007). Virtual reality systems that provide more information (across all sensory modalities) and tracking capabilities (that match the users proprioceptive feedback about body movements) are likely to be more immersive, as they preserve fidelity in relation to equivalent sensory modalities in the real-world (Slater &

Wilbur, 1997). Whereas immersion is an objective and quantifiable description of what a VR system provides, presence can be seen as a subjective correlate, and is defined as a state of consciousness, and the psychological sense of being in the virtual environment (Slater & Wilbur, 1997).

2.3.2 Immersion

The term ‘immersion’ appears to have originated in 1995, and was used as a description of the technology as opposed to psychological characteristics experienced by the user as a result of using a VR system (Slater et al., 1995). As described by Slater et al. (1995) immersion includes the extent that display systems (VR systems) are: inclusive; extensive; surrounding; vivid; and matching. Inclusive refers to the extent that systems block out sensory information from the real world, whereas extensive refers to the sensory systems that the system is able to accommodate (Slater et al., 1995). Immersion can therefore be seen as greater in systems that substitute real sensory information (i.e., vision, auditory, touch) with computer generated ones (Slater & Sanchez-Vives, 2016). Surrounding refers to the extent that information can arrive at the user from any given direction, and that the user is able to turn (physically) towards any direction yet still remain in the virtual environment (Slater et al., 1995). Immersion can therefore be seen as greater in a system that offers a panoramic view as opposed to a narrow field, for example head mounted displays compared to two-dimensional (2D) display screens, respectively (Slater & Wilbur, 1997). Vivid refers to the resolution, richness, information content, and overall quality of the system displays (Slater et al., 1995; Steuer, 1992). Matching is the final aspect of immersion, and refers to the system’s ability to match the proprioceptive feedback regarding body movements, and the information that is displayed to the user (Slater et al., 1995).

However, not all authors have viewed immersion as an objective description of the VR system (Witmer & Singer, 1998). Instead, Witmer and Singer (1998) defined immersion as a psychological state characterised as perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences. While Witmer and Singer (1998) agree that equipment configuration is essential for allowing immersion, they perceive immersion to be something that a person experiences. As a result of the disagreement in terminology, Slater (1999) responded by suggesting the extension of the term immersion to include either ‘system immersion’ in reference to Slater et al. (1995), and ‘immersive response’ to reference Witmer and Singer (1998). Unfortunately, the expanded definitions of Slater (1999) made minimal impact within the literature, with many authors referencing immersion without indication of the terminology they are referring to. Furthermore, some authors have referred to immersion when what they are actually referring to is ‘presence’.

2.3.2.1 Measuring immersion

Despite the definitions of system immersion and immersive response being introduced by Slater (1999), no unified measure exists for either of these definitions. For system immersion, authors have previously categorised the level of immersion into high, moderate, and low levels depending on the aspects of immersion that was introduced (Slater et al., 1995). For instance, for the 'extensive' aspect of immersion, Miller and Bugnariu (2016) categorised a low VR system as accommodating one sensory modality; moderate for accommodating 1-2 sensory stimuli that may or may not be spatially orientated; and high for accommodating two or more sensory modalities that are spatially orientated. By categorising the level of system immersion that a VR system offers, authors have been able to explore how immersion (by means of the hardware and software) effects task performance (Rose et al., 2018; Shen et al., 2021), and enjoyment (Madden et al., 2020). As for immersive response, a paucity of literature is available regarding the measurement of this construct. This is perhaps due to the similarity that the definition shares with the definition of 'presence' (Slater, 1999). However, the authors who defined immersion (immersive response) did develop the immersive tendencies questionnaire, a psychological measure of an individual's tendency to experience presence (Witmer & Singer, 1998). Immersive response, and its shared definition, presence, are subsequently discussed.

2.3.3 Presence

Presence can be defined as a person's illusion of being in the scene displayed by the VR system (Slater et al., 2010), and is often referred to as a psychological sense of 'being there' (Slater et al., 1995; Steuer, 1992). The concept of presence comes from the term 'telepresence', coined in the 1980s, which is defined as 'when a person is objectively present in a real environment that is physically separate from the person in space' (Schloerb, 1995). It was then in the 1990s that this term was adapted for VR and referred to as 'presence', so instead of having the sense of being at the remote physical environment, the user has a sense of being in a virtual environment (Held & Durlach, 1992; Sheridan, 1992). To further refine the terminology of 'presence', Slater (2009) introduced two concepts, the first being 'place illusion' for the type of presence that refers to the sense of 'being there', despite knowing that you are not. The second concept is known as 'plausibility illusion' which refers to illusion that what is apparently happening in the virtual environment is actually happening, and that the events within the virtual environment respond to you and correlate with your own actions (Slater, 2009; Slater & Sanchez-Vives, 2016). For instance, Slater and Sanchez-Vives, (2016) use the analogy of a virtual avatar approaching the VR user and smiling, resulting in the user smiling back despite knowing (even if subconsciously) that no one is there.

It is commonly agreed among those interested in VR, that one of the main goals of VR is the ability to promote a sense of presence (Grassini & Laumann, 2020; Steuer, 1992). In fact, the ability to induce presence is what distinguishes VR from other forms of media (Slater & Sanchez-Vives, 2016). Presence is therefore an important area of research. For instance, understanding factors related to presence and how these factors differ between immersive environments and exocentric displays (such as computer screens) is important, as the findings guide the technologies development (Slater & Wilbur, 1997). However, the importance of studying presence goes beyond improving developments in the technology, as it has been studied in direct relation to applicable research (Grassini et al., 2020). For instance, Slater and Wilbur (1997) state that understanding presence is important because with greater levels of presence comes more opportunity for users to behave physically and emotionally as they would in the real-world, which is important for training tasks so that behaviours created/learned in VR transfer to the real-world. However, that is not to say that higher levels of presence correlates with training transfer, to which research has been inconclusive (Stevens & Kincaid, 2015). For example, while Grassini et al. (2020) found positive correlations between presence and model airplane building performance (quality assessment, errors in production, and speed of build) (Rho = 0.917, $p < 0.001$), whereas others, such as Persky et al. (2009) found that no correlations existed between presence and learning comprehension outcomes following an active and didactic virtual learning environment. Irrespective of these findings, others have argued that a causal direction between presence and performance might not ever be determined as they are mutually reliant on one another (Maneuvrier et al., 2020). For instance, take a fictious VR ball catching test - similar to that of Zaal and Bootsma (2011). Assuming the VR environment is designed in such a way that it relies on one's sensorimotor skills as they would use them in the real-world (i.e., judgement of ball speed, depth-perception, etc), more balls caught would relate to enhanced task performance. However, the more balls caught could result in an awareness of one's sensorimotor skills in VR and how they compare to the real-world, enhancing presence. Therefore, whilst it is possible that presence could promote performance, its equally argued that performance could promote presence.

2.3.3.1 Measuring presence

To date, several methods of assessing presence exist within four categories: questionnaires; physiological measures; assessing user behaviours; and interviews (Nichols et al., 2000). The most common method of assessment are questionnaires whereby users of VR provide a post-test subjective response to the VR experience, usually by answering questions on a Likert scale (Souza et al., 2021). Widley applied presence questionnaires include the Slater, Usoh, and Steed questionnaire (Usoh et al., 2000), the Presence Questionnaire (Witmer & Singer, 1998), and the International Television Commission Sense of Presence Inventory (ITC-SOPI) (Lessiter et al., 2001). The popularity of questionnaires derives from the minimal cost, and the administrational, analytical, and interpretational ease, yet are limited by the reliance on users' memory which can be incomplete, biased, and influenced by events towards the end of the VR experience (van Baren & IJsselsteijn, 2004). As such, it is recommended that multiple methods of assessing presence are used alongside subjective questionnaires (van Baren & IJsselsteijn, 2004)

2.3.4 Cybersickness

Cybersickness, also referred to as motion sickness or simulator sickness in a virtual environment, is a side effect associated with VR use (LaViola, 2000; Yildirim, 2020). Although symptoms of cybersickness can vary, commonly reported symptoms include nausea, dizziness, stomach discomfort, and vomiting (LaViola, 2000). While the symptoms of cybersickness are generally agreed on, it is still not clear why cybersickness occurs. Several theories have been suggested to explain why cybersickness occurs such as the poison theory, postural instability theory, and the sensory mismatch theory (Yildirim, 2020). The latter theory is regarded as the most accepted and suggests that the cause of cybersickness is a conflict between the visual and vestibular systems during exposure to motion in virtual environments (Yildirim, 2020). That is, it appears (visually) that motion is taking place, when in reality (vestibular) motion is not. While this theory has been supported, it may not fully explain cybersickness occurrence, with research showing that human factors such as age and gender (Chang et al., 2020), display type (Yildirim, 2020), and task (Chang et al., 2020) may all contribute. As such, the variety of explanations for cybersickness occurrence means that it is possible that different virtual environments will impose different levels of cybersickness to the user.

The prevalence of cybersickness when using VR is problematic. First, it is without question that the prevalence of cybersickness is an ethical issue for participant safety. However, beyond the wellbeing of participants, cybersickness prevalence has been shown to negatively impact perceived presence within the virtual environment (Weech et al., 2019), which may influence engagement with the task. Additionally, it has been indicated that cybersickness may have an adverse effect on cognitive function such as reaction time (Nalivaiko et al., 2015). For this reason, studies have investigated subjective measures of cybersickness, using questionnaires such as the nausea profile (Muth et al., 1996), the motion sickness assessment questionnaire (Gianaros et al., 2001) and the most commonly cited, simulator sickness questionnaire (Chang et al., 2020; Kennedy et al., 1993). Although these questionnaires have been used successfully in understanding the prevalence of cybersickness, they are structurally limited to use post-test, and can not be administrated during the test due to their length and complexity (Keshavarz & Hecht, 2011). This comes at a disadvantage, as it means that onset of cybersickness occurrence can’t be investigated. For this reason, Keshavarz and Hecht (2011) developed a unidimensional measure of cybersickness known as the fast motion sickness scale. The fast motion sickness scale has been validated against the simulator sickness questionnaire, and requires participants to rate perceived levels of nausea, stomach discomfort, and general discomfort on a 0-20 scale. That said, given the clinical likeness between motion sickness and cybersickness where moderate relationships have been shown between nausea ratings (r = 0.58) (Gavgani et al., 2018), it might be possible to predict cybersickness occurrence in those who are susceptible to symptoms of motion sickness. However, further studies are required to confirm this hypothesis.

Several studies that have used VR in sport have referred to cybersickness when introducing the virtual environment to the participants for the first time (Michalski et al., 2019; Pagé et al., 2019). Formally known as familiarisation (discussed further in section 2.5.1), studies have indicated the importance of customising participants to potential adverse effects of VR use. For instance, ahead of prescribing participants to a VR table tennis program, Michalski et al. (2019) included a 30 minute familiarisation to allow participants opportunity to see if they could tolerate the VR exposure. However, the authors did not attempt to measure cybersickness in the study, which is problematic because despite it appearing that no participant had to withdraw from the study due cybersickness, it is possible that cybersickness still existed but at tolerable levels. If so, participants might have been unable to attain the most out of the training intervention, which may have impeded the results of the study. To date, only one study has quantified cybersickness prevalence during a VR sport task, where a VR tennis simulation was found to illicit low levels of cybersickness (mean = 1.26, SD = 0.45, 95% CI [1.02, 1.50]) where 1 equals the lowest possible score (no cybersickness) and 5 equals the highest possible score (evidence of cybersickness) (Le Noury et al., 2021).

2.3.5 The use of virtual reality in soccer

In professional soccer, VR might have several potential use cases. Virtual reality could be used for non-interactive (with the virtual environment) observational learning such as during performance analysis sessions, by allowing players and coaches to review simulated match footage from their own perspective, which could have benefits over conventional video analysis (Craig, 2014; Düking et al., 2018). For instance, Pagé et al. (2019) reported that VR training (viewing basketball scenarios from a first person perspective) improved post-test basketball decision making accuracy in basketball players to a greater extent than computer screen training (viewing basketball scenarios on a monitor from a third persons perspective), and no training at all (control group). However, it is acknowledged that despite reporting a significant group x time interaction (*p* = 0.001), no post-hoc data was provided to show the magnitude of change from pre to post in any of the groups (i.e., the change in decision making accuracy in the VR training group). Instead, the author included a graph, and only reported non-significant differences between groups at pre intervention (*p* > 0.9; *d* < 0.41), and then reported the decision-making accuracy scores at post intervention for the VR (78.9%), computer screen (60.9%), and control (60.2%) groups, along with the effect size difference for the VR group compared to other groups (*d* = > 1.69). As such, while it appears that there is an effect of VR observational training on decision making ability (as suggested by the graph included in the paper), it is unclear what the magnitude of improvement is from pre-to-post when compared to other modalities. Similarly, Kittel, Larkin, Elsworthy, Lindsay, and Spittle, (2020) showed that decision making performance improved in Australian football umpires after a VR training intervention, however when compared to a conventional broadcast training intervention, the VR condition was not any more effective. However, the authors highlighted that the use of VR training was agreed as being more relevant (r = 0.71) and enjoyable (r = 0.56) than conventional broadcast training, as shown by large effect sizes between the two interventions. Yet, limitations exist surrounding how these perceptions were defined and measured. Firstly, unvalidated questions to measure relevance and enjoyment were adapted from previous research, so it is unclear if the questions measured what they intended to measure. For instance, the question of enjoyment: 'would you happily use this as a coaching tool', might be misleading as it is argued that a coach might happily use a technology (i.e., VR) but not necessarily enjoy using it, or in contrast, wouldn't use it as a coaching tool but thoroughly enjoyed it. Next, to measure these perceptions, the authors used a two-point scale, where participants had to provide a dash between 1 (low) and 10 (high), with the mean of each group used in the results. From these, statistically significant (*p* = 0.01) differences were shown between the VR and broadcast training groups for relevance (8.11 ± 1.42 vs 6.75 ± 1.84) and enjoyment (8.93 ± 1.09 vs 7.68 ± 1.72), and conclusions derived that included VR being more enjoyable and relevant. Yet, this does not allow for interpretation of what the data means in a pragmatic sense. For instance, for the perception 'enjoyment', enjoyment was found to be greater in VR by a mean difference of 1.25, but it is unclear if this difference is meaningful to infer that VR is more enjoyable.

Virtual reality might also be used in training sessions with soccer players to improve soccer related skills. Although studies have not explored soccer skill development through use of VR, several other sports including baseball (Gray, 2017), table tennis (Michalski et al., 2019; Oagaz et al., 2022), and darts (Tirp et al., 2015) have shown sport-specific skill improvement in the real-world as a result of VR training. Arguably, one of the most pivotal studies examining VR use to enhance sport skill in the real-world is that of Gray (2017), who evidenced the potential of VR training through its ability to provide a stimulus that is impractical or impossible to achieve in the real-world. In the study, competitive baseball players participated in a real-world baseball batting test, followed by being assigned to one of three training interventions (real-world, VR, or VR adaptive) or a control group that was conducted alongside their normal training, for a 6-week period. The real-world training consisted of extra baseball batting, whereas the VR training consisted of extra baseball batting within a virtual environment. The VR adaptive group also trained in a virtual environment, however, two design features were incorporated that ensured the difficulty level matched the performers skill level (challenge point hypothesis) and included variability in practice conditions. This included the speed of the ball and the pitch crossing location (vertical and horizontal), which were consistently matched to the ability of the participant. Such a feat would be possible, but impractical to replicate in the real-world. Post-test results showed that the VR and real-world training groups improved the number of balls hit to a similar extent when compared to the pre-test (*d* = 0.8), which was larger than experienced by the control group (*d* = 0.4), but much less than shown in the VR adaptive group (*d* = 1.9). Elsewhere, VR has also been shown to improve anticipation in cricket (Discombe et al., 2022), and improvements in rowing performance (Hoffmann et al., 2014; Parton & Neumann, 2019).

Another area that soccer clubs might use VR is through its potential to test and/or monitor soccer players. In this context, VR could be used in conjunction with other technologies or used entirely independently to capture data. Used in conjunction with other technologies, VR could be used to promote a biomechanical response or action that wants to be explored further. For instance, DiCesare, Kiefer, Bonnette, and Myer (2020), used VR to simulate a soccer corner-kick scenario where the participant had to jump and head a virtual ball into the goal. Kinematic data on landing at the hip, knee and ankle was captured through an optical motion camera system (used adjacently to the VR), analysed, and compared to landing kinematics captured following a standardised drop vertical jump (without using VR). Results showed differences between the two landing conditions, where the VR condition exhibited reduced sagittal plane range of motion at the ankle during the latter stages of landing when compared the standardised drop jump (*d* = 1.31), and differences at the ankle (*d* = 1.04) and hip (*d* = 0.84) at the fontal plane during the beginning and middle landing phase. It was concluded that VR-based assessments may provide a better indicator of biomechanical deficits than standardised jump assessments, that could predispose athletes lower-extremity injury (DiCesare et al., 2020). When used independently, VR can provide an immersive environment, while also recording data from the hardware and software, simultaneously. For instance, studies by Wirth et al., (2021) and Ferrer, Shishido, Kitahara, and Id (2020) used VR to measure visual exploratory activity (such as head turn frequency and excursion) of soccer players during a VR soccer match scenario through head-movement characteristics recorded by the head-mounted display. By comparison to the real-world, VR was used as it allowed for the task and environment to be standardised and repeated with precision, between participant tests. Whilst it would have been possible to repeat match scenarios in the real-world between participants, it is likely that variability in design factors such as player running speed, ball speed, and pass accuracy (from VR avatars) would have existed, which could have impacted the results of the study. Further, whilst researchers have previously examined the influence of visual exploratory activity in a competitive soccer match (Aksum et al., 2021), it is difficult to compare these behaviours between players due to contextual factors such as playing position, and team and opponent tactics that are shown to influence visual exploratory activity (Jordet et al., 2020). Both studies found that VR was able to distinguish between soccer players of differing levels of playing ability by comparing the differences in visual exploratory activity. For example, Wirth et al. (2021) found that when comparing lesser experienced soccer players to more experienced ones (categorised by playing leagues), they were shown to exhibit a higher frequency (9.81 ± 0.70 vs 8.93 ± 0.68) and excursion (66.72 ± 4.61 vs 58.94 ± 5.02 degrees) of head turns. This might have been explained by the more experienced group using a more efficient orientation of their bodies for action and observation of each scenario, requiring less dependence on head turn activity. While studies are yet to apply these findings into the applied soccer setting, it is possible that such insights could contribute to talent identification or training programs of soccer players in the future. However, both studies also acknowledged negative aspects of VR, that soccer clubs must be mindful of. For instance, the mass and the lack of peripheral vision of the headsets used in the studies were suggested as potential influential factors on the studies' findings. To expand, during the study by Wirth et al. (2021) each scenario (18 in total) involved a pattern recall task whereby on receiving the ball the screen was blanked, and players were asked to provide verbal feedback on the number of virtual players that were present on the screen, with the lesser experienced group having a statistically better recall ability (97.67 ± 2.23%) than the more experienced group (93.79 ± 5.44%). Given professional soccer players have been shown to utilise their peripheral vision more than their amateur counterparts (Williams & Davids, 1998), it was suggested the reduced peripheral vision in the headset could have impacted the groups to perform in the task. However, it should be noted that the unavailability of player perceptions on headset weight or lack of peripheral vision and their influence on task performance make these theories speculative, highlighting the added value of adopting a mixed-methods approach in research study design (Abt et al., 2022).

Virtual reality has also been used to measure characteristics in goalkeepers including the initiation of hand movement (through use of trackers placed on the gloves) when attempting to save a free-kick (Dessing & Craig, 2010). In the study by Dessing and Craig (2010), recreational and expert goalkeepers took part in a VR goalkeeping task with the aim of saving as many balls delivered from a free kick. Hand movements were shown to be biased towards the initial direction of the ball trajectory, which during a curved free kick resulted in a bias towards the direction opposite to where the ball would end up. Further the study reported that one expert goalkeeper waited longer until initiating movement and suggested that by waiting longer the goalkeeper was able to better predict the end placement of the ball towards the goal. However, by only reporting the findings of a single expert goalkeeper, these results cannot be generalised to all goalkeepers. Further, data that is easily interpreted for the reader was not provided, making it difficult to assess the findings of this study. Similar to Wirth et al., (2021), the study highlighted the benefit of using VR in soccer research, by allowing for the trajectory of the ball to the repeated across trials with precision, which is impossible to replicate in the real-world.

Soccer VR has also been used to measure an ‘overall diagnostic score’ which is purported to reflect the overall ability of a soccer player (Wood et al., 2021). Three groups of soccer players (professional, academy and novice) completed four soccer related tests, with data captured used to compute scores relating to the players passing accuracy, reaction times, composure, and adaptability. The study found that performance scores in three out of the four tests could be differentiated between playing groups, as could an overall diagnostic score calculated from an algorithm using the four performance tests. For instance, the overall diagnostic score showed statistically significant differences with very large effect sizes between novice and academy players ($d = 0.98$, $p = 0.27$), novice and professional players ($d = 2.73$, $p < 0.001$) and academy and professional players ($d = 1.45$, $p < 0.001$). However, it is worth noting that the author did not provide any data to illustrate the uncertainty in the magnitude of differences between players. For example, 95% confidence intervals can be derived from the mean, standard deviation, and sample size, and allows the reader to assess the uncertainty in the data and what the true population difference could be. For instance, a 95% confidence interval to alleviate the uncertainty in the data, between academy and professionals would have shown the true population difference to be between two large effect sizes ($d = 0.7 – 2.2$), confirming that it is very likely that strong evidence for a difference exists, albeit with large uncertainty. In contrast, if a 95% confidence interval was included between novices and academy players, the true population difference would have shown small to large effect sizes ($d = 0.28 – 1.7$). That said, the study’s finding that characteristics of soccer players can be measured and differentiated between groups has possible applications in areas such as talent identification, although further evidence is required before this claim can be made. However, this area of practical application was not suggested by the authors, who instead suggested that because construct validity was demonstrated, the soccer-specific simulator might have potential for use during player rehabilitation as a means of maintaining perceptual cognitive skill whilst avoiding physical load (Wood et al., 2021). Yet, this study did not investigate VR in relation to rehabilitation, which is a clear indicator of ‘hype’, whereby exaggerations are made on the benefits of the research, usually beyond the aims and scope of what was originally set out (Caulfield & Condit, 2012). Whilst it is possible that the authors exaggeration has been made unintentionally, others have suggested reasons as to why researchers might hype their findings, such as to best position and represent their research that is attractive to journals and press, often at the expense of scientific enquiry (Caulfield & Condit, 2012). Irrespective of why a researcher might choose to hype their findings, the consequence of this falls on the reader, especially if information is interpreted at face value. For instance, a soccer club might decide to use VR for the purpose of rehabilitation, knowing that a study has validated and recommended its use in that area. Although VR might provide several use case possibilities, it is important to consider the dynamics of the soccer environment when considering a new technology such as VR.

2.3.6 Social and cultural considerations for virtual reality use in soccer.

Although supporting evidence for the efficacy of VR in soccer is considered important for it to be successfully implemented in the sport, it is not the only consideration that individuals working in soccer clubs will have. For instance, in accordance with the diffusion of innovation theory, technology adoption is likely to be dependent on the end-users' subjective perceptions of usability, complexity and observability (Donaldson et al., 2012; Rogers, 1995). It is therefore plausible that decisions on using VR in soccer could be made from initial observations of the technology at conferences or trade demonstrations, as opposed to evidence for its efficacy. Soccer clubs might also be influenced by the behaviour, attitudes, and actions of others when it comes to technology adoption decision making (Venkatesh et al., 2003; Windt et al., 2020). An example of such influence is the 'neighbour effect', where it's possible that clubs might implement VR simply because others are also implementing it and do not want to lose out on a competitive advantage. Previous research has demonstrated this phenomena, whereby hospitals were more likely to adopt surgical robotics if their nearest neighbour (hospital) had also adopted them (Li et al., 2014). Interestingly though, the use of surgical robotics in hospitals has been questioned due to inconclusive evidence that they improve patient outcomes to any greater extent than compared to other traditional methods such as open or laparoscopy surgery (Köckerling, 2014). Furthermore, with such a large implementation and operational cost, the cost to benefit ratio alone should raise caution. Yet, as shown by Li et al. (2014), this is not always the case.

In contrast to the reasons why a soccer club might implement VR irrespective of its efficacy, there may also be reasons for a soccer club being resistant to the implementation of VR. First, soccer clubs are by nature old-fashioned constructs with deep local roots, and a high degree of cultural preservation (Gibbons, 2016; Plattfaut & Koch, 2021). While the integration of technology in soccer is seen as important, some of these technologies could conflict with the traditions of the club, or key stakeholders that have a particular method of operation (Plattfaut & Koch, 2021). In the context of VR used in the soccer training ground environment, it is plausible that the technology could be seen as a step away from what is seen as tradition, especially if it is believed that such technology will be used to replace existing methods (e.g., tactical analysis purposes). That being said, soccer clubs are showing interest towards technological advances, such as methods to explore big data, and artificial intelligence, which are 'traditionally' not part of soccer. Even Ph.D., sport scientists are now being embedded into soccer clubs to answer research questions, which might include investigation of new technologies that are adopted in the soccer environment (Champ et al., 2020). In the next section, the thesis will introduce literature relating to passing in soccer, and the assessment of this skill. This skill is specifically discussed as it is an integral part of the studies in Chapter 3 and Chapter 4.

2.4 Passing in Soccer

Ball passing is a fundamental skill in soccer, and is defined as the transfer of the ball from one player to another of the same team (Aquino et al., 2017). There are multiple factors that influence the number of passes performed by a player during a 90 minute game, such as playing position (Carling, 2011), team formation (Carling, 2011), and percentage of ball possession (Bradley et al., 2013). Successful ball passing between teammates is essential as it allows for the team to keep possession and is a more efficient way of moving the ball around the defending opposition than dribbling. Furthermore, in the five minutes proceeding a goal, scoring teams have been reported to have a significantly greater number of passes than the mean of the half played (Redwood-Brown, 2008). For these reasons, soccer coaches and scouts understand the importance of identifying players that are competent in this skill. However, quantifying ball passing ability in players through observations made in a soccer game is complicated as the number and success of these passes is likely to be affected by several factors, as described above. For that reason, researchers have designed passing tests that can be replicated in a controlled environment.

2.4.1 Passing skill assessment

One of the most commonly recognised passing skill tests is the Loughborough Soccer Passing Test (LSPT) developed by Ali et al., (2007). Other soccer passing tests include the passing accuracy test, and passing and visual recognition tests (Bekris et al., 2020), the short and long passing tests from the FIFA Medical Assessment and Research Centre (F-MARC) battery (Rosch et al., 2000), and the 20 m pass, 35 m pass, wall-pass, and short passing performance with rebound boards at an angle of 90 degrees (Rebound 90), and 135 degrees (Rebound 135) from the University of Queensland Football Skill Assessment (UQFSA) (Wilson et al., 2016). Several studies have attempted to demonstrate the reliability of these soccer skill tests by comparing two or more trials completed by the same participant, commonly referred to as test-retest reliability (Hopkins, 2000). For the LSPT, the authors reported poor to moderate reliability (ICC: 0.42 - 0.64) for their overall measure of performance, which varied depending on soccer playing level (Ali et al., 2007). Incorrectly, the authors reported good reliability of the LSPT in their discussion, which contradicts the interpretation of intra-class correlation coefficients (ICC) (Koo & Li, 2016), where a 'good' ICC would be classed as one between 0.75 and 0.9. Similarly, the passing tests developed by Bekris et al., (2020) showed varied reliability which ranged from poor to good in both the passing accuracy test (ICC: 0.21 - 0.76) and passing and visual recognition tests (ICC: 0.33 - 0.83). In contrast, the UQSFA passing tests demonstrated good to excellent reliability across all tests (ICC: 0.88 - 0.94)(Wilson et al., 2016), and the F-MARC passing tests demonstrated good to excellent reliability in the short pass (ICC = 0.82, 95% CI = 0.68 - 0.9) and good reliability in the long pass (ICC= 0.82, 95% CI = 0.66 - 0.89), (Padrón-Cabo et al., 2019). Notably, only Padrón-Cabo et al. (2019) reported the 95% confidence intervals for the ICC, which allows for improved interpretability of the data so the reader does not have the rely entirely on the point estimate (Koo & Li, 2016). As reported above, the intra-class correlation coefficients for both the short (0.82) and long (0.82) passing tests might be interpreted as having good reliability, if interpreted by using the point estimate alone. However, the inclusion of the confidence intervals indicates that the reliability for the true population (not the point estimate) is in fact moderate to good (Koo & Li, 2016).

2.5 Methodological considerations for testing

Measuring the performance of an individual can serve multiple purposes including being able to evaluate an intervention programme (Russell et al., 2011), establish performance profiles to evaluate the strengths and weaknesses of athletes (Robertson et al., 2015), inform talent identification processes (Goto et al., 2015), or provide diagnostic information related to the health of individuals (Stølen et al., 2005). To ensure the researcher or practitioner is confident in the test findings, it is important that the test displays the appropriate criteria of measurement properties (Robertson et al., 2017). Measurement properties typically refer to the quality of the data obtained from a test, such as its reliability, need of familiarisation, and validity (Hopkins, 2000; Robertson et al., 2017). It is generally agreed that a test's measurement properties are assessed before it is used in practice, as otherwise there is risk of imprecise or biased results being obtained that lead to the wrong conclusions (Mokkink et al., 2010). Measurement property checklists have been proposed in the fields of patient health (Mokkink et al., 2010), and sport science (Robertson et al., 2017). In the follow sections, the measurement properties, familiarisation, and validity are discussed in more detail.

2.5.1 Familiarisation

To understand what is meant by the term familiarisation, it is first important to explain the term 'reliability'. Hopkins (2000) defined reliability as the repeatability or reproducibility of a measure or variable, whereas Portney and Watkins (2000) refer to reliability as the extent to which a measurement is consistent and free from error. The term 'error' as alluded to by Portney and Watkins (2000), refers to measurement error which is common across all human participation research and is a combination of random and systematic error. Random error of measurement is generally accepted as inevitable when measuring human participants, and is introduced by factors acting independently between individuals, that are caused by the variable nature that humans respond to repeated tasks (Hopkins, 2000). In this sense, it is therefore not possible for a measurement to be truly free from error. However, Portney and Watkins (2000) were possibly referring to systematic error, a non-random change in the mean between two or more trials, and refers to predictable errors in measurement (Portney and Watkins, 2000). A systematic error that is often discussed within literature is the learning effect, whereby participants perform better in the second trial because they've benefitted from the experience of the first trial (Hopkins, 2000). Research aiming to establish the reliability of a measure should therefore look to minimise or eliminate the learning effect where possible, by ensuring 'familiarisation' to the test. That said, familiarisation to a test is also necessary even when the reliability of a measure is not necessary (such as talent identification in soccer) as it ensures that optimal results are obtained that best reflect the capabilities of the individual being tested (Ramsay et al., 2001).

Familiarisation can be affected by several transient factors such as prior experience (practice), fatigue, boredom, and environmental conditions (Dubrowski, 2005). However, the most commonly researched factor is time, as generally speaking, the performance of a repetitive task changes over time (Ramsay et al., 2001). Research that investigates the effects of familiarisation over time (duration or trials) will often refer to the term 'learning curve' (Anzanello & Fogliatto, 2011), which has been investigated extensively in the VR medical surgery field (Feldman et al., 2009; Hogle et al., 2007).

2.5.2 The learning curves

The learning curve refers to a phenomenon used to describe improvements in performance that are rapid to begin with, before tailing off until a plateau is reached (Ramsay et al., 2001). Several studies have investigated the learning curve of VR tasks, particularly with surgical simulator disciplines including, hysteroscopy sterilization (Janse, Goedegebuure, Veersema, Broekmans, & Schreuder, 2013), gastrointestinal endoscopy (Eversbusch & Grantcharov, 2004), laparoscopic (Hogle et al., 2007) and vascular intervention (Patel et al., 2006). To assess the learning curves, researchers typically investigate multiple trials of the VR simulation to examine how performance changes over time. For example, in a study exploring the learning curve of a hysteroscopy sterilization VR simulator, results showed it took nine trials for the outcome across all variables to statistically and significantly improve from the first trial, with most variables improving rapidly within 3-5 trials (Janse et al., 2013). Similar findings have been reported in laparoscopic surgery simulators (Hogle et al., 2007) and gastrointestinal endoscopy simulators (Eversbusch & Grantcharov, 2004), where a performance plateau emerges by the 8th and 7th trias, respectively.

2.5.3 Familiarisation to virtual reality sport

Familiarisation to a test serves the purpose of minimising or alleviating the learning effect, thereby reducing systematic error required for repeated trials of a test, such as when the test is used to measure a variable of interest in an intervention study (Hopkins, 2000). Several studies have used VR as part of an intervention, whereby VR is used to measure performance related variables. However, these studies did not familiarise participants to the VR test before the intervention took place. For instance, Gray (2017) did not familiarise participants to a VR baseball hitting test that was used as part of an intervention study to understand the effects of VR training. In the study, participants completed a VR baseball hitting test before being split into one of four training groups (VR training, adaptive VR training, real world training, and a control). After the intervention period, all groups repeated the VR baseball hitting test, to which all groups significantly improved on the number of balls hit including the control group who did not do additional real-world and VR training. It is plausible that the control group improved due to becoming more familiar with the test and interaction with the virtual environment. While the magnitude of improvement was still greater in the three training groups (VR training: $d = 1.5$, adaptive VR training: $d = 2.8$, real world training: $d = 1.7$) by comparison to the control group ($d = 1.1$), it should be acknowledged that it is possible the training groups also improved because of familiarisation, and thereby the training groups magnitude of improvement (from pre to post) should be questioned.

Familiarisation also serves the purpose of ensuring accurate results are obtained that best represent the capabilities of the individual being tested (Ramsay et al., 2001). Several studies that have used soccer specific VR to measure metrics of interests have adopted a familiarisation period. Familiarisation methods differed by number of trials: two (DiCesare et al., 2020), five (Wirth et al., 2021), and twenty (Dessing & Craig, 2010), or by time: one minute of practice (Shimi et al., 2021). However, two of these studies did not actually familiarise participants to the test itself, but instead provided two tutorials (Ferrer et al., 2020) or modifications to the actual test (Shimi et al., 2021). As such, participants were most likely familiarised to the environment and how to interact with the environment as opposed to the specifics of the test. Lastly, Brault, Kulpa, Duliscouët, Marin, and Bideau (2015) indicated a training period took place without further elaboration, whereas Wood et al. (2020) provided no familiarisation period. Of the studies that did provide a familiarisation period, none of them justified why the length or duration of the familiarisation period was chosen, and therefore it is unclear what the optimal familiarisation period is to ensure the learning effect is removed. This is problematic, as studies might under familiarise participants to the test, to which systematic bias (i.e., a learning effect) could still exist in the data. In contrast, over familiarisation might affect the data in an entirely different way, so that participants are fatigued due to repetitive activity, which might deteriorate the ability to perform in the test. To date, no study has explored the learning curve of a soccer specific VR simulator, and therefore not determined the point at which familiarisation is achieved.

2.5.4 Validity

Numerous definitions of the term validity, exist. For example, Lohr (2002) defined validity as to the degree to which an instrument or test measures what it purports to measure. Others such as Messick (1990), defined validity to the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations based on the test scores. Messick (1990) defines validity further to contrast with the definition of Lohr (2002), by suggesting that validation is not necessarily for the test or device that is being used (e.g., VR), but the inferences derived from the tests scores, such as its meanings and interpretations, and the resulting implications these have. Taking both definitions into account, a test or instrument could be considered as valid, but only when its meanings and interpretations are considered. Validation should therefore take into account the practical applications before a test or instrument is assessed (Messick, 1990).

Validation of a test is generally agreed as being important before the test is used as part of a research intervention, or used in a practical setting (Robertson et al., 2017). This is because without established validity the results of a test can be misleading and consequential, such as when decisions are made from the basis of a tests score (Im et al., 2019). In other words, the test outcome(s) could indicate an individual's score despite that score being non-existent or false (Patrick & Beery, 1991). To illustrate this point, the intervention study by Gray (2017), used a VR baseball batting test that was included to measure baseball related performance (such as number of balls hit) without any mention of it being validated. The study found that following a VR or real-world training intervention, performance in the VR test improved. However, it can not be determined with confidence that the test was measuring what it intended to measure (i.e., baseball hitting). For instance, it is unknown if the number of hits in VR has any relation to the number of hits that could be achieved in the real-world, to which a criterion method of validity would look to achieve. Alternatively, assessing differences between groups on the VR test, that are expected to differ at baseball batting in the real-world (construct validity) could have been investigated. These types of validity are explored further.

2.5.5 Construct validity

Construct validity is theory dependent and is used to assess a test or measure that is to be interpreted as a measure of some attribute or quality which is not operationally defined (Cronbach & Meehl, 1955). To establish that a test or measure has construct validity, researchers need to present correlations between a measure of a construct and a number of other measures that should theoretically be associated with it (convergent validity) or vary independently of it (divergent validity) (Westen & Rosenthal, 2003). Construct validity can also be obtained through group differences if there is knowledge to suggest that two or more groups are expected to differ in the measured construct (Cronbach & Meehl, 1955). This form of validity is also referred to as ‘known group differences technique’ (Impellizzeri & Marcora, 2009).

Historically, the known group differences technique has been used to validate VR medical simulators, by differentiating surgical related variables between experienced surgeons and junior surgeons or medical students (Albrecht et al., 2022; Jacobsen et al., 2015; Khanduja et al., 2017; Palet et al., 2021). For instance, Palet et al. (2021) compared the performance of a VR knee arthroscopy simulator between novice (n = 20) and expert (n = 10) orthopaedic surgeons. Following the completion of 11 VR simulation tasks that recorded a total of 44 variables, the results showed that experts performed equally or greater than novices in 43 of the measured variables (98%), and statistically significantly ($p < 0.05$) greater than novices in 41% of the tasks. Additionally, both groups also completed seven advanced simulator tasks consisting of a total of 75 variables, with experts showing a superior performance in 48 of the 75 variables (64%, $p > 0.05$), compared to novices who outperformed experts in one variable (1.3%, $p > 0.05$). However, limitations in the study include a small sample size that is justified through convenience samples as opposed to precision. This is problematic because small sample sizes are likely to result in underpowered study designs, which are likely to increase false positives, false negatives, and produce overestimated effect sizes (Mesquida et al., 2022). Unfortunately, the study did not report effect sizes that would identify the magnitude of difference between groups, and these can not be calculated as the study only reported the median in each group, without any data that reports the uncertainty (e.g., confidence intervals, standard deviation). Further, there is a possibility that the expert group might have had previous exposure to the VR simulator, as it was highlighted that none of the novices had previous experience with surgical simulators but no information was provided on the experts previous experience with simulators (Palet et al., 2021). As such, while differences were shown between groups, this might have been a result of one group not being familiarised with the test (novices) and the other being familiarised (experts).

Sport VR applications have also adopted the construct validity approach. For instance Kittel, Larkin, Elsworthy, and Spittle (2019), assessed the construct validity of VR by comparing the decision making abilities between two groups of Australian football umpires, those considered experts (officiating in the Australian Football League) and those considered amateurs (officiating in Australian football Metropolitan division one). Participants viewed 60 video clips though a VR head mounted display of small-sided games, designed to represent match play scenarios. Each video contained a possible infringement, which required participants to provide a verbal response that would indicate their decision, and each response was compared to the decision of two full-time AFL umpire coaches, resulting in a correct or incorrect decision. The results showed that the decision-making ability was greater in the expert group (75.18 ± 6.37 %) compared to the amateur group (65.21 ± 7.29 %, p = 0.001), with the authors concluding 'strong' evidence of construct validity. Yet, the authors provided no information as to why this method of validity (construct validity – known group differences) was chosen, considering that it appears as though the aim of the study was to compare between methods of video footage (VR vs broadcast) as opposed to what they assessed, being between skill level (expert vs amateur). As previously alluded to by Impellizzeri and Marcora (2009), this example shows why it is important that authors have a clear theoretical framework in place for assessing the validity of a test, regardless of how sophisticated or novel the statistical analysis is. Similar expert to novice comparisons have been adopted in golf as reported by Harris et al. (2021) who assessed VR putting ability between these two groups. Participants in both groups were required to putt a golf ball towards the hole from a stationary distance of 3.05 m away. In the VR scenario the hole was covered (but still visible to the participants), meaning the ball would not drop into the hole. The aim was therefore to putt the ball as close to the hole, with the outcome measure being the radial error of the ball from the hole. The results of the study showed that in the VR condition, expert-amateur golfers (44.1 ± 18.5 cm) outperformed the novice golfers (74.2 ± 29.3 cm), by displaying significantly lower radial error scores ($t(34) = 3.68$, $p = 0.001$, $d = 1.23$).

To date, only three VR soccer validation studies exist, with all using a construct validity approach of the known-groups difference technique. Wood et al. (2020) investigated the construct-validity of four VR soccer tests by comparing the performance between three groups of soccer players that were theoretically expected to differ in their soccer-related abilities (professionals, academy, and novices). The study found that performance scores in three out of the four tests could be differentiated between level of player, as could an overall diagnostic score calculated from an algorithm using the four performance tests. The authors concluded that the results demonstrated construct validity and that the system was suitable for use in professional soccer. In another study, Wirth et al. (2021) examined visual exploratory activity through evaluating head motion characteristics from the head mounted display across 18 VR soccer scenarios, between two groups of soccer players (low experienced, and experienced). The results showed significant differences between the two groups for several head turn measures such as head turn frequency (low experience: 9.81 ± 0.70 vs. experienced: 8.93 ± 0.68) and excursion (low experience: 66.72 ± 4.61 deg vs. experienced: 58.94 ± 5.02 deg), and that such findings demonstrate the validity of the measures captured through VR for assessing the visual exploratory activity of soccer players. Similar to Wirth et al. (2021), Ferrer et al. (2020) also investigated visual exploratory activity between two groups of soccer players (beginners and amateurs). Participants completed a single trial of an in-game soccer situation where the user had seven seconds to search the soccer field consisting of virtual teammates and pressing opposition and follow up with a decision on which team-mate to pass the ball to. Head excursion data was used as an outcome measure and was represented as a percentage of time spent exploring areas greater than 45˚ in either rotated direction of the head. Amateur players were shown to have a higher average visual exploratory score (score = 49.83, SD = 19.86) than beginners (score = 30.83, SD = 17.61), with an obtained significant difference (diff = 19, 95% CI = [3.1, 34.8], *p* = 0.021). However, it is important to note the limitations in these studies methods and how this can lead to a misinterpretation of the results. For instance, two of the studies failed to include a measure of uncertainty (such as confidence intervals) when reporting the differences between groups and relied entirely on the point estimate in their interpretation of the data.

To highlight this issue, Wood et al. (2020) reported large effect sizes (*d* = 0.98) for differences in overall diagnostic scores between academy and novice players, without any measure of uncertainty. With the inclusion of a 95% confidence interval (CI, to alleviate uncertainty in the data), the data would have shown that the true difference in the population could lie anyway between 0.28 (small effect size) to 1.7 (large effect sizes). Further, when applied to one of the individual tests ‘shoulder sums’, that was reported as having a significant (p = 0.049) and large difference (*d* = 0.82) between professional and academy players (indicating the test can differentiate between playing level), the uncertainty is broadened to an effect size range of trivial (*d* = 0.1) to very large (*d* = 1.5). In other words, the true population difference for that specific test could be non-existent or even larger than the study has reported. Similarly, this issue is demonstrated in the study by Wirth et al. (2021) who reported that low experienced soccer players tended to focus their visual attention significantly longer on the opposition than the experienced players (*d* = 0.93), suggesting that experienced players deployed different visual exploration strategies such as focusing more on team-mates and the ball. However, a calculation of the effect sizes with a measure of uncertainty would show that the true population difference falls between trivial and very large (*d* = 0.93, 95% CI = 0.15 to 1.76). The large variability in true population differences makes it difficult to interpret the results in both studies and could be a consequence of the small sample sizes reported. For instance, Wood et al. (2020) had an equal numbers in each group (n = 17) which was based on an *a priori* examination of effect sizes reported from a similar study (Harris et al., 2021) and not from an *a priori* sample size calculation, whereas Wirth et al. (2021) sample size (low experienced, n = 17; experienced, n = 11), and Ferrer et al. (2020) sample size (n= 12 in both groups) were based entirely on convenience without any indication of sample size calculations. In the case of Wood et al. (2020), the issue of a low sample size could have been easily addressed given the author reported having access to a 250 participant sample made up of 125 professional players, and 125 academy players. This number was only reduced to match the sample size of the novice group, where only 17 participants were recruited. Yet, this method is argued as illogical, as the larger sample would have been likely to have decreased the confidence interval width and provided greater certainty on true population differences between those groups, even if uncertainty still existed when comparing against the novice group.

2.5.6 Criterion validity

Criterion validity relates to evidence that shows the scores of the test are related to a separate criterion measure (Lohr, 2002). Different types of criterion validity exist, such as predictive validity, whereby the scores of a test are used to predict the scores of an alternative measure at a future time point (Messick, 1990). Other forms of validity include concurrent validity which compares the scores of a test, to another test with already established validity, at the same point in time (Messick, 1990). As the name implies, this is typically achieved by assessing two scores that are obtained through different mediums or approaches, simultaneously during the same test. For instance, GPS sprint acceleration data has been simultaneously compared to the similar data recorded through laser and radar devices, following maximal straight line sprinting (Nagahara et al., 2017). Alternatively, during instances where it is not possible or practical to assess two scores obtained through different methods simultaneously during the same test, some approximation of time is considered acceptable (e.g., comparing two tests during the same day, or days apart). For instance, the concurrent validity of two metabolic gas analysis systems for use during maximal oxygen consumption (VO_2max) testing can not be obtained simultaneously (due to the fitting of the individual masks), and so a repeated measures design would have to be undertaken. However, it is important for researchers to be aware of the limitations of this approach, such as the impact of fatigue on repeated tests. For example, VO_2max testing on the same day is likely to impact the results due to the physiological and mental fatigue of repeating such a test (Stølen et al., 2005).

To date, only one sport VR validation study has adopted a criterion based, concurrent validity approach. The study by Harris et al. (2021), compared golf putting ability between real-world and VR environments and found a positive relationship (r = 0.46, p = 0.004) for the radial error (distance of golf ball to the putt) between conditions. In other words, those with a lower radial error score in the VR condition, also had a lower radial score in the real-world. Such a finding is important as to evaluate a sport skill in VR it must be compared to the equivalent skill in the real-world. However, based on a sample size of 36 participants, the 95% CI (0.16 to 0.69) would suggest that the true population relationship could be much weaker or stronger than reported. Further, it is worth noting that the study used an unvalidated golf putting test in the real-world, and so it can not be determined with certainty that the real-world test was an effective assessment of golf putting ability. For this reason, it is recommended that the criterion chosen to be validated against, be one that already has evidence of validity (Impellizzeri & Marcora, 2009). For example, Harris et al. (2021) could have replicated the Golf Australia Putting test by Robertson, Gupta, Kremer, and Burnett, (2015) in the virtual environment, that has evidence of construct (known-groups differences) and criterion (predictive) validity. That said, it is duly noted that Harris et al. (2021) were not intending to validate the golf putting task for use in the sport of golf, per se, but as an illustration of the different methods in which VR could be assessed against the real-world. Yet, one method they did not consider was the test-retest reliability of the real-world task. This is important as it could be argued that if VR is to be compared to an equivalent real-world task, but the real-world task is not reliable (i.e., variability in the scores exist between repeated trials), then it is possible that a correlation between the two conditions (VR and real-world) might not exist, simply because random variability in the two scores prevent them from doing so.

2.6 Perceptions of virtual reality

2.6.1 Implementation science

While establishing the validity of VR is important, it does not guarantee its uptake into routine practice (Bauer & Kirchner, 2020). Early work by Rogers (1962), conceptualised the diffusion of innovation as a social process, with multiple determinants that are beyond the efficacy of the innovation itself. For instance, aside from the effectiveness of a technology, organisational policies and individual opinions are likely to contribute towards its adoption. This field of understanding what barriers exist towards a technologies adoption, along with strategies to overcome these barriers to increase uptake, is known as implementation science (Bauer &

Kirchner, 2020). Implementation science can be defined as the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practice into routine practice (Eccles & Mittman, 2006). In the initial stages of implementation science, identifying behaviours that contribute to the evidence-practice gap is required, and can be evaluated through quantitative measures such as structured surveys (Bauer et al., 2015; Handley et al., 2016). To assess the acceptance and implementation of technology, several technology acceptance models have been proposed in the literature, that utilise surveys as a means of understanding perceptions.

2.6.2 Technology acceptance models

Technological devices with established validity and efficacy may not be always be accepted by the end users (Mascret et al., 2022). For this reason, models have been devised to investigate the psychological determinants that may influence the intention to use a technology. One of the earliest and most widely used models is the technology acceptance model (Davis, 1986). The technology acceptance model was designed on the beliefs that perceived usefulness and perceived ease of use are the most important contributors towards technology acceptance. Other models such as the theory of reasoned action is a widely studied model from the field of social psychology, and is concerned with the determinants of consciously intended behaviours (Fishbein & Ajzen, 1975). According to the theory of reasoned action, a person's behaviour towards technology (e.g., intention to use a technology) is determined by their behavioural intention to perform the behaviour (e.g., use the technology), which is jointly determined by a person's attitude (e.g., positive or negative feelings about using the technology), and subjective norms (e.g., the person's perception that most people important to them think they should or should not use the technology) (Davis et al., 1989; Fishbein & Ajzen, 1975). Unified models, such as the Unified Theory of Acceptance and Use of Technology (UTAUT) also exist, that have been designed to take the most meaningful constructs that explain technology acceptance from other models, such as the technology acceptance model and the theory of reasoned action mentioned above (Venkatesh et al., 2003). In the UTAUT, four constructs are determined to play a significant role in technology acceptance and usage behaviour, including the technology's perceived usefulness (performance expectancy), its perceived ease of use (effort expectancy), the influence of social surroundings such as the users peers (social influence), and barriers towards the technologies use (facilitating conditions) (Venkatesh et al., 2003). In the sport and health sciences field, the UTAUT has been used to explain contributing factors towards using golf technology, such as GPS devices and screen golf where it was found that performance expectancy, effort expectancy, and facilitating conditions all positively contributed towards consumer intent on using golf technology, whereas social influence did not have a contributing effect (Seol et al., 2017). In other words, consumer intention to use golf technology originated from their own personal beliefs of the technology (e.g., its performance benefit and ease of use) as opposed to socially orientated ones such as their peers who might suggest using golf technology.

2.6.3 Perceptions of virtual reality in soccer

Only two studies have explored the perceptions of VR in soccer. The first was by Thatcher et al. (2021), who explored the perceptions of VR use in professional soccer. In the study, coaches and performance analyst staff working at professional soccer clubs took part in a semi-structured interview to understand how VR could be used in the sport, and to understand the barriers to its widespread adoption. From the interviews, several themes emerged on virtual reality's usefulness in soccer, including its use for tactical awareness, player development, and rehabilitation. Additionally, barriers were acknowledged that would prohibit its use in the sport, such as quality and accuracy of the current technology, the practicality of its implementation, and the lack of empirical evidence that is available. However, it is worth noting that the study was conducted with a small sample size (n = 6) that lacked any *a priori* sample size calculation, and in a specific cohort of staff working in professional soccer which does not fully represent the multi-disciplinary nature of professional sport. For instance, physiotherapists and strength and conditioning staff who are more knowledgeable on the player rehabilitation were not consulted, and so it is possible that their opinions of using VR during rehabilitation differs to those interviewed in the study. It is therefore difficult for the study's findings to be generalised to the wider population of staff that work in professional soccer, which would also include soccer players who are an essential stakeholder towards technology adoption such as VR yet were not represented in the study. Furthermore, although the study by Thatcher et al. (2021) highlighted some of the key opportunities and barriers to VR's adoption in soccer, it is unclear how they compare to one another. For instance, it is unclear how valuable VR is for player development in comparison to rehabilitation, or how much of a barrier a lack of evidence is by comparison to the quality of the technology. By quantifying these insights, VR developers, researchers, and practitioners can isolate the key areas of VR use in soccer and alleviate the top barriers to improve its accessibility. Finally, it is unclear what the mechanistic factors influence the intentions to use VR in professional soccer. For instance, referring to the UTAUT described in section 2.6.2, it is unclear from the study how much performance related factors (such as the belief that VR will improve performance) are compared to socially orientated ones (such as the influence that other peers, or soccer clubs have on adoption).

In contrast, Mascret et al. (2022) did investigate the mechanistic factors towards VR use in sport. In the study, athletes from a variety of competitive levels (recreational to international) and sports, completed a survey to investigate the acceptance of VR before first use (i.e., participants had no prior experience of VR). The survey consisted of questions derived from the technology acceptance model, and was used to examine the perceived usefulness, perceived ease of use, perceived enjoyment, social norms (i.e., social influences), and how these related to the intention to use. The results of a structural equation model revealed that intention to use VR was positively predicted by each of the constructs previously mentioned, with perceived usefulness being the strongest predictor towards intention to use. In other words, those who believe VR to be more beneficial have a greater intention to use VR in the future. The study is strengthened by the exceptionally large sample size used (n = 1162), that is rarely seen in sport and exercise science survey literature (Malone et al., 2019; Weston, 2018). However, this sample consisted of athletes from 62 different sports, meaning that the results can not be generalised specifically to professional soccer, especially considering only a small sample represented this population (n = 17), assuming 'regional level' soccer players (n = 54) are not professional players. Further, it is worth mentioning that the results of the study could have been influenced by positively orientated messages on the 'potential' benefits of VR that was read by participants before completion of the survey. For instance, the authors included a paragraph on what could be achieved through VR use, such as 'carry out tasks that will allow you to improve your performance in specific areas', and 'improve your own technique'. Such information might be misleading, as it implies that VR can improve these areas, when in fact only a paucity of evidence exists in only limited sports such as baseball, table tennis, and darts (see section 2.3.5). This is especially so given that 57% of participants had never used VR before, and no participants had used it in a sporting application, meaning that participants could be more suspectable to influence. As such, it is of opinion that the authors biases could have influenced the participant responses, and it is suggested that more neutral language is used in future.

2.7 Summary points

- Virtual reality consists of hardware and software components, and provides the user with an experience. Baños et al. (2000) defines VR as a computer-simulated environment that aims to induce a sense of being mentally and/or physically present in another place.
- Research has suggested the VR could be used to test specific characteristics of soccer players, due to VR being able to control and repeat the task/environment between players in a way that is not possible in the real-world. However, it's important that tests are assessed for their measurement properties such as familiarisation and validity.
- Familiarisation to a test serves the purpose of minimising or alleviating the learning effect, whereby participants perform better in a second trial simply due to knowledge

of the first trial. To date, no study has explored the learning curve of a soccer specific VR test.

- Validity as the degree to which an instrument or test measures what it purports to measure (Lohr, 2002), and is important as without established validity the results of a test can be misleading or consequential.
- Several studies have assessed the construct validity of soccer specific VR tests, showing they are capable of distinguishing between soccer players of differing ability levels. However, no study has explored the concurrent validity of VR in soccer.
- To date, only two studies have explored the perceptions of VR in soccer with key findings including its perceived usefulness such as for improving tactical awareness, amongst barriers towards its adoption such as a lack of evidence base. However, these studies used small sample sizes, or samples that don’t best represent the coaches, backroom staff, and players who work in professional soccer.

Chapter 3 Examining the learning effect associated with a virtual reality task in professional soccer players.

3.1 Introduction

Virtual reality is defined as a computer-simulated environment that aims to induce a sense of being mentally and/or physically present in another place (Baños et al., 2000, Neumann et al., 2018). A notable feature of VR is that it allows the user to interact with virtual objects in a controlled and repeatable environment (McMenemy & Ferguson, 2007). It is this feature that has attracted several industries to adopt VR, particularly for training purposes. For example, the medical surgical industry have used VR for the purpose of training and testing novice surgeons in complex procedures that would be unsafe to practice in the real world (Sutherland et al., 2006). However, given that VR is a novel environment for most people, it is difficult to assess if changes in task performance over time are the result of learning or simply that users have become familiarised to the task (Levac & Sveistrup, 2014). To examine this, studies have investigated what are known as 'learning curves', whereby the performance of a task is assessed over a period of time or number of trials (Anzanello & Fogliatto, 2011). Typical features of these learning curves include a rapid improvement in performance between the first two trials, followed by a gradual improvement in performance before a plateau occurs, referred to as familiarisation (Bartlett et al., 2020; Gallagher & Satava, 2002; Janse et al., 2013; Jokinen et al., 2020). Investigation of the learning curve, specifically the point at which familiarisation has been achieved, allows for the performance of a test to be evaluated knowing that systematic error has been removed (Hopkins, 2000).

The use of VR has been investigated in various sports as a technology for training and testing athletes, with studies in baseball, table tennis, and darts demonstrating cognitive-perceptual transfer of skills to the real-world (Gray, 2017; Michalski et al., 2019; Tirp et al., 2015). However, there is no information to date on the learning effects and familiarisation observed in a sport-specific VR task. Despite this, several studies have conducted familiarisation sessions before collecting data, yet the number of trials or time to familiarise participants seems arbitrary. For example, when evaluating the validity of a VR golf putting test, no evidence supported the rationale of completing three familiarisation trials before the data was collected (Harris et al., 2021). It therefore remains unknown if three familiarisation trials were enough to reduce systematic error. Studies have also been conducted without any reference to familiarisation before testing the athlete. For example, Gray (2017) did not familiarise participants to a VR baseball hitting test before the pre-intervention (baseline) scores were taken. After an intervention period, participants in all groups improved the number of successful hits in the VR baseball-hitting test (Gray, 2017). However, it could be argued that the improvements were a result of knowing how to better perform the baseball hitting task due to completion of the pre-test. That is, a learning effect has probably occurred, as both the control and intervention groups improved the number of successful hits over time (Gray, 2017). The lack of familiarisation prior to main trials means that systematic error might be present in the data, suggesting the results are not representative of the participants' true VR performance. To date, only one VR study in soccer has included a familiarisation protocol in the research design (Ferrer et al., 2020). Ferrer et al. (2020) used a familiarisation protocol consisting of two VR scenarios used to (1) help the user acclimate to the gaze and point mechanics of the head-mounted display, and to become comfortable wearing it during movements such as head rotations, and (2) to familiarise the participant with the passing mechanics of the ball. However, the impact of these on the results is unclear. Therefore, the aim of this chapter was to quantify the learning effect during familiarisation trials of a soccer-specific VR task.

3.2 Methods

3.2.1 Pre-registration

The study question, hypotheses, sampling plan, data collection procedure and data analysis plan were pre-registered on the Open Science Framework (https://osf.io/ya5r8). Any deviation from the pre-registration is clearly outlined in the relevant section of the method.

3.2.2 Participants and sample size calculation

Thirty-four professional soccer players (mean (SD) age, stature, and body mass: 20 (3.4) years; 180 (7) cm; 79 (8) kg) from an English third division soccer club took part in the study between August and November 2020. In relation to sample size, the study used a Bayesian sequential analysis approach where the precision of the population estimation (the credible interval) was the goal rather than a frequentist approach of power (Kruschke & Liddell, 2018). The minimum sample size of 30 participants was identified, which represented half of the total number of participants available at the soccer club where the study took place. Further details on the Bayesian sequential analysis including our justification for stopping data collection are discussed in the statistical analysis section (Section 3.2.5). Participants were excluded from the study if they had previously completed the VR task 'Rondo Scan' within the last 12 months (to desensitise participants who might already be familiar with the task), which was the original version of the VR task used in this study (Section 3.2.4). Participants were also excluded if they had any injuries at the time of testing which prevented normal soccer-related activities such as lower extremity kicking, and upper extremity body rotations. These injury exclusions were confirmed by the medical team. Participants were provided details on the aims, risks, and benefits of the study and then provided written informed consent. The Faculty of Health Sciences Research Ethics Committee at the University of Hull approved the study (REF: FHS246), which conformed to the Declaration of Helsinki.

3.2.3 Virtual reality instrumentation

Participants wore a VR head-mounted display (Valve Index Headset, Valve Corporation, Washington, United States) and trackers (VIVE tracker, HTC corporation, Taoyuan City, Taiwan) which were securely fixed to trainers (Revolution 4, Nike, Oregon, United States) provided by the researcher (Appendix 1). The VR soccer passing task was a modified version of the task 'Rondo Scan', available through Rezzil (Version.1.3, Rezzil Inc, Manchester, UK). Details of the modifications to the original version are provided in section 3.2.4. Both hardware and software were powered by a custom-built computer (Processor: i7-8700K 6-Core 3.7, GPU: 1080 Ti, Memory: 16GB, OS: Windows 10 Pro). The VR operating space was calibrated as a 4 x 4 m rectangle using SteamVR (Valve Corporation, Washington, United States). Three lighthouse stations (Base station, HTC corporation, Taoyuan City, Taiwan) tracked three-dimensional movement, and were fitted at heights ranging between 30 cm and 247 cm at both sides (along the coronal plane) of the participant (room set up is available in Appendix 2), ensuring tracking was maintained throughout the procedure. To ensure participant safety while using the VR system, participants completed the task within a playing circle with a radius of 0.74 m. Leaving the circle resulted in the head-mounted display screen being dimmed, providing participants with a visual warning that they had left the safe playing area. The head-mounted display screen reverted to full contrast on returning to the playing circle. Participants could still interact with the VR ball if they stepped out of the playing circle, although they were encouraged to stay within the circle throughout each trial. Audio from the software was delivered through speakers built into the head-mounted display.

3.2.4 Procedure

Prior to each trial, participants were provided with detailed information on the VR task using a poster to illustrate the task requirements (Appendix 3). Training shoes were provided to participants, and head-mounted display fitting instructions were provided visually and verbally. Participants confirmed that the head-mounted display did not cause discomfort at the time of fitting or throughout testing. Calibration of the foot trackers was completed following the software guidelines (Rezzil Inc, Manchester, UK) which started by entering the participant's regular shoe size into the software. Calibration took place while the participant wore the head-mounted display, with instructions provided visually and verbally by the software. The calibration protocol required the participant to stand with their feet shoulder width apart within a virtual footprint positioned on the floor. Calibration took approximately three seconds, and once completed, participants were presented with virtual soccer boots within the head-mounted display. Participants were asked to confirm if the virtual boots were aligned with their own foot position and were at a height that represented where their feet would be in the real world.

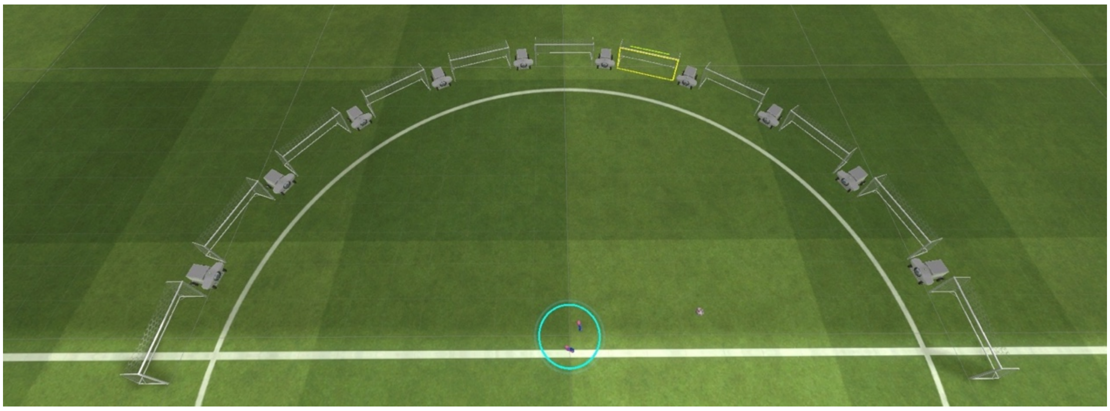


Figure 3.1 Aerial view of the VR soccer task featuring the mini goals, separated by ball launchers. The blue circle represents the playing zone that participants received and passed the ball from.

After calibration, a virtual soccer training pitch was displayed. Within the soccer training pitch, 11 mini goals were positioned in a semi-circle that surrounded the participant. Aligned with, and separating each goal, were ball launchers with a further four ball launchers positioned '12.5 m' from the participant, directly behind goals two, four, eight and ten (from left to right) (Figure 3.1) Participants were placed '9 m' in front of the goals and completed the task within the playing circle mentioned above (Section 3.2.3). To begin the task, participants were instructed to kick a virtual soccer ball which was placed on the pitch in front of them. This was the first point at which participants interacted with the virtual ball. Participants then commenced Level 1, which consisted of 15 balls randomly delivered one at a time from one of the ten ball launchers positioned between the mini goals. Participants were made aware of which launcher would deliver the ball by a 1 s green highlight before the ball was released. As the ball launcher highlighted green, a mini goal also highlighted yellow to indicate where the ball should be passed to (Figure 3.2). On receiving the ball, participants were required to pass the ball into the highlighted mini goal within 5 s from the moment the ball was released. Participants were instructed that they could play one touch (pass directly into the goal without stopping the ball), or two touch (stop the ball with the side of the foot and then pass the ball). On completion of Level 1, participants immediately commenced Level 2 which also consisted of 15 balls. However, Level 2 involved (1) balls also being delivered from behind the mini goals; and (2) participants were required to pass the ball within 4 s of being released from the ball launchers. In the original version, participants would only progress to level 2 if they passed 7 out of 15 balls into the correct mini-goals, and so the modification ensured that the number of balls delivered was standardised between participants. For both levels, successful passes were indicated to the participant by the mini-goal turning green, whereas unsuccessful passes turned the mini-goal red. At the end of Level 2, participants were presented with a score board that indicated the number of successful passes completed during each level, the mean pass reaction time (s), and the number of passes attempted with the left and right foot. Each trial took approximately three minutes to complete, and each trial was separated by a 1-minute rest period, allowing enough time to take on fluids and adjust the head-mounted display if needed. Participants completed six trials, in total. The dependent variable was the total number of goals scored in each trial, accumulated from Levels 1 and 2. Throughout the trial, feedback was not provided to the participant on their technique nor guidance on how to improve their score. No participants were tested within 72 h of being involved in a match.

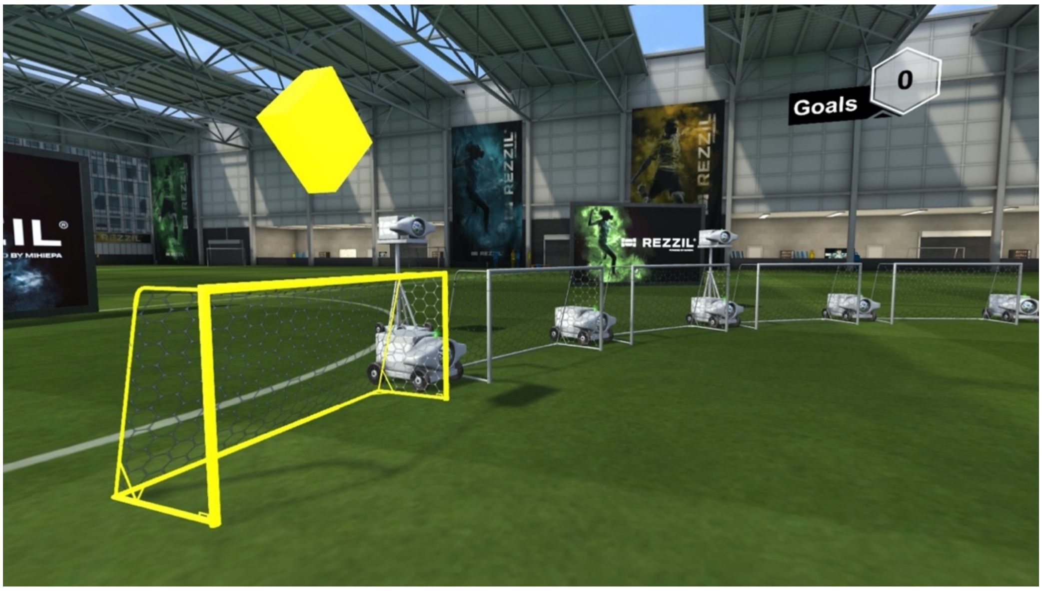


Figure 3.2. Virtual environment of the soccer task including the mini-goals and ball launchers. Visual cues are provided for the release of the ball (green launcher), and where to pass the ball (yellow mini-goal).

3.2.5 Statistical analysis

Data were confirmed as being normally distributed using a Shapiro-Wilk test and visual examination through quantile-quantile plots. Normally distributed data are presented as mean (standard deviation). Inferential analyses were conducted using a Bayesian approach conducted in JASP v.0.14.1 (JASP team, 2021). RStudio (RStudio Team, 2020) was used for data visualisation using the ggplot package (Wickham, 2016). We examined the changes in total goals scored across trials both in the form of a hypothesis test using Bayes Factors (BFs) and as a parameter estimation for the posterior distribution of the standardised difference. Uncertainty in the standardised mean difference is reported as the 95% highest density interval (95% HDI;

commonly referred to as credible interval). For model comparison we used BFs which quantify the relative predictive performance of the null hypothesis (H_0) compared to the alternative hypothesis (H_1). When computing BFs (Schönbrodt et al., 2017) or estimating the highest density interval of the posterior distribution (parameter estimation), Bayesians are free to monitor the data as often as they wish as it is being collected (Wagenmakers et al., 2018). As such, data analysis took place on achieving the minimum sample (see section 7.2.2), and then after every five participants. A stopping rule was set on data collection when the 95% HDI for the standardised effect size was at most 0.4 either side of the point estimate, or until the maximum number of participants available was reached. The rationale for using a HDI width of 0.4 as a measure of precision and as a stopping rule was derived by examining the changes in total goals across each of the six trials in a pilot study. The data from this pilot study suggested that the standard deviation across trials was approximately 5. We reasoned that the smallest effect size of interest would be a difference of two goals. Calculating Cohen's *d* using two as the numerator and 5 as the denominator resulted in d = 0.4.

For our first pre-registered hypothesis, a one-sided Bayesian paired samples t-test was used to confirm if there was an increase in total goals between trials one and two. We used a one-sided test because previous studies suggested that there would be an increase in goals scored between trials one and two. For our second pre-registered hypothesis, Bayesian paired-samples equivalence t-tests were used to confirm if the last two consecutive pairs of trials (4 vs 5; 5 vs 6; 4 vs 6) could be considered equivalent. Our equivalence zone was set as a standardised effect size of 0.4 either side of 0. As outlined above, this equated to a two-goal difference between each trial assuming a pooled standard deviation of 5 across trials. For hypothesis 1 and 2, a zero-centred Cauchy distribution prior was used with a scale 0.707. This prior was compared to a wide and ultrawide prior to assess the robustness of the data. Although BFs are a continuous measure of evidence of H_1 relative to H_0 with higher values indicating increasingly stronger support for H_1, we used a classification scheme as set out by Lee and Wagenmakers (2013) to aid interpretation. In this classification scheme, a BF_{10} (which indicates evidence for H_1) of between 1 and 3 is indicative of anecdotal evidence, between 4 and 10 is indicative of moderate evidence, between 11 and 30 is indicative of strong evidence, between 31 and 100 is indicative of very strong evidence, and > 100 is indicative of extreme evidence. To describe the magnitude of the standardised effect we used the classification scheme of Cohen (1988) with values of 0.2, 0.5 and 0.8 representing small, moderate, and large effects, respectively. The generic values of *d* suggested by Cohen are used because we have no other measure of the distribution of standardised effect sizes for this VR task.

3.3 Results

All the participants (n = 34) recruited for this study were included in the analysis. Across the six trials an improvement in goals scored was observed in the initial trials, followed by a plateau in performance (Figure 3.3). Our stopping rule criteria for data collection was shown between trials 5 and 6, but not between trials 4 and 5; and 4 and 6 (Figure 3.5). As set out in our pre-registration, data collection should have continued until the stopping rule criteria of all three equivalence tests had been achieved, or we reached the maximum sample available to us. However, due to soccer club not renewing the license agreement to use the VR software required in this study, which was unforeseen at the time of the pre-registration, data collection could no longer take place after November 2020.

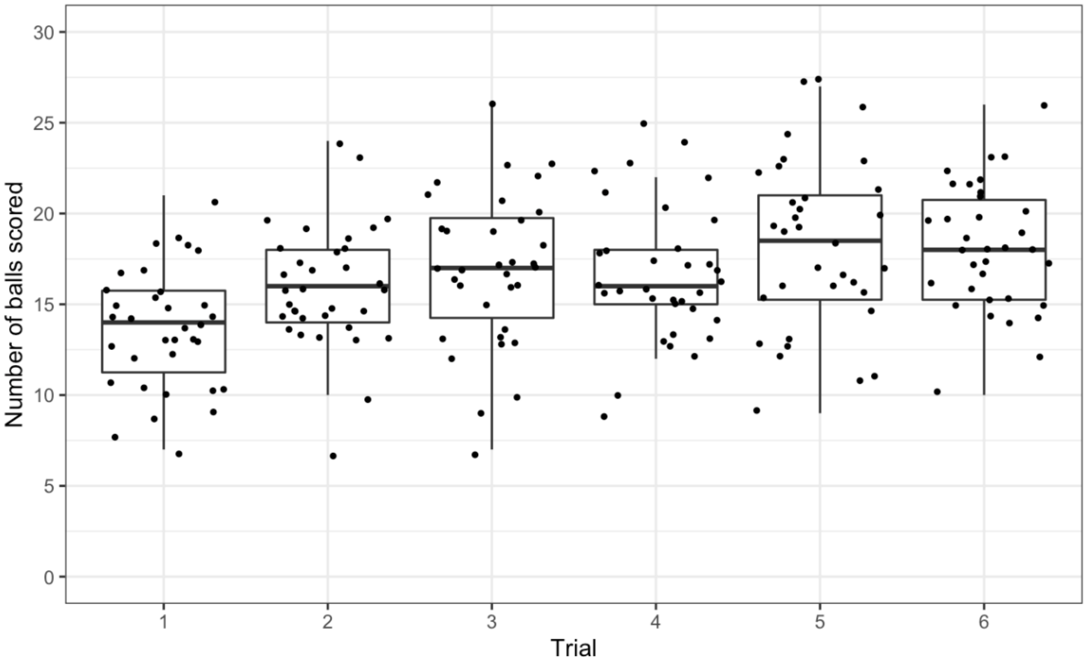


Figure 3.3. Number of goals scored in each trial of the VR task. Boxplots represent the 25th-75th percentiles, and the horizontal black line represents the median. Dots represent participant score in each of the trials. Whiskers extend to indicate the most extreme data values, with dots beyond the whiskers considered as outliers (1.5 times the interquartile range from the edge of the box).

For the first pre-registered hypothesis the one-sided Bayesian paired samples t-test indicated very strong evidence in favour of H_1 (BF_{10} = 46.5, d = 0.56 [95% HDI = 0.2 to 0.92]) for improvements in total goals scored between trial 1: 13.6 (3.3) and trial 2: 16 (3.3) (Figure 3.4). Cohen's *d* would suggest there was a moderate increase in goals scored between trials one and two, but there is considerable uncertainty in this estimate based on the 95% HDI. The BF robustness test remained relatively stable across a range of priors, suggesting that the choice of prior had little effect on the evidence in support of H_1 (Appendix 4).

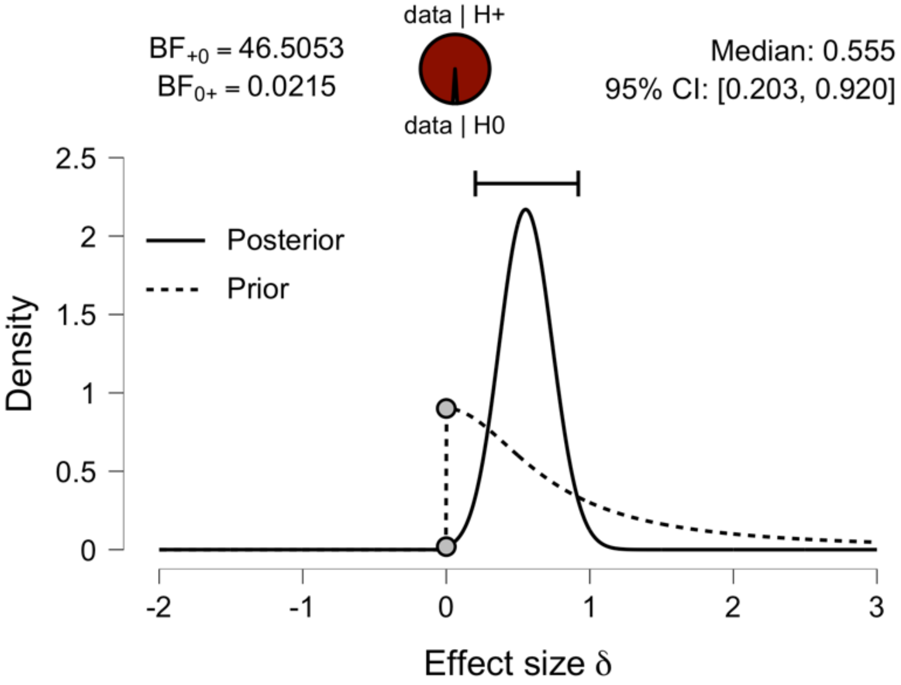


Figure 3.4. One-sided Bayesian paired samples t-test, prior and posterior distribution plot. The probability wheel provides a visualisation of the evidence that the data provides for H_1 (top left). Grey dots show the prior and posterior density values, with the median effect size and 95% high density interval (termed 'CI') in the top right.

For the second pre-registered hypothesis, the Bayesian paired-samples equivalence t-tests indicated strong evidence in favour of H_1 (BF_{10} = 10.2, d = 0.24 [95% HDI = -0.09 to 0.57]) for equivalence between trial 4: 16.7 (3.7) and trial 5: 18.2 (4.7); extreme evidence in favour of H_1 (BF_{10} = 132, d = -0.02 [95% HDI = -0.34 to 0.30]) for equivalence between trials 5 and 6: 18.1 (3.5); and moderate evidence in favour of H_1 (BF_{10} = 8.4, d = 0.26 [95% HDI = -0.08 to 0.59]) for equivalence between trials 4 and 6 (Figure 3.5). The sequential analysis with robustness check indicated instability across a range of priors between trials 4 and 5, and 4 and 6, whereas convergence was demonstrated between trials 5 and 6 (Figure 3.6).

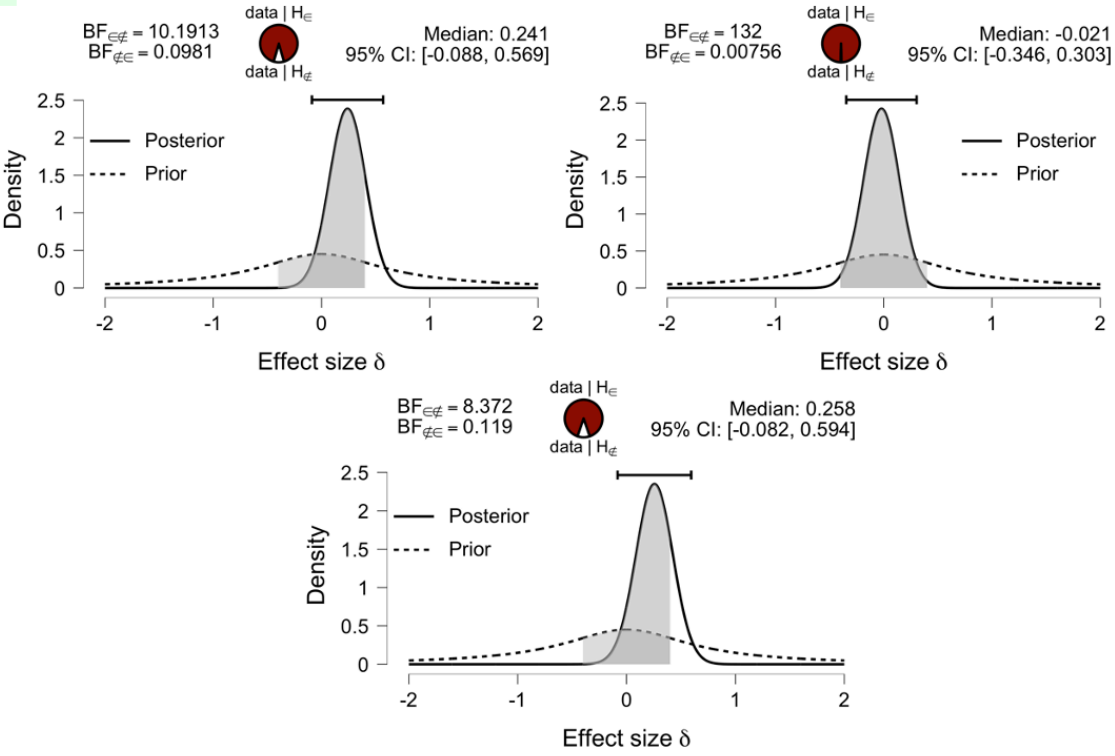


Figure 3.5. Bayesian paired samples equivalence tests, prior and posterior distribution plots. The probability wheel provides a visualisation of the evidence that the data provides for H_1 (top left). Grey dots show the prior and posterior density values, with the median effect size and 95% highest density interval (termed in the figure as ‘CI’) in the top right. The grey shaded area represents the equivalence region (-0.4 to 0.4). Plots are shown for trials 4-5 (top left); trials 5-6 (top right); and trials 4-6 (bottom middle).

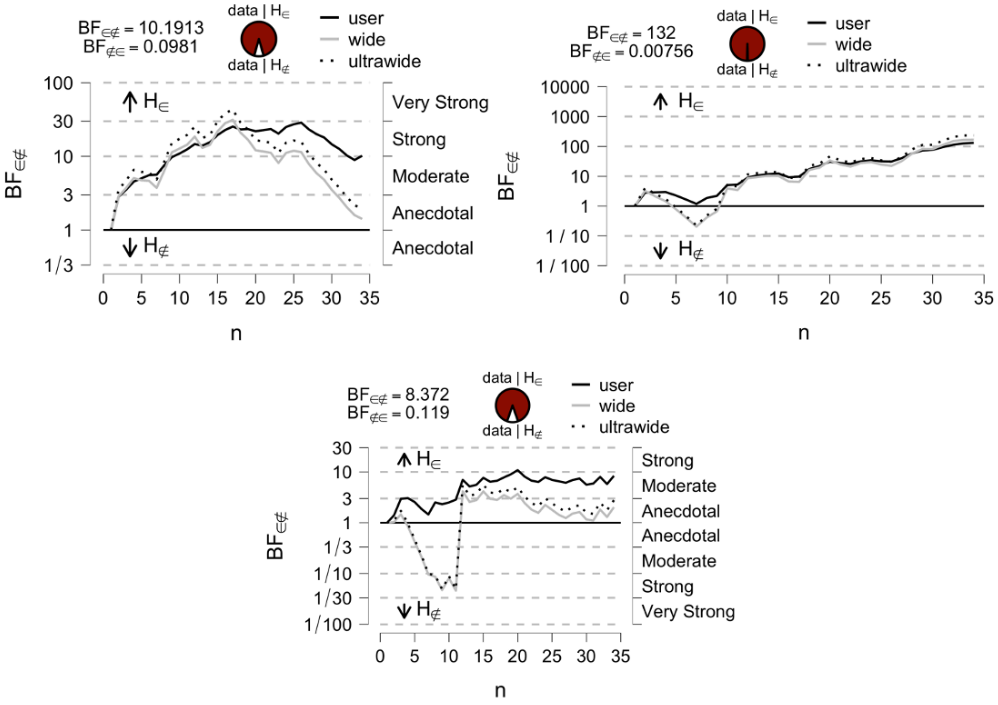


Figure 3.6. Equivalence test sequential analysis and robustness plots. Strength of evidence is shown as data accumulates over a range of priors. Plots are shown for trials 4-5 (top left); trials 5-6 (top right); and trials 4-6 (bottom middle).

3.4 Discussion

Despite a rise in interest in using VR in sports such as soccer (Thatcher et al., 2021), no studies have investigated or questioned the need to familiarise participants to a VR task before collecting data. Therefore, the aim of this study was the quantify the learning effects during familiarisation trials of a soccer-specific VR task. The main findings of this study were (1) evidence for a learning effect as shown by an increase in total goals between trials 1 and 2 (Figure 3.4), and (2) familiarisation might be achieved by the fifth trial, indicated by evidence of equivalence between trials 5 and 6 (Figure 3.5).

The learning effect associated with VR tasks has been well documented in non-sport settings (Bartlett et al., 2020; Janse et al., 2013). Moreover, our observation of a learning effect in the present study is comparable to that reported in learning curve studies where a substantial improvement in performance is observed in the first two trials (Bartlett et al., 2020; Gallagher & Satava, 2002; Janse et al., 2013; Jokinen et al., 2020). Learning to perform a task in VR provides the user with a unique challenge, as not only is the user learning the specific goals of the task (as they would in the real world), but users are also faced with learning challenges that are: (a) irrelevant to the aims of the task, and (b) unique to VR. For instance, participants had to learn how to control and pass a virtual soccer ball that provided no haptic feedback to the user. While a study by Harris et al. (2021) demonstrated that users' perceptions of controlling a golf ball in

VR was no different to controlling a golf ball in the real world, studies from the surgical literature have highlighted the importance of VR with haptic feedback for procedures where force applications are necessary (Alaker et al., 2016). Further, as VR is a computer simulation, and therefore the laws of physics can not be directly applied to the VR soccer task, it is possible that the ball may have responded differently to the way a player would interact with a ball in the real world. Other learning challenges include the potential for VR to cause perceptual strain on the user, as had been shown during the completion of a VR golf putting task (Harris et al., 2021).

Our study hypothesised that familiarisation would be evidenced by a plateau in performance between trials 4 to 6. Our finding of the need to complete multiple trials of a VR soccer passing task before a plateau in performance is reached is consistent with VR medical surgery literature that has explored the learning curve. For instance, a minimum of five trials were required in a VR hip arthroscopy simulator for a plateau in performance to be reached (Bartlett et al., 2020), whereas up to eight trials were necessary in a VR salpingectomy simulator (Bharathan et al., 2013). However, the evidence to support equivalence between trials 4-5 (strong) and 4-6 (moderate) were not as convincing as that for equivalence between trials 5 and 6 (extreme) (Figure 3.5). Given our study only managed to demonstrate strong to extreme evidence for equivalence between the last two trials, it is unknown if further trials would have affected the outcome of this research. For example, the inclusion of a seventh trial may have demonstrated a continued plateau in performance as shown in trials 5-6, strengthening the argument that familiarisation to the task had been established. In contrast, further improvement in performance would discredit that familiarisation was reached between trials 5-6. Similar arguments have been made by Bartlett et al. (2020), who also established familiarisation in the final two trials of a study.

3.4.1 Strengths and Limitations

Several strengths and limitations are presented from this study. First, our study was completed in a professional soccer club training ground, with participants completing the familiarisation trials on a regular training-day. As such, it is hoped that our findings can be generalised to professional soccer clubs currently using or planning to use VR. However, this approach did not allow us to thoroughly standardise variables that could affect task learning such as caffeine intake; room temperature; or circadian effects (Foskett et al., 2009). While this can be considered as a limitation, our findings reflect how we expect VR to be used by practitioners and therefore shows high ecological validity. For the purpose of ecological validity and due to previous research in VR golf highlighting the low physical demands compared to a real-world equivalent (Harris et al., 2021), a warm-up was not included in the present study. We consider this a limitation, as previous research supports the benefit of physical activity on perceptual-cognitive abilities such as visual search and information processing (Lambourne & Tomporowski, 2010). It is possible that if the VR task provided sufficient physical stimulation, the first task could have acted as a warm-up, facilitating an improvement in performance as seen in the second trial. Unfortunately, no information is available on the physical demands of a VR soccer-specific task. Future research should look to investigate the effects of a dynamic warm-up on VR performance. A study delimitation that needs to be considered in our study design was the width of the equivalence zone used in the statistical analysis plan (section 3.2.5). Our justification was based on pilot data collected using the VR system and a logical rationale as to what we considered equivalent in soccer - a two-goal difference between trials. That said, our rationale was subjective, and a different equivalence zone criterion could have changed the results in this study. For example, changing the equivalence zone to a standardised effect size of .5 either side of 0, would have changed the BF_{10} in trials 4-6 from 8.4 (moderate evidence) to 19.1 (strong evidence), suggesting stronger evidence for a plateau in performance being achieved.

3.5 Conclusion

Our study demonstrated sufficient evidence for a learning effect early in the repeated trials. Further, the findings indicated that up to five trials might be necessary for performance to plateau in a VR soccer passing task. Future studies examining VR with professional soccer players should consider how a lack of task familiarisation could lead to systematic error in the data and that several familiarisation trials (in the region of five) should be considered.

Chapter 4 The efficacy of a soccer virtual reality measurement system: A comparison of soccer skill performance in real and virtual worlds

4.1 Introduction

Ball passing is a skill in soccer, and is defined as the transfer of the ball from one player to another of the same team (Aquino et al., 2017). The development of this skill is fundamental, as it allows for the team to keep possession (a factor that is linked positively to match outcomes) and is a more efficient way of moving the ball around the defending opposition than dribbling (Collet, 2013). The importance of this skill is put rather simply by the late Johan Cruyff who famously quoted "you can't win without the ball, if we have the ball, they can't score", a reference linked to his style of high possession/passing style of coaching ('You can't score if you don't shoot,' 2016). As such, soccer coaches and scouts understand the importance of identifying (and developing) players that are competent at ball passing (Larkin & O'Connor, 2017). However, quantifying this skill in players through observations and analysis of a soccer game, is complicated as the number and success of these passes is likely to be affected by several factors, such as playing position, and team formations (Carling, 2011). For this reason, researchers have designed soccer passing tests that can be conducted in a controlled environment, such as the Loughborough Soccer Passing Test (LSPT) by Ali et al. (2007), or the short and long passing tests from the FIFA Medical Assessment and Research Centre (F-MARC) battery (Rosch et al., 2000). However, it is argued that such tests lack ecological validity, by the fact that passing is assessed in situations that are not reflective of the sport (e.g., passing a ball against a board). For this reason, professional soccer has begun exploring the use of VR due to the capability of producing ecologically valid environments (e.g., passing to a team-mate in game situations) that can be tested under controlled and repeatable conditions.

Virtual reality has attracted attention in sport, primarily in training to increase athletes' proficiency in their sport, as it allows for the creation of safe and repeatable tasks to be carried out in an artificial, virtual environment (Thatcher et al., 2021; Wood et al., 2021). Ultimately, when used for the purpose of training, the typical objective of VR is to provide a stimulus that results in a transfer of skill and/or performance to the real-world. For example, studies in baseball (Gray, 2017), table tennis (Michalski et al., 2019; Oagaz et al., 2022) and darts (Tirp et al., 2015) have demonstrated this transfer by assessing the changes in real-world skill following a VR intervention, and comparing these to a real-world training or control (no training) group. However, training is just one aspect of how VR can be used, and certain simulators have demonstrated VR's measurement capabilities (i.e., its use as a test). For instance, VR has been used in surgical education to test trainee surgeons during routine operations (Gurusamy et al., 2008). This is important as it allows stakeholders to assess the competency of individuals before they progress to real-world situations.

In the context of sport, performance tests hold a number of purposes which include assessing the effectiveness of an intervention (Russell et al., 2011), to gather objective evidence of an athletes' strengths and weaknesses (Robertson et al., 2015), to provide diagnostic information related to the health of an athlete (Stølen et al., 2005), or to identify talent (Robertson et al., 2017). It is generally agreed that a performance test displays evidence of measurement properties that match the test's content and purpose (i.e., does the test intend to identify talent, or evaluate the effect of an intervention (Robertson et al., 2017). Validity, which is defined as the degree to which a test measures what it purports to measure (Lohr, 2002) is important to establish, because without it the results of a test can be misleading and consequential, such as when decisions are made from the basis of a test score (Im et al., 2019). That is, without establishing test validity, the test outcome(s) could be false (Patrick & Beery, 1991). For example, Gray (2017) included a VR baseball hitting test as part of an intervention study to understand the effects of VR baseball training on virtual and real-world baseball performance, yet provided no information on the validity of the VR test. It is therefore difficult to determine whether the test scores from the study accurately represented what the test intended to measure (e.g., baseball hitting accuracy.) Evidence for validity can be acquired in multiple ways, such as through inherent characteristics of the test (content validity), the test's relation to a criterion (e.g., predictive or concurrent validity), or its relation to a construct (e.g., convergent or divergent validity) (Cronbach & Meehl, 1955). Studies exploring the validation of VR have often used construct validity, notably the known-group differences techniques whereby groups that should theoretically differ in the real-world are shown to differ in VR. For instance, Harris et al. (2021) reported that experienced golfers performed better at a VR golf putting task than novice golfers.

To date, only two VR soccer validation studies exist, with both using a construct validity approach of the known-groups difference technique. Wood et al. (2020) investigated the construct-validity of four VR soccer tests by comparing performance between three groups of soccer players who were theoretically expected to differ in their soccer-related abilities (professionals, academy, and novices). The study found that performance scores in three out of the four tests could be differentiated between level of player, as could an overall diagnostic score calculated from an algorithm using the four performance tests. The authors concluded that the results demonstrated construct validity and that the system was suitable for use in professional soccer. Similarly, Wirth et al. (2021) examined visual exploratory activity through evaluating head motion characteristics from the head mounted display across 18 VR soccer scenarios, between two groups of soccer players (low experienced and experienced). The results showed statistically significant differences (albeit without including a measure on the magnitude of difference, e.g., an effect size) between the two groups for several head turn measures such as head turn frequency and excursion, and that such findings demonstrate the validity of the measures captured through VR for assessing the visual exploratory activity of soccer players. However, it is important to note the limitation in both studies methods and how this can lead to a misinterpretation of the results. For instance, both studies failed to include a measure of uncertainty (such as confidence intervals) when reporting the differences between groups and relied entirely on the point estimate in their interpretation of the data.

To highlight this issue, Wood et al. (2020) only reported large effect sizes (*d* = 0.98) for differences in overall diagnostic scores between academy and novice players, whereas the inclusion of a 95% CI would have shown that the true difference in the population could lie anyway between *d* = 0.28 (small effect size) to *d* = 1.7 (large effect sizes). Further, when applied to one of the individual tests 'shoulder sums', what is reported as a significant and large difference (*d* = 0.8) between professional and academy players (indicating the test can differentiate between the ability of a player) is broadened to a range of trivial (*d* = 0.1) and an even greater effect size (*d* = 1.5). In other words, the true population difference for that specific test could be non-existent or even larger than the study has reported. Similarly, this issue is demonstrated in the study by Wirth et al. (2021) who reported that low experienced soccer players tended to focus significantly longer on the opposition than the experienced players (*d* = 0.93), suggesting that the latter deploy different visual exploration strategies such as focusing more on team-mates and the ball. However, a calculation of effect sizes with a measure of uncertainty would show that the true population difference falls between trivial and very large (*d* = 0.93, 95% CI = 0.15 to 1.76). The large variability in true population differences makes it difficult to interpret the results in both studies and is likely to be a consequence of the small sample size reported. For instance, Wood et al. (2021) had equal numbers in each group (n = 17) which was based on an *a priori* examination of effect sizes reported from a similar study (Harris et al., 2021) and not from an *a priori* sample size calculation, whereas Wirth et al. (2021) sample (low experienced, n = 17; experienced, n = 11), was based entirely on convenience without any indication of sample size calculations.

Although some evidence exists to show the validity of soccer VR, it is clear from reviewing the two papers outlined above that further studies are required that provide clear justifications on sample size parameters, and take into account any uncertainty when making interpretations on the data (Abt et al., 2020; Greenland et al., 2016). Further, it is evident from these studies (and other VR validation studies) a reliance on construct validity techniques exists, whereby evidence is provided for interpretation/validity arguments (in these cases, differences in performance based on soccer playing level) without showing clearly how it supports the validity of a test (Colliver et al., 2012). Both studies suggest evidence has been met for validity by differentiating between groups, yet this evidence does not directly establish with confidence that each of these tests measures what it purports to measure (Colliver et al., 2012), such as visual exploratory activity as suggested by Wirth et al. (2021). Visual exploratory activity assessed using VR might be more appropriately validated by comparing the VR scores to visual exploratory activity that is measured in the real-world (criterion validity). Lastly, the type of validation method should have a sound theoretical framework in place to justify why the method was chosen (Impellizzeri & Marcora, 2009). Unfortunately, neither study sufficiently rationalised why they choose their method of validation, in relation to the aims of the VR task. Although the studies by Wood et al. (2020) and Wirth et al. (2021) attempt to show VR's ability to discriminate between groups, and so might have use in talent identification, such methods (i.e., known-groups validity) are insufficient for evaluative purposes. As such, while it is encouraged that more studies adopt the known-group difference approach if the purpose is to establish scores at different stages of soccer development (e.g., playing level or age), other validation approaches are required if the purpose is to evaluate a skill in VR in relation to the same skill in the real-world.

According to Harris, Bird, Smart, Wilson, and Vine (2020), validation is just one factor required when examining the suitability of a VR, and proposed a framework to guide researchers exploring the use of the technology in sport. Although intended for assessing VR used for training purposes, the framework suggested by Harris et al. (2020) involves establishing a taxonomy of evidence for fidelity and validity in the VR simulator that would facilitate transfer of skills to the real-world. For instance, within these frameworks authors have suggested assessing the subjective experience of 'presence', which is defined as the illusion of being in the VR environment and is often referred to as the psychological sense of 'being there' (Slater et al., 1995). Understanding if the VR environment has induced a sense of presence is important as the user is more likely to engage with the environment and respond similarly to stimuli as they would in the real world (Slater & Sanchez-Vives, 2016). This is important, as a VR test that demonstrates a low level of presence might mean the user does not engage with the task and that the outcome measure (e.g., a measure of soccer skill) does not reflect their true capabilities. Other examples in the frameworks include the assessment of psychological fidelity, which is defined as the degree to which the VR environment replicates the perceptual and cognitive demands of the real world (Harris, Bird, et al., 2020). One way of assessing the psychological fidelity is by comparison of subjective effort between the virtual and real world. For instance, the simulator task load index (SIM-TLX) was specifically designed to assess and compare multiple sources of subjective demand (e.g., physical and mental) between virtual and real-world tasks (Harris, Wilson, et al., 2020). That's not to say that the subjective demand needs to be identical for the VR task to be a good representation of the real world. In fact, under certain circumstances it might be desirable for subjective demand to be lower in VR than in the real world for applications such as rehabilitation, where reducing physical intensities could be favourable (Thatcher et al., 2021). However, an awareness of any substantial differences between conditions could be useful to understand why an individual could perform well under one condition (e.g., VR) but not under another (e.g., real world).

To that end, the purpose of this study was to investigate the efficacy of using VR to assess soccer skill performance. First, the study assessed performance scores of tests completed in the virtual world versus those performed within the real-world with the specific aims being: 1) to examine the difference in soccer passing skill in VR versus the real-world; 2) to examine if passing skill performance in VR is equivalent to that in the real world; and 3) to examine the agreement between passing skill in VR and in the real world. For this part of the study a Bayesian sequential parameter estimation approach was taken and therefore, null hypothesis tests were not conducted or reported. Second, the study took an exploratory approach to examine the suitability of VR for use as a testing tool by using existing frameworks (Harris, Bird, et al., 2020). These frameworks aimed to 1) explore the psychological fidelity by comparing the subjective task demands between the VR and real-world conditions, and 2) to explore the sense of presence experienced by the participants.

4.2 Methods

4.2.1 Pre-registration

The study question, sampling plan, data collection procedure and data analysis plan were pre-registered on the Open Science Framework (https://osf.io/krdc7). Any deviation from the pre-registration is clearly outlined in the relevant section of the method.

4.2.2 Participants and sample size calculation

Twenty-nine (male = 28, female = 1) soccer players (recreational = 19, academy = 10) (mean (SD), age: 22 (5) years, stature: 176 (6.9) cm, and body mass: 73.1 (6.9) kg) were recruited between August and November 2021, via e-mail that included an attached poster (Appendix 5). In our pre-registration we set a minimal sample size of 30 participants, so in this aspect we have deviated from our plan set *a priori*. Participants were provided details on the aims of the study verbally, and through a video (https://youtu.be/toGC-I7kDi0) and gave written informed consent before data collection began. In relation to sample size, the study used a Bayesian sequential analysis approach where the precision of the population estimation (the credible interval) was the goal rather than a frequentist approach of power (Kruschke & Liddell, 2018). Further details on the Bayesian sequential analysis including our justification for stopping data collection are discussed in the statistical analysis section. Participants were excluded from the study if they had an injury that would prevent them from taking part in regular soccer training. Additionally, participants were required to be between the ages of 16 and 37. This age range was chosen based on research exploring 14500 match observations across three soccer seasons, where the mean (SD) age of players was 25.5 (3.7) (Sal de Rellán-Guerra et al., 2019). For this study, 3 standard deviations were applied either side of the mean (thereby capturing ~99% of the population ages) to give an inclusion zone of 14.4 to 36.6. As soccer players do not commence full-time training until aged 16, the inclusion zone was increased to this minimum age. The Faculty of Health Science research ethics committee at the University of Hull approved the study (REF: FHS331), which conformed to the Declaration of Helsinki.

4.2.3 Virtual reality instrumentation

Participants wore a head-mounted display (Valve Index Headset, Valve Corporation, Washington, United States) and trackers (VIVE tracker, HTC corporation, Taoyuan City, Taiwan) which were securely fixed to the participants' trainers (Appendix 1). The VR soccer skill task was custom-made by a VR software company (Rezzil Inc, Manchester, UK) following guidelines as set out by the lead researcher. This included an outline of the drill being presented to the VR software company, that included specific instructions on the dimensions, and aesthetics of the task, and how the participant would interact with the environment (e.g., ball passing). These details were taken from the passing test described by Wilson et al., (2016) (see section 4.2.5 for more details). Several iterations of the software were designed, that each included a process of feedback being provided to the company to allow further refinement to take place. Feedback mostly consisted of improvements to the efficiency of the task (such as the addition of boundary lines, discussed in section 4.2.7, that allowed the ball to respawn quicker), or removal of bugs that were found during testing. Both hardware and software were powered by a custom-built computer (Processor: i7-8700K 6-Core 3.7, GPU: 1080 Ti, Memory: 16GB, OS: Windows 10 Pro). The VR operating space was calibrated as a 2 x 2 m rectangle using SteamVR (Valve Corporation, Washington, United States). Two lighthouse stations (Base station, HTC corporation, Taoyuan City, Taiwan) tracked three-dimensional movement, and were fitted at heights of 180 cm at both sides of the participant ensuring tracking was maintained throughout the procedure. It is noted that the room set up for calibration differed to that of Chapter 3, although this does not affect tracking performance, and was simply due to the parameters of the room and the nature of the installation (e.g., permanent vs temporary usage). To ensure participant safety while using the

VR system, participants completed the task within a playing circle (radius of 1 m). Leaving the circle resulted in the head-mounted display screen being dimmed, providing participants with a visual warning that they had left the safe playing area. The head-mounted display screen reverted to full contrast on returning to the playing circle. Participants could still interact with the VR ball if they stepped out of the playing circle, although they were encouraged to stay within the circle throughout each trial. Data was still collected if participants left the circle. Audio from the software was delivered through speakers built into the head-mounted display.

4.2.4 Protocol

Participants completed two conditions: a real-world soccer test, and an equivalent VR soccer test. Counterbalancing was used to determine which condition participants completed first, and participants completed both conditions on the same day, consecutively, and within a 40-minute period. Immediately after completing each condition, participants completed the SIM-TLX, and after completing the VR condition, participants completed a sense of presence questionnaire. Finally, participants took part in a semi-structured interview to discuss their experience of the real-world and VR passing tests.

4.2.5 Passing Test

To compare soccer skill between the real-world and VR conditions, a short passing test from the University of Queensland Football Skill Assessment Protocol (UQFSA) was used (Wilson et al., 2016). This comprised of a rebound 135 skill test which was devised to measure a player’s ability to receive a ball and execute a subsequent pass with speed and accuracy (Wilson et al., 2016). This test was chosen due to, (1) the ease of replicating the test in VR, (2) the established construct validity in the real world, and (3) the established reliability in the real world (intra-class correlation coefficient, 0.94) (Wilson et al., 2016). However, for the present study an updated version of the rebound 135 test was used by following an unpublished protocol devised by the first author of the UQSFA. This version was used over the original version because it was easier to implement (i.e., time penalties for technical errors were removed), required less time to complete, yet was shown to be just as reliable using the preliminary data shared by authors of the UQSFA (Wilson et al., 2016).To ensure systematic error was minimised, four familiarisation trials were completed immediately prior to the recorded trials in both the real world and the VR condition. Using findings from Chapter 3 which explored the learning effect of a soccer test in VR, it was considered that four familiarisation tests would be sufficient to fully accustom participants to the protocols and negate any possible learning effects. To standardise trials between the two conditions, four familiarisation trials were also performed in the real-world condition. After familiarisation, two recorded trials were performed in each condition, with the best trial score used for analysis. A 2 min rest period was set between each familiarisation and recorded trial.

4.2.6 Real-world condition

The real-world condition passing test took place in an indoor biomechanics laboratory (Figure 4.1). Two rebound boards (2 m long x 0.6 m high; Diamond Football Co Ltd, Molesey, UK) were positioned at an angle of 135° to each other, with the central focal point (where the participant stands) 4 m from each rebound board. To account for a limitation of the playing surface which caused the ball to rise off the floor and bounce back towards the participant, four 20 kg plates (Eleiko, Halmstad, Sweden) were placed under each of the four board holders, resulting in the boards being at an 80° vertical angle. Before starting the first familiarisation trial, information was provided on the aims of the test, including a specific passing technique which each participant was required to follow. The study used a passing technique described in more detail by Wilson et al. (2016), which simulates passing when pressure from an opponent is from the direction that the ball is received, thereby requiring the participant to turn with the ball to protect it. Each trial started by passing the ball (Size 5, Strike football, Nike, Oregon, United States) to rebound board 'a' with their left foot. As the ball returned, participants received the ball with their right foot, and turned with the ball to set up a pass with their right foot. Participants passed the ball with their right foot to rebound board 'b', before receiving the ball with their left foot and turning back to rebound board 'a' to set up another pass with their left foot. For each trial, participants were given 45 s to complete as many passes as possible using this technique, with the timer commencing when contact was first made with the ball. The total number of passes using this technique was taken as the participant's score, and for a pass to count it had to hit the rebound board and return to the participant's foot. If the ball missed the rebound board, the timer did not stop but a new ball was rolled in for the participant to continue.

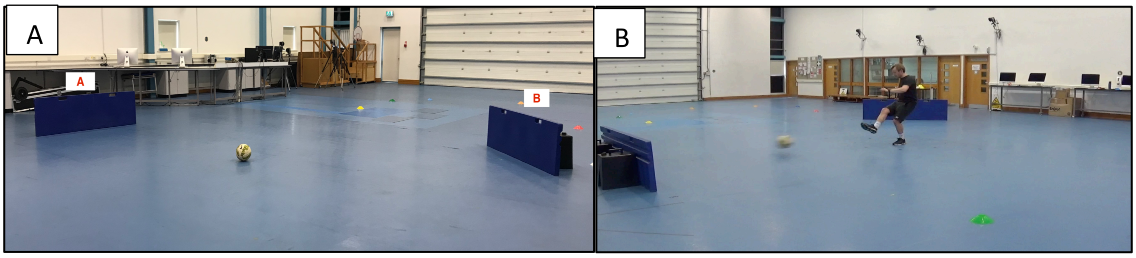


Figure 4.1. (A) Set up of rebound board 'a' and 'b' relative to the central focal point. (B) Participant kicking the ball to rebound board 'a'.

4.2.7 Virtual reality condition

The VR environment and set up are outlined below (Figure 4.2). Participants were first fitted with trackers strapped to the participants' trainers and provided instructions regarding fitting the head-mounted display. Participants confirmed that the head-mounted display did not cause discomfort at the time of fitting or throughout testing. Calibration of foot trackers was completed following the guidelines set by the VR software (Rezzil Inc, Manchester, UK) which included entering the participant's regular shoe size into the software. Calibration took place while the participant wore the head-mounted display, with instructions provided visually and verbally by the software. The calibration protocol required the participant to stand with their feet shoulder width apart within a virtual footprint positioned on the floor. Calibration took approximately three seconds, and once completed, participants were presented with virtual soccer boots within the head-mounted display. Participants were asked to confirm if the virtual boots were aligned with their own foot position and were at a height that represented where their feet would be in the real world. Once calibration was finished, participants were taken to a virtual indoor soccer pitch. On the virtual soccer pitch were two rebound boards, set at the same dimensions and the same distance from the participant as described in the real world, except for them not being placed at an 80° vertical angle. Additionally, two boundary lines were set 0.3 m either side of each rebound board, running parallel towards the participant. These lines acted as boundaries that the ball must stay within, with the ball resetting automatically to the start position if the ball crossed the line or was kicked over the rebound board. As such, participants would always start by kicking a reset ball to rebound board 'a' with their left foot. In each trial, participants were required to complete as many passes as possible in 45 s, using the same passing technique used in the real-world condition. In contrast to the real-world condition, the timer was stopped if the ball was kicked over the rebound board, or beyond the boundary lines. This was due to a limitation in the software meaning the ball would not always reset to the start position instantly but could take up to 3 s. In such a case, participants were instructed to continue with the trial as soon as the ball appeared.

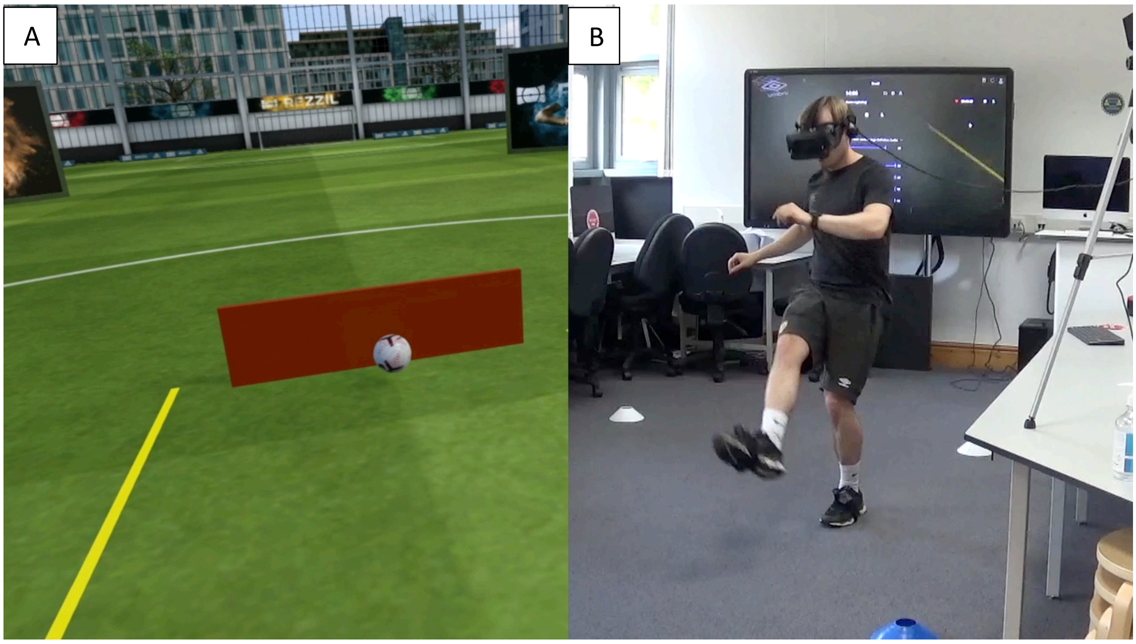


Figure 4.2. (A) Virtual reality environment depicting a virtual ball approaching a rebound board. (B) Room set up showing participant wearing a head mounted display and foot trackers, having just kicked a virtual ball.

4.2.8 Subjective task demand

To measure the subjective task demand, the SIM-TLX was used (Harris, Wilson, et al., 2020). Whilst the term ‘load’ was retained when referencing the name of the measurement index because it is the name of the measure created by the authors, researchers should be aware that the use of the term ‘load’ is contested (Staunton et al., 2022). As such, we refer to ‘load’ as subjective task demand throughout. The SIM-TLX has been validated as a multidimensional measure of subjective task demand for assessing VR environments (Harris, Wilson, et al., 2020). Participants rated the level of demand they experienced on nine, 21-point Likert scales that reflected different task demand dimensions: physical demands; mental demands; temporal demands; frustration; task complexity; situational stress; distractions; perceptual strain; and task control. Likert scales were anchored between very low (score of 5) to very high (score of 100). Participants did not complete the pairwise comparison part of the SIM-TLX, which is used to assess the relative importance of each dimension. This part was omitted from the study to reduce the time burden placed on participants. Furthermore, this method of removing the pairwise comparisons has been used with the NASA-TLX (Hart & Staveland, 1988) to create what is known as the Raw TLX, with research showing it to be equally as effective measure of task demand.

4.2.9 Sense of presence questionnaire

To measure the sense of presence, participants completed The Independent Television Commission – Sense of Inventory (ITC-SOPI) questionnaire (Lessiter et al., 2001). The ITC-SOPI is a valid and reliable method of measuring presence across a range of different media types, including VR (Lessiter et al., 2001). The questionnaire consisted of 43 Likert scale questions, each having five response anchors labelled from strongly disagree to strongly agree. Questions were divided across four constructs that are believed to contribute towards the sense of presence: sense of physical space (18 questions), engagement (13 questions), ecological validity (5 questions), and negative effects (6 questions). Sense of physical space indicates a sense of being in the physical environment, including the ability to interact and control aspects of the environment. This construct is closely linked to the definition of presence, which is the sense of 'being there' (Slater et al., 1995). Engagement identifies the feeling of psychological involvement and enjoying the content, whereas ecological validity indicated the tendency to perceive the virtual environment as lifelike and real, including their own movements and involvement. For these three constructs, means were created and a score of 1 indicated the lowest level of presence, and 5 considered the highest. Negative effects were related to adverse physiological reactions to the virtual environment such as headaches, nausea, and dizziness. Although less relatable to the first three constructs, presence has been shown to be inversely related to cybersickness and so should be considered when assessing VR applications (Weech et al., 2019). Means of 1 indicate that no negative effects were experienced, whereas 5 would indicate the maximum amount experienced.

4.2.10 Semi-structure interviews

Qualitative interviews were conducted so that participants could describe their experience of completing the real-world and VR tests. Interviews consisted of two prescribed questions that were asked to all participants. These aimed at understanding the similarities and differences between the two conditions, and the influence of haptic feedback on virtual reality task performance. These questions were: “What are your thoughts on the virtual reality soccer skill test you've just completed compared to the real-world version?”, and “Did the haptic feedback (touch of the ball) influence your ability to perform the soccer skill test in virtual reality?”. Additional questions were asked depending on how participants responded. Interviews lasted between 3-7 minutes. Recorded interviews where later transcribed, verbatim, and thematic analysis was conducted.

4.2.11 Statistical analysis

A Bayesian sequential parameter estimation approach was used (as opposed to hypothesis testing) consisting of descriptive and inferential analyses. The inferential analysis consisted of *a priori* confirmatory analyses (as outlined in the pre-registration) with the option of conducting exploratory analyses following data collection. Data analysis was conducted using R (4.1.0) (RStudio Team, 2020) and JASP v.0.14.1 (JASP team, 2021).

4.2.12 Confirmatory analysis

To analyse the difference between conditions (virtual reality and real world), Bayesian regression models using 'Stan' package (brms) were used (Bürkner, 2017). The first confirmatory inferential analysis consisted of comparing the means of the two conditions. Because the outcome variable (number of passes) is discrete data (counts), two Bayesian generalized multilevel models (Gaussian and a Poisson model) were constructed. For each model, Leave-One-Out Cross-Validation (LOO) criterion was applied, and posterior predictive checks produced to compare the observed data to simulated data from the posterior predictive distribution. To determine the best models in terms of the out-of-sample predictive performance, Leave-One-Out Cross-Validation was used to estimate the Expected Log-Predictive Density and Pareto Smoothed Importance Sampling (PSIS). Estimated marginal means were obtained from our model using the package ‘emmeans’ (Lenth, 2021), and log-estimates back-transformed if the Poisson model was used. The estimated marginal means (EMM) and highest density interval (HDI) intervals are reported for each condition and for the differences between conditions.

The second confirmatory analysis examined if the conditions (virtual reality or real-world) are 'different' or 'equivalent' to each other, by setting a region of practical equivalence (ROPE) using the package ‘bayestestR’ (Makowski et al., 2019). The ROPE was set as a two-pass difference either side of the point estimate (between -2 and 2). The rationale for the ROPE was based on preliminary unpublished data (by independent researchers) collected during repeated trials of the real-world passing test. Based on the repeated measures means (trial 1 and 2) and the pooled standard deviation, a two-pass difference was calculated to equate to a standardised effect size of 0.4 either side of 0. Additionally, a two-pass difference was determined as an acceptable parameter of equivalence because a passing difference of this magnitude would probably not be considered worthy of intervention by a coach. Further, to accept equivalence between the two conditions, the 95% HDI should fall within the ROPE (Makowski et al., 2019).

The third confirmatory analysis examined the overall distributions of the model as opposed to just the means. As the distributions for the two conditions could differ outside of their central tendency (i.e., the tails of the distributions), a hierarchical shift model was used with regressions calculated for each 10th quantile between 10 and 90 (9 quantile regressions). For the hierarchical shift model the same ROPE was applied as outlined above to assess if the two conditions were equivalent at each quantile regression. The fourth confirmatory analysis examined the agreement between test scores in each condition using Bland-Altman 95% limits of agreement analysis (Bland & Altman, 1986). The analysis was conducted using the package 'blandr' (Datta, 2017), with the mean bias examined to see if it differed from 0 using a Bayesian one-sample t-test using JASP (0.14.1). For this one-sample t-test, Bayes Factors (BFs) were used which quantify the relative predictive performance of the null hypothesis (H_0) compared to the alternative hypothesis (H_1), and a classification scheme as set out by Lee and Wagenmakers (2013) was used to aid interpretation. In this classification scheme, a BF_{10} (which indicates evidence for H_1) of between 1 and 3 is indicative of anecdotal evidence, between 4 and 10 is indicative of moderate evidence, between 11 and 30 is indicative of strong evidence, between 31 and 100 is indicative of very strong evidence, and > 100 is indicative of extreme evidence. Limits of agreement (LOA) were reported with 95% credible intervals, and heteroscedasticity was examined through linear regression.

4.2.13 Exploratory analysis

The sense of presence questionnaire and SIM-TLX were explored visually using the GGPlot 2 package (Wickham, 2016) with descriptive data reported. Due to the small sample size from this study, we refrained from conducting inferential data analysis with the exploratory data. Transcripts from the semi-structured interviews were explored for themes manually using NVivo (QSR International Pty Ltd, 2020).

4.3 Results

4.3.1 Confirmatory analysis

Models (Gaussian and Poisson) were first compared to one another to assess which model best fit the data. The Expected Log-Predictive Density difference between the two models was -10.6 (4.6) in favour of the Gaussian model and so this model was used throughout the analysis. Posterior predictive checks confirmed that the observed data matched the simulated data (Appendix 6). For the first confirmatory analysis a difference between conditions was observed (EMM = -3.9, 95% HDI = -5.1 to -2.7) with the number of passes being greater in the real-world condition (EMM = 19.7, 95% HDI = 18.6 to 20.7) than the VR condition (EMM = 15.7, 95% HDI = 14.7 to 16.8).

For the second confirmatory analysis, a ROPE was set to examine if the two conditions were ‘equivalent’ or ‘different’ to one another, based on the criteria set. The posterior distribution showed that 100% of the data fell outside of the ROPE, and therefore could not be classified as ‘equivalent’ (Figure 4.3a). For the third confirmatory analysis a hierarchical shift model was used for each quantile along with the ROPE (Figure 4.3b). All nine quantile distributions fell outside of the ROPE, and we observed increased differences between the 40th and the 90th percentiles, such that the best performing participants had a greater difference between their real-world and VR scores than worse performing participants (those <40th percentile).

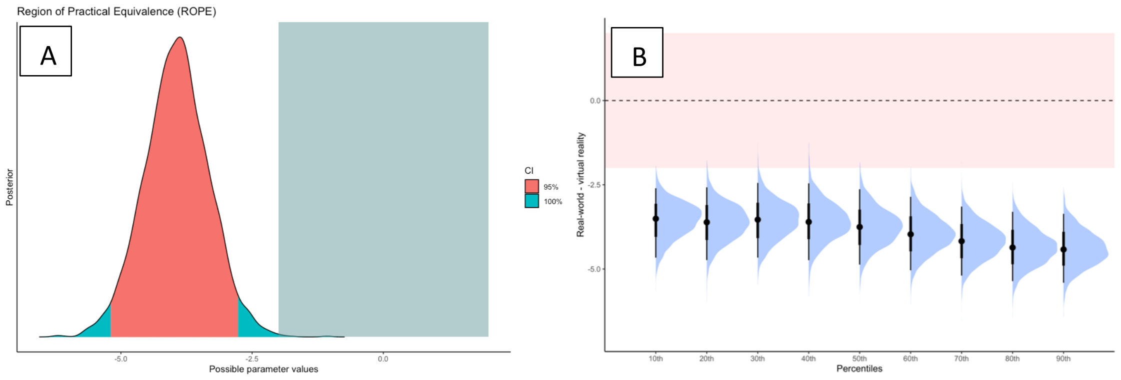


Figure 4.3. (A) The posterior distribution for the difference between conditions, relative to the region of practical equivalence. (B) The posterior distribution for differences between conditions at each quantile relative to the region of practical equivalence.

For the fourth confirmatory analysis, Bland-Altman plots were produced (Figure 4.4). The analysis showed that VR conditions differed to real-world conditions by -3.9 passes (difference 95% CI = -2.8 to -5), with an upper LOA of 2 passes (95% CI = 4 to 0) and a lower LOA of -9.9 passes (CI = -7.9 to -11.9). The Bayesian one-sample t-test showed extreme evidence that the mean bias differed from 0 (BF_{10} = 111014, *d* = 1.2 [95% CI = 0.73 to 1.73]).

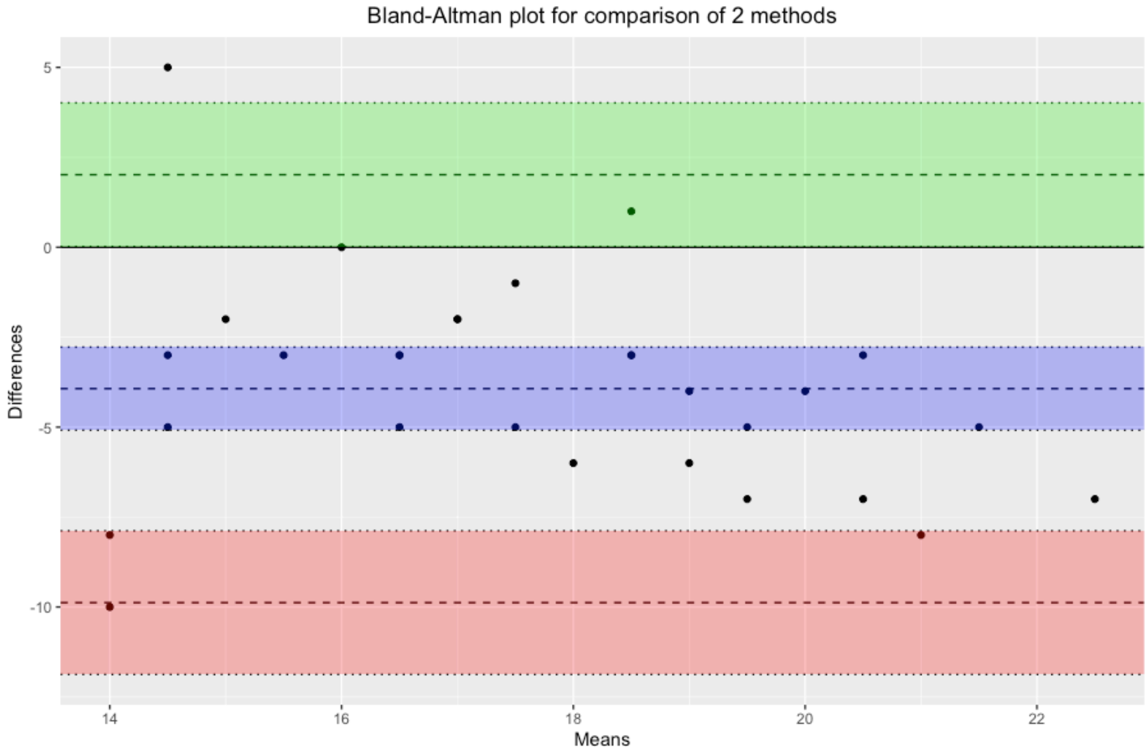


Figure 4.4. Bland-Altman plot for differences between the number of passes in the real world and virtual reality. Black dots represent differences in passes for each participant. The differences for the mean (purple), upper limit of agreement (green), and lower limits of agreement (red) are each represented by the dashed lines (95% CI = dotted lines). CI, confidence intervals.

4.3.2 Exploratory analysis

4.3.2.1 Task load dimensions

Comparisons were made between each dimension of the SIM-TLX between VR and real-world conditions (Figure 4.5). Across all dimensions of the SIM-TLX and in both conditions, large variability of participants scores are reported. Observational differences between conditions are accounted for in most dimensions for the SIM-TLX, with notably large differences for task control.

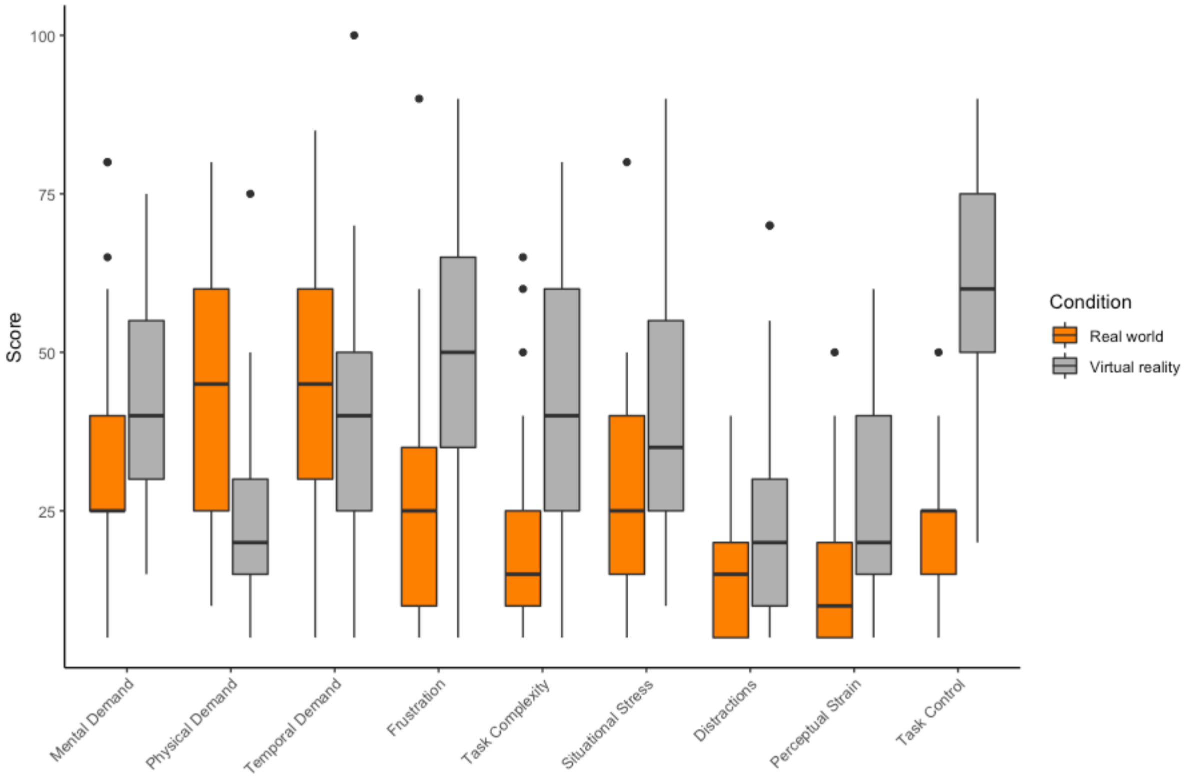


Figure 4.5 Grouped box plots for the nine subscales of task load for real-world and virtual reality tasks. Boxplots represent the 25th-75th percentiles, and the horizontal black line represents the median. Whiskers extend to indicate the most extreme data values, with dots beyond the whiskers considered as outliers (1.5 times the interquartile range from the edge of the box).

4.3.2.2 Presence

Participants' spatial presence (mean = 3.39, SD = .47, 95% CI = [3.22 to 3.57]), engagement (mean = 3.64, SD = .41, 95% CI = [3.48 to 3.8]), and ecological validity (mean = 3.48, SD = 0.6, 95% CI = [3.25 to 3.7]) factor scores were all considered to be in the high range (Figure 5). Negative effects factor scores were low (mean = 1.65, SD = .55, 95% CI = [1.44 to 1.86]) (Figure 4.6).

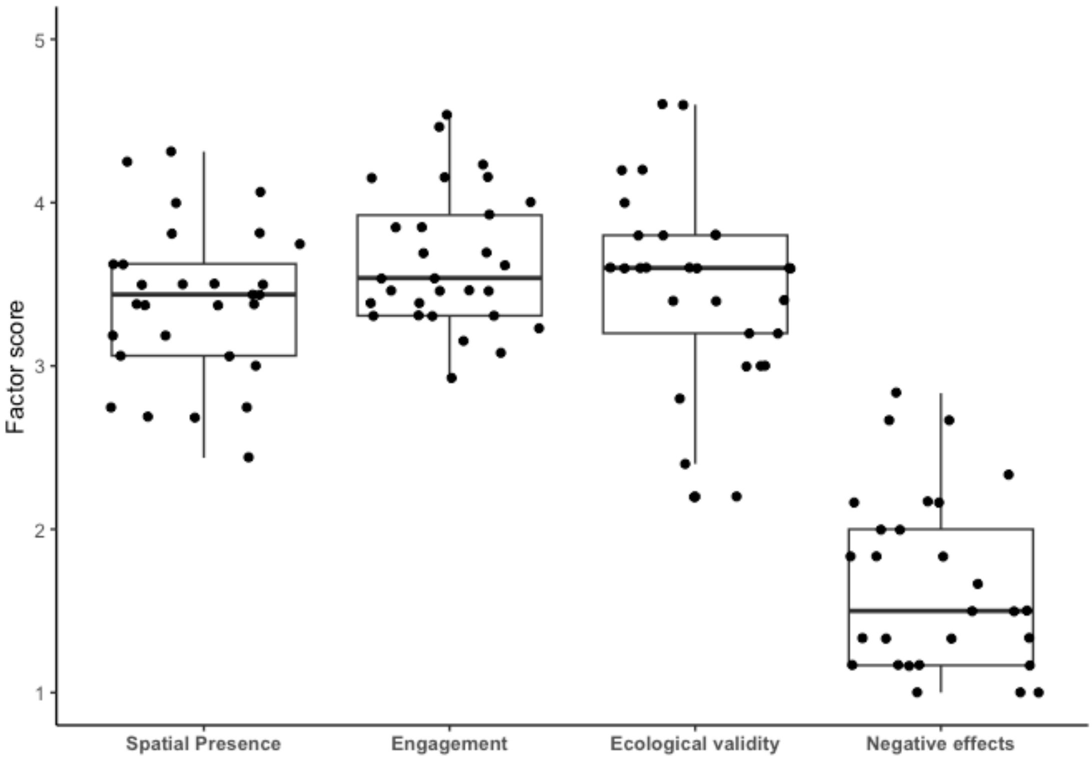


Figure 4.6 Presence scores for each construct of the ITC-SOPI questionnaire. Boxplots represent the 25th-75th percentiles, and the horizontal black line represents the median. Whiskers extend to indicate the most extreme data values, and each dot represents individual participant scores.

4.3.2.3 Semi-structured interviews

Following the interviews, several themes emerged that were used to explain differences and similarities between the two conditions, along with general opinions of the experience of using VR and its utility. Several participants identified differences between the two conditions, most notably the control of the ball which was reported as more difficult:

> “There's a really big difference in the amount of power that you use. Like, I couldn't kick a ball how I normally would because it would go flying because the weight of the ball isn't there. Whereas in the real life I can kick the ball to the board with the correct amount of power I want to put into it”

> “So the main differences from participating in was there was no control over the weight of the pass, where it was sort of less realistic, whereas in the real one I'm punching the ball and its coming back the way I want it too”

When asked if the haptic feedback of the VR simulator had any influence on the perceived performance of the task, participants identified that not feeling the ball may have contributed to the task difficulty:

> "Yes, I think not feeling the weight of the ball in virtual reality it puts you off a little bit as you don't actually know how much force to put on the ball when you play it to your back foot, whereas obviously in the real world you can feel the weight of the ball"

That said, comments on the same question alluded to a small number of participants identifying that although the feel of the ball wasn't there, the sound of the ball was helpful in completing the task:

> "I think one thing that makes it feel more real was the sound. Because I did notice that the sound differed depending on how hard you was hitting it [the ball]. So that gave it a bit of a guide"

> "it did feel like a realistic ball to kick. It's not haptic feedback but the sound as well on the ball helped"

When asked about similarities between the two conditions, participants acknowledged similar dimensions such as the distance and size of the rebound boards, and the size of the ball. Notably, several participants reported similar movement patterns between the two conditions, suggesting that participants responded biomechanically to a VR ball as they would a real ball:

> "The similarities were obviously the body shapes to receive and pass the ball, the thinking and thought processes were very similar"

> "Similarities, I think the movements to receive the ball and how I turned my body. I think that would be what any player would do. So I think movement wise, I think nothing changed as such"

4.4 Discussion

The aim of this study was to investigate the validity of a VR soccer passing test by comparing the performance to an equivalent test in the real-world. Furthermore, the study aimed to explore the construct of 'presence' while completing the VR condition and compare the psychological fidelity between the two conditions. The main finding of this study was that the number of passes completed in VR was lower compared to the real-world (EMM = -3.9, 95% HDI = -5.1 to -2.7), and that 100% of the differences fell outside the ROPE criteria that was set *a priori*.

These main findings are comparable to Harris et al. (2021) who reported that following a 3 m golf putting task in real and VR conditions, distance of the ball to the hole was lower in the real-world. In addition to assessing the absolute difference between conditions, our study reported that the number of passes completed between conditions was greater than a two-pass difference which was defined as the ROPE (Figure 4.3a). Several reasons are proposed in this discussion as to why differences were observed between the two conditions. First, our study identified that participants found controlling the environment in VR to be more difficult, as observed by the SIM-TLX dimension, task control (Figure 4.5). One explanation as to why the task was more difficult to control in VR, might have been due to the lack of haptic feedback produced when stopping or passing the ball. Haptic feedback refers to the feeling of 'touch', and is an essential component in VR surgical applications where applying force is necessary (Gani et al., 2022). In soccer, players often use terms such as 'weight' and 'hardness' to describe the feel of the ball while interacting with their feet, with many having a preferred style of ball to interact with (Thompsett et al., 2016). These ball properties could not be experienced in VR, meaning participants could not determine how much force to produce when stopping or passing the ball. Unfortunately, the VR hardware used on the foot could not produce haptic feedback, and to the best of the authors' knowledge, no other commercially available VR hardware would be able to produce haptic feedback on the foot. In contrast, Harris et al. (2021) reported no significant difference for task control in a 3 m golf putting test between their VR and real-world conditions, despite their VR condition also having no haptic feedback. These contrasting results could be due to differing visual strategies used across tasks, with the golf putting task allowing for long and sustained fixations on the ball and hole before contact is made with the putt, commonly referred to in the literature as quiet eye (Vickers, 1996). The soccer passing task did not allow for this opportunity as it intended to replicate passing ability when put under pressure from the opposition, meaning passes had to be executed quickly (Wilson et al., 2016). In competitive soccer matches, these types of passes would often see players fixating on team-mates or open space as the pass is executed (Aksum et al., 2020). It is plausible that task control might have been more similar between conditions if a passing test that allowed for longer and sustained eye fixations to be made on the ball and target was used, such as the 20 m passing test from the UQSFA (Wilson et al., 2016). Task complexity in the present study was also observably different to task complexity perceived in the study by Harris et al. (2021). Clearly, there is a difference between the closed nature of putting a golf ball that is performed in the participants own time and has no outside interference, to the open nature of the soccer task in this study which required anticipation of the rebounded ball in terms of its trajectory, the need to stop the ball, and pass in the opposite direction.

Another reason that task control was observably higher in VR could be due to the ball physics applied in VR not being representative of the real-world. Participants in the present study referred to the 'weight' of the pass not being there, causing the ball to travel faster and further than would be possible in the real-world. Because this study used a commercially available software that is currently in use within professional soccer (Thatcher et al., 2021), the physics of the VR ball were not questioned prior to the study. However, as shown by studies in table tennis

(Oagaz et al., 2022), and billiards (Haar et al., 2021), the validation of the ball and cue ball physics, respectively, in relation to the real-world is an important first step in VR that requires physical interactions with an object. For example, Haar et al. (2021) validated the physics of the cue ball in VR by merging the virtual and real-world environment so that interaction with the environment is simultaneously aligned. In other words, a real pool table and cue ball was linked to a virtual pool table so that interactions with the cue ball in VR corresponded to interactions with a cue ball in the real-world. The results showed with-in shot comparisons between conditions to be highly consistent for the direction of travel (Pearson correlation r = 0.99) and velocities (Pearson correlation r = 0.83). Future studies in VR soccer should consider validating the ball physics applied in VR.

The sense of presence questionnaire revealed that participants experienced a high sense of presence during the VR soccer passing task, as noted by high scores for spatial presence, engagement, and ecological validity (Figure 4.6). These findings are comparable with Le Noury, Buszard, Reid, and Farrow (2020) who also noted a high sense of presence for a tennis VR simulator. However, it is acknowledged that all three constructs of presence were observably higher in the study by Le Noury et al. (2020) and might be explained by the environment which included a scoreboard, crowds, and large buildings surrounding the court, and were purposefully added to enhance presence. These features may have added to the realism of the environment (ecological validity). Additionally, the tennis simulator involved participants engaging in a competitive match scenario, against a virtual opponent that could be controlled by the research team. Unlike the current study that did not include a competitive element, the competitive nature of the tennis task may have resulted in greater motivation and consequently enjoyment (engagement). Previous research on VR rowing has demonstrated greater motivation when the VR environment includes an on-screen competitor to challenge the participant (Parton & Neumann, 2019). Finally, it is important to consider the hardware set up in relation to the demands of the task. It is possible that the wire from the head-mounted display could have contributed to differences in spatial presence between the two studies. In the present study, several participants reported ‘feeling’ the wire as they were completing the task, which may have caused restricted movement and a worry of disconnecting the power. These findings are comparable to Ferrer, Shishido, Kitahara, and Id, (2020), who also indicated participants reporting the feel of the wire. Ferrer et al. (2020) theorised that the distraction of the wire might have severed the concentration of the participant, disconnecting them momentarily from the virtual environment. This theory is somewhat supported in the present study, by one participant’s response to the wire stating that the wire reminded them they were in the real world, not the virtual environment that was presented to them. Although Le Noury et al. (2020)

also used a tethered head mounted display, the demands of the tennis task would not have required rotation of the head, and so was unlikely to have disturbed the participants.

4.4.1 Strengths and Limitations

The present study includes several strengths and weaknesses. First, the study employed the open science practice of pre-registration, including an outline of our approach to the study's data collection, sample size estimation, and method of data analysis prior to the data being collected. This approach is designed to increase the methodological rigour of the study, especially when combined with the study's outcomes and raw data which would allow others to evaluate the robustness, precision and interpretation of the study (Abt et al., 2022). However, despite stating our approach to the study's sample size, we failed to reach the minimum number of participants, which was one criterion on which our sequential analysis plan would have determined the point at which data collection finished. That said, our pre-registration plan involved a stopping criterion of "the highest density interval (HDI) for the mean difference between conditions is at most a two-pass difference either side of the point estimate". As can be observed in the results, the 95% HDI for the mean difference (-3.9) was -5.1 to -2.7, with both upper and lower bounds within the pre-registered two pass difference. As such, our goal of precision was achieved, despite not meeting our *a priori* minimum sample size. Unfortunately, COVID-19 government restrictions at the time of data collection did have an impact on participant recruitment.

For our study, a dynamic passing test was selected based on evidence of reliability in the real-world along with its ability to be replicated in VR. Our rationale for using a test with established reliability was on the basis that if variability of scores was high in the real-world then it would be plausible to find high variability of scores in VR, which would have reduced the possibility of equivalence. Therefore, this ruled out tests with un-established reliability, including the creation of new tests due to the availability of time. We consider this rigorous approach to defining our methods a strength of this study. However, it's important to consider the generalisability of our findings in relation to the validation of soccer VR. Given our study only examined one test that measured repetitive passing, our findings can only be generalised to VR being used for the purpose of measuring repetitive passing ability.

4.5 Conclusion

Our study finds a lack of evidence to support the validation of a dynamic passing test in VR. Further, our study reports several subjective differences for psychological fidelity between the two conditions, notably the ability to control the ball in VR which, was suggested to have been more difficult. Future studies should examine the relationship between these dimensions and

VR task performance. Our study also reported a high sense of presence through the ITC-SOPI questionnaire, albeit less than has been reported in a different VR sport study. However, it is inconclusive if an improved presence would have improved VR task performance in our study design, and so future research should explore this relationship.

Chapter 5 Perceptions of professional soccer coaches and players toward virtual reality and the factors that modify their intention to use it.

5.1 Introduction

Soccer ‘performance’ is a multifactorial construct comprised of tactical, technical, cognitive and physical components (Pruna & Bahdur, 2016; Russell et al., 2013; Stølen et al., 2005). To improve these components (and therefore performance), information technology systems have been invaluable to coaches and players (Dellaserra et al., 2014; Liebermann et al., 2002). For example, athlete tracking devices that include GPS and micro electrical mechanical systems have allowed a greater understanding of the physical demands imposed on soccer players during training and match-play, with the data collected used to optimise training sessions and physical performance (Dellaserra et al., 2014; Malone et al., 2017). Elsewhere, video analysis applications such as team statistical analysis and real time event analysis have allowed coaches and performance analysts to examine individual and team tactics, providing players with constructive feedback before, during or after competition (D’Orazio & Leo, 2010).

Athlete tracking devices and video analysis are now entrenched in professional soccer, with new technology introduced at an increasing rate (Almulla et al., 2020). One technology that is starting to become commercially available and used in professional sport, including soccer is VR (Thatcher, Ivanov, Szerovay, & Mills, 2020). Virtual reality is defined as a computer-simulated environment that aims to induce a sense of being mentally and/or physically present in another place (Baños et al., 2000, Neumann et al., 2018). Although VR has been available for over thirty years, professional sport teams and organisations have only recently become interested in the technology due to the increased accessibility and mobility of these systems (Cotterill, 2018). Despite this increased awareness and use of VR in professional sport, the evidence base for its efficacy is still small. That said, there is a small evidence base that does support its use as a technology to enhance athletic performance. For example, transfer of perceptual motor-skills learned within the virtual world to the real-world has been demonstrated, illustrating the effectiveness of VR as a training tool (Gray, 2017; Michalski et al., 2019; Tirp et al., 2015). Virtual reality has also been shown to be effective as a tactical analysis tool to improve decision-making when compared to conventional methods such as video recordings viewed through a computer screen (Kittel et al., 2020; Pagé et al., 2019). As such, VR presents multiple opportunities to be used within the professional sport, not least within a soccer training program. For example, physiotherapists might implement VR training with injured soccer players to train perceptual motor skills or maintain functionality of movement in a controlled environment, while minimising excessive physical loading on the skeletomuscular system (Stone et al., 2018). Alternatively, technical coaches and performance analysts may adopt VR with players as a method of analysing match performance. In doing so, players are able to re-visit the match from their own-perspective (Craig, 2014).

While there appears to be a rationale for the use of VR in professional soccer training, there is a paucity of knowledge on the current practices and perspectives of coaches and support staff who are currently using, or are likely to use this technology (Thatcher et al., 2021). To date, only one study has explored the perceptions of VR use in professional soccer clubs. The study by Thatcher et al. (2020) identified key opportunities for its implementation including rehabilitation of players and for youth development, as well as barriers towards its use, such as lack of empirical evidence and the practicality of using it. However, the perceptions represent a very small sample (n = 6), nor did the study attempt to quantify the perceptions of the individuals. For example, it is unclear how much of a barrier lack of evidence is, compared to other barriers reported in the study. Further, it is unclear what the consensus was on the technology, as the benefits may not be enough to create a desire to use it. Studies that have explored the perceptions of using VR in other industries have reported that despite the evidence-based effectiveness of the technology, conflicting opinions on the desire to use it are evident. For example, students in medical surgery have previously stated a preference for using more traditional learning modalities such as black box trainers or animal models over the more sophisticated VR systems (Kanumuri et al., 2008; Palter et al., 2010). This could be due to the lack of knowledge of such systems at the time (2006) where only 46% of skills laboratories were using VR, compared to the other modalities mentioned that had a 99% exposure rate (Palter et al., 2010). Perceptions have also been investigated in psychotherapists, where the majority (66%) indicated low ratings on the likelihood of them using VR in their applied setting (Schwartzman et al., 2012). These findings are in accordance with the diffusion of innovation theory, which suggests that the rate of adoption for a technology depends more on the end-users' subjective perceptions of usability, complexity and observability of the technology rather than objective evidence of the technology's efficacy (Donaldson et al., 2012; Rogers, 1995). It is therefore of interest to understand the related contributing factors that might facilitate or prevent the use of VR in professional soccer. For example, it was shown that the benefits of using VR in psychotherapy practice only narrowly outweighed the negative aspects, with monetary cost, technical difficulties and the need for training being important barriers (Segal et al., 2011). Contributing factors towards VR use have also been explored in orthopaedic trauma surgeons and neurosurgeons, where realism and usability of a specific VR system was viewed as positive but a lack of haptic feedback was generally viewed as negative (Koch et al., 2019).

Although it is important to understand the facilitators and situational barriers to using VR, studies that have explored these factors have not linked them to the user's intention to use the technology. This is an important limitation of previous studies because it is critical that we understand the mechanistic factors that influence intention to use VR. In contrast to a mechanistic approach, previous studies have largely taken a purely descriptive approach to examining the facilitators and barriers to using VR. In order to understand intention, studies on technology acceptance typically refer to theoretical models from the social sciences, specifically designed to examine behavioural intention to use technology and therefore explain why people adopt technologies (Liu et al., 2015). Examples of such theories include the Theory of Reasoned Action (Fishbein & Ajzen, 1975), Technology Acceptance Model (Davis, 1989), the Social Cognitive Theory (Davis, 1989) and the Unified Theory of Acceptance and Use of Technology (UTAUT; Venkatesh, Morris, Davis, & Davis, 2003). Such models are also valuable as they provide additional insights that contribute to VR acceptance that have not previously been considered. For example, previous research has not considered social influence as a contributing factor to VR acceptance, despite theoretical evidence to support it as such (Venkatesh et al., 2003).

Studies examining the contributing factors to VR use have also not considered those that are in-directly related to the specific technology. One of these factors is technology readiness, which is defined as a person's propensity to embrace and use new technologies for accomplishing goals in home life and at work (Parasuraman, 2000). Previous research has quantified technology readiness through valid and reliable scales such as the technology readiness index 2.0 (TRI 2.0), where it has been demonstrated to be a robust predictor of technology-related behavioural use intentions (Parasuraman & Colby, 2015). That is, those who score higher on the TRI 2.0, are also more likely to engage with new technology. It is therefore possible, that although there are likely to be facilitators and barriers that are directly related to VR, an individual may disengage with VR simply because they demonstrate a disinterest in technology in general.

As outlined previously (see Chapter 2), the UTAUT was developed by Venkatesh et al. (2003) in response to the wide range of differing exploratory models and theories that were used to explain an individual's perspective and behaviours when using new technologies. Through integrating eight of the most prominent theories and models such as the theory of reasoned action (Fishbein & Ajzen, 1975), technology acceptance model (Davis, 1989), and the social cognitive theory (Davis, 1989), the UTAUT identifies four core constructs that are direct determinants of behavioural intention to use a technology (Venkatesh et al., 2003). These include performance expectancy, defined as the belief that using the system will help one attain gains in job performance; effort expectancy, defined as the ease associated with using the system; social influence, defined as the belief that important others believe one should use the system; and facilitating conditions, defined as the belief that organisational and technical infrastructure exists to support the system's use (Venkatesh et al., 2003).

Whereas the UTAUT is generally used to understand the behavioural use intentions of a specific technology, the TRI 2.0 is a measure of an individual's propensity to embrace new technologies in general (Parasuraman & Colby, 2015), and is an updated and streamlined version of the original technology readiness index (Parasuraman, 2000). The TRI 2.0 is conceptualised through four dimensions: innovativeness and optimism, contributing towards 'motivators' for using technology, and discomfort and insecurity, contributing towards "inhibitors" of using technology. While the TRI 2.0 provides a measure of each of the dimensions, such as a score of innovativeness, the TRI 2.0 also provides an overall score of technology readiness.

Although it is important to understand the perspectives of those who implement new technologies in professional soccer, understanding the perceptions of players is of equal importance, as players are a key stakeholder in the adoption of a new technology (Ringuet-Riot et al., 2013). While there is currently a lack of information available on the perceptions of soccer players on the use of technology, research has demonstrated conflicting perceptions between the coaches and players in other areas. For example, academy soccer players were shown to perceive the coach-created motivational climate as performance oriented and less mastery-oriented, whereas coaches indicated an attempt on creating a mastery-orientated climate over a performance-oriented climate. (Møllerløkken et al., 2017) In this scenario, although coaches may attempt to create a mastery-orientated climate, the players perceived this differently. This difference could result in players being unhappy within the created climate, potentially causing them to leave the team, or the sport entirely. As such, it is important that coaches are aware of the views of players to avoid any conflict. Therefore, in the context of implementing a new technology, such as VR, if the users (players) perception of its overall application or how it should be used differs to those who implement it (coaches), it is possible that conflicting views will reduce the 'buy-in' to its implementation in professional soccer. To date, only two studies have looked to explore the perceptions of VR in the context of soccer training. However, one of these studies focused on VR acceptance in athletes from a variety of sports, and a variety of competitive levels, meaning the results cannot be generalised to professional soccer (Mascret et al., 2022). In the second study, Thatcher et al. (2021) did explore the perceptions of those working in professional soccer, however the results were based on a small sample (n = 6) and only reflected opinions of technical coaches and performance analysis staff, not the multidisciplinary team (i.e., soccer players, sport scientists, rehabilitation and medical staff) that professional soccer clubs regularly employ.

5.1.1 Aims

The present study has three aims. First, to examine the knowledge, perceptions, influences, and barriers of those responsible for implementing VR (coaches and support staff) and those who would use VR (players) within professional soccer clubs. Second, to identify how VR is presently used within professional soccer. Third, to examine factors that could contribute to behavioural intention to use VR in professional soccer coaches and support staff who do not currently have access to the technology.

5.2 Methods

A cross-sectional survey of coaches, support staff and players working within professional soccer was conducted between December 2019 and April 2020. The study received ethics approval (FHS203) from the Faculty of Health Sciences Research Ethics Committee at The University of Hull and conformed to the Declaration of Helsinki.

5.2.1 Sample size calculation

To support the third aim of the study, a sample size calculation was conducted. To determine the minimum sample size, the statistical model used, type 1 and 2 error rates, and the expected effect size should be considered (Abdi et al., 2013; Hair et al., 2019). As will be explained in more detail later, data were analysed using partial least squares structural equation modelling (PLS-SEM). It has been suggested that because PLS-SEM builds on ordinary least squares regression, statistical power analyses for multiple regression models will result in a satisfactory sample size estimation (Nitzl, 2016). As such, *a priori* sample size estimations were carried out using G*Power software (Faul et al., 2007), using the F tests, linear multiple regression: Fixed model, R^2 deviation from zero option (Appendix 7). Sample size was estimated to detect a moderate effect size (f^2 = 0.15), using a statistical power of 0.9 and with alpha set to 0.05. A moderate effect size was chosen for the sample size calculation as this was determined as the smallest but most meaningful effect size that we wished to find. The effect size chosen was also supported by similar studies that have used the UTAUT model and found moderate effect sizes (Liu et al., 2015). In determining the number of predictors, the construct with the largest number of formative indicators, or the dependent construct variable with the largest number of independent construct variables influencing it, is chosen (Chin & Newsted, 1999). As such, seven predictors were determined based on the number of formative indicators required for the facilitating conditions construct (Table 5.1). From these inputs and considerations, the G*Power analysis suggested that a minimum sample size of 130 participants was required.

5.2.2 Survey design

Due to the nature with which participants interact with VR, two separate surveys were designed to be completed by coaches/support staff and players, respectively. Both surveys included the following sections: (1) general information; (2) technology acceptance (3) knowledge of VR; (4) performance expectancy; (5) social influence; and (6) barriers to using VR. For the technology acceptance section, participants completed the 10 question version of the TRI 2.0 (Parasuraman & Colby, 2015). This was used to assess participant's technology readiness. Depending on whether participants had access to VR or not, a further section asked questions related to either (7) use of VR; or (8) intention to use VR.

Except for five questions which required a binary response (yes/no), all questions were multiple choice or Likert scales. For Likert scales, four to seven response label anchors were used. Fully labelled Likert scales were used instead of partially labelled scales due to their improved validity and reliability (Krosnick & Presser, 2010). The Likert scale labels reflected the different sections of the survey (e.g. agreement, influence, barriers) and each section was defined precisely (Krosnick & Presser, 2010). For example, the social influence section used a five point Likert scale ranging from 1 – 'Not at all influenced', to 5 – 'Extremely influenced'. To reduce cognitive load while completing the survey, guidelines from a theoretical model devised by Sweet (2016) were used. This included ensuring all parts of the survey were visible to the participant, i.e., did not include hidden text that the participant had to open, such as drop-down boxes; ensured all text was clear, with a high contrast font and background; and only used text that was familiar and clear to the respondents. As such, jargon and acronyms were avoided throughout.

To assess the content validity of the survey and its ease of completion, a pilot study was conducted using five support staff (three sport scientists and two performance analysts) and two players employed by a UK second division soccer club, and four individuals (three lecturers, one student) working at a UK higher education institution. Pilot study feedback was collected in person or via email. The feedback suggested that both surveys were time consuming and that the time to complete the survey may result in participant incompletion. As such, 8 questions were removed from the coaches/support staff survey, and 9 questions were removed from the players survey. The final surveys therefore consisted of 54-63 questions for coaches/support staff, and 33-39 questions for players. The number of questions completed depended on the participants knowledge of VR and if they currently used it within their clubs. Although some questions were the same in both surveys to allow comparisons, other questions were specific to each group. Both surveys followed the same order and themes consisting of (1) general information (4-8 questions); (2) technology acceptance (10 questions); (3) knowledge of VR (3-6 questions); (4) performance expectancy (5-16 questions); (5) social influence (3-5 questions);

(6) barriers to using VR (4-9 questions); and (7) use of VR (4-8 questions); or (8) intention to use VR (3 questions).

Questions assessing the inclusion criteria were placed at the start of the survey after reading the information sheet and providing participation consent. The inclusion criteria required the participant to be 17 years of age or older and have not submitted the survey previously. Additionally, at the time of completion, participants had to be working for a professional soccer club, national association team (coach/support staff survey), or be in a professional contract at a professional soccer club (players survey).

After completing the general information and technology acceptance sections, VR was defined as, "including a headset worn by a user that covers their eyes, allowing them to experience a virtual world that is created by a computer". To further clarify our point on the definition, VR was also defined as "not referring to non-immersive technology that results in an output through a television or other electronic interface, and does not include augmented reality, whereby computer-generated images are placed into the real world and viewed live". Images were included alongside the definitions so that participants were aware of the type of VR being referred to. To define the context in which questions should be answered, the definition "Virtual reality used by coaches, support staff and players as part of training or personal use within professional football training grounds" was used. This was to differentiate from answering in other contexts where VR is used in soccer such as fan engagement (Kim & Ko, 2019).

5.2.3 Survey distribution

The survey was distributed electronically via email to soccer clubs in England, Scotland, Australia, and the United States of America. The survey was first sent to known contacts of the research team at soccer clubs in each of these countries. For clubs where no relationship existed with the research team, an invitation email was sent to the Head of Medicine and Sport Science, or the equivalent position. Within the invitation email, information was provided on the aims and benefits of the research, and details regarding confidentiality of data. Participants were asked to circulate the surveys in their club, specifically to coaches (technical coaching staff), support staff (all staff that work in medical or performance, including performance analysts) and players. Additionally, the survey was circulated openly on the social media platform, Twitter. As such, the survey was made open to individuals working in clubs outside of the countries outlined above. Data was collected using an online survey platform (www.onlinesurveys.ac.uk, Bristol, UK), with links to both surveys included in the email or on social media. A reminder email was sent out one month before the closing date. The coach/support staff and player surveys took approximately 10 and 6 minutes to complete, respectively. As participants were encouraged to share the survey with those in their team, and the surveys were circulated openly through social media, it was not possible to determine the response rate to the surveys.

5.2.4 The partial least squares structural equation model specification

Specification of the PLS-SEM model involved the development of the inner (also known as structural) and outer (also known as measurement) models. In the first step, the inner model was created, involving the specification of the path models between the independent (exogenous) construct variables and the dependent (endogenous) construct variable. Secondly, the outer model was specified by connecting the indicator variables that correspond with the constructs specified in the inner model. The indicator variables that corresponded to the constructs were based on theoretical knowledge in the literature, and the experience of the researcher who was based at a professional soccer club. The independent and dependent construct variables are outlined below with their corresponding indicator variables.

5.2.4.1 Dependent construct variable

Likeliness to use VR was specified as the dependent construct variable and consisted of two indicator variables (Table 5.1). Likeliness to use construct was specified as a reflective construct, that is, the indicator variables chosen for the construct reflect likeliness to use VR.

5.2.4.2 Independent construct variables

The UTAUT constructs were adapted to be appropriate for this study. As such, the performance expectancy, social influence and facilitating conditions constructs were used, under new definitions. Performance expectancy was defined as “the degree of belief that virtual reality will improve performance”; social influence as “the degree of being socially influenced to use virtual reality”; and facilitating conditions as “the degree that barriers are in place to use virtual reality”. Effort expectancy was not included in this study under the assumption that most participants would not have access to VR, nor would most be directly responsible for its operation. Performance expectancy and social influence consisted of five indicator variables, whereas facilitating conditions consisted of seven indicator variables (Table 5.1). All the adapted UTAUT constructs were specified as formative constructs, that is, the indicator variables chosen form each of the constructs. Technology readiness was also included as a single item construct using the overall score of the TRI 2.0 questionnaire. The overall score for technology readiness was calculated through methods outlined by Parasuraman & Colby (2015).

Table 5.1 Specification of the PLS-SEM constructs and indicator variables.
Also included are the questions used to make each construct.

Construct	Indicator variable	Question
Performance expectancy	Physical	Physical fitness (i.e., Virtual reality used with players to improve areas such as strength, power, aerobic fitness etc.)
	Cognition	Cognition (i.e., Virtual reality used with players to improve cognition such as decision making, reaction time, visual awareness etc.)
	Technical	Technical skill (i.e., Virtual reality used with players to improve technical ability such as passing & shooting accuracy etc.)
	Tactical	Tactical development (i.e., Virtual reality used with players to improve awareness of team tactics etc.)
	Mental Wellbeing	Mental wellbeing (i.e., Virtual reality used with players to improve mental wellbeing such as stress and anxiety etc.)
Social influence	To be seen using	To be seen using an innovative technology (i.e., I would be influenced to use virtual reality so that others see me using an innovative technology)
	Influential others	Influential others use virtual reality (i.e., I would be influenced to use virtual reality if individuals that influence me also use it)
	Influential clubs	Influential clubs use virtual reality (i.e., I would be influenced to use virtual reality if clubs that influence me also use it)
	Seniors want it used	Those senior to me (i.e., I would be influenced to use virtual reality if individuals that are senior to me want it to be used)
	Players enjoy using	Player enjoyment (i.e., I would be influenced to use virtual reality if players enjoyed using the system)
Facilitating conditions	Player buy-in	Player buy-in (i.e., Getting players to engage with the virtual reality system is a barrier to using it)
	Coach buy-in	Coach and support staff buy-in (i.e., Coaching staff buy-in to virtual reality being used with players is a barrier to using it)
	Space to operate	Personnel to operate (i.e., requiring personnel to operate the virtual reality system is a barrier to using it)
	Personnel to operate	Space to operate (i.e., space within the training ground to operate the virtual reality system is a barrier to using it)
	Limited evidence	Limited evidence base (i.e., limited research available on virtual reality used in professional football is a barrier to using it)
	Time available	Time available (i.e., time available to use within schedule is a barrier to using virtual reality)
	First impression	First impression (i.e., my first impression of using, seeing or hearing about virtual reality is a barrier to using it)
Technology readiness	TRI 2.0 overall score	Overall score of the 10 technology readiness questions
Likeliness to use	Intention	If virtual reality technology was made available to you, how likely are you to use it within your club?
	Opinion	What is your overall opinion of virtual reality technology for use by coaches, support staff and players within the training ground setting?

5.2.5 Data analysis

Data were analysed in two stages. In the first stage, descriptive statistics of the participant's knowledge, perceptions, influences, barriers and opinions of VR were produced. For categorical, multiple choice and Likert scale questions, frequency analysis was conducted with percentages and number of participants reported. Likert scales were treated as ordinal variables, due to the non-normally distributed data. Data were confirmed as being not normally distributed using the Shapiro-Wilk test and quartile-to-quartile plots. For testing the multivariate research model, a PLS-SEM technique was used. PLS-SEM is a second-generation data analysis technique, which unlike first generation techniques, such as linear regression, allows for the modelling of relationships in multiple independent and dependent construct variables to take place simultaneously. As such, research questions can be answered in a single, systematic and comprehensive analysis (Gefen et al., 2000). Unlike linear regression, PLS-SEM allows for the structural model and the measurement model to be analysed in a single process. In other words, confirmatory factor analysis is combined with hypothesis testing in a single operation (Gefen et al., 2000). Unlike the alternative method of conducting structural equation modelling such as covariance based structural equation modelling (CB-SEM) which aims to confirm theories by determining how well a model can estimate the covariance matrix for the sample data, PLS-SEM operates similar to a multiple regression by maximising the variance of the dependent variable construct (Hair et al., 2011). Over the past decade there has been a considerable debate as to which SEM approach is most appropriate (Reinartz et al., 2009; Rigdon et al., 2017). However, it is suggested that PLS-SEM should be used when the analysis is testing a theoretical framework from a predictive perspective, when the path model includes one or more formatively measured constructs and when the research objective is to better understand increasing complexity by exploring theoretical extensions of established models (Hair et al., 2019). Other advantages of PLS-SEM are the ability to use non-normally distributed data due to the method being non-parametric. While the maximum likelihood estimation method with CB-SEM has been noted to be robust against violations of normality, larger sample sizes are often required as if the sample is limited, CB-SEM can produce abnormal results when non-normally distributed (Reinartz et al., 2009).

The PLS-SEM was assessed through the evaluation of the inner and outer models. As formatively and reflectively measured constructs are based on different concepts, different evaluation measures took place (Hair et al., 2019). In the assessment of the reflective construct, evaluation of the internal consistency reliability, convergent validity and discriminant validity took place. To assess the internal consistency reliability, Cronbach's alpha was calculated, providing an estimate of the reliability based on intercorrelations of the constructs observed indicator variables (Hair et al., 2013). In addition, composite reliability was assessed due to Cronbach's alpha being a conservative measure of reliability (Peterson & Kim, 2013). In contrast, composite reliability overestimated reliability. Both reliability measures are interpreted in the same way, with scores above 0.7 being considered as satisfactory. Next, convergent validity was assessed which is the extent that a measure correlates positively with alternative measures of the same construct. As reflective indicators of the same construct are treated as different approaches to measure the same construct, it is expected that indicator variables of the same construct should converge with one another and share a high proportion of variance. To assess this, the outer loadings of the indicators were considered, and the mean variance extracted was calculated. For reflective constructs, outer loadings should be statistically significant and above 0.7. The mean variance extracted value is calculated as the grand mean value of the indicator loadings associated with the construct (Hair et al., 2013). A value of 0.5 and above is considered satisfactory and indicates that a construct explains more than half the variance of its indicators. Finally, discriminant validity was assessed and is defined as the extent that the construct is truly distinct from other constructs, implying that it is unique within the model (Hair et al., 2013). First, the indicator's outer loadings should be higher than all its cross-loadings of other constructs. Secondly, the Fornell-Larcker criterion should demonstrate that the square root of the mean variance extracted in the intention to use construct is higher than its highest correlation with the other constructs within the model. In other words, the intension to use construct should share more variance with its associated indicators than with any other construct.

In assessing the formative constructs, evaluation of collinearity, significance, and relevance of the formative indicators was conducted. First, collinearity was assessed. Collinearity refers to high correlations between two or more indicator variables and its occurrence can be problematic during the analysis and interpretation of the PLS-SEM. This is due to high levels of collinearity having an impact on the estimation of weights and their significance (Hair et al., 2013). In assessing collinearity issues, variance inflation factor was used. In terms of PLS-SEM, a variance inflation factor higher than 5 is an indication of potential collinearity. However, the ideal threshold is a variance inflation factor below 3.3. In the second stage, the statistical significance and relevance of each indicator was assessed. As PLS-SEM is a non-parametric method, a 5000 bootstrapping procedure was carried out to determine statistical significance. The confidence intervals (CI) of the indicator weights were used to determine if the indicators were significant, with a CI including a 0 interpreted as non-significant. Where an indicator's weight was not significant, the indicator's absolute contribution to the construct was considered via assessing its outer loading. That is, the correlation between the indicator variable and the construct when no other indicators are taken into consideration. Indicators with weights and outer loadings that were both not significant, were removed from the construct as the indicator provided no significant explanation in forming the construct. For indicator variables with a weight that was not significant, but the outer loading was significant and above .5, remained in the construct. For indicators with significant outer loadings but below .5, strong theoretical support should be in place to support its inclusion in the formative construct model.

For the inner model, collinearity between constructs was first assessed using the variance inflation factor criteria as outlined previously. The inner model was then assessed by means of (1) the size and statistical relevance of the path coefficients (β); (2) the explained variance (R^2); and (3) the path coefficient effect size (f^2) (Hair et al., 2013). The R^2 value represented the independent construct variables combined effect on the dependent construct variable. The R^2 values of 0.25 'weak', 0.5 'moderate' and 0.75 'substantial' were used. Effect sizes (f^2) were reported as 0.02 'small', 0.15 'medium' and 0.35 'large', and represented the independent construct variables contribution to the dependent construct variable R^2 value (Cohen, 1992). Descriptive statistics were analysed using R Studio (RStudio Team, 2020), with figures produced using the 'Likert' package (Bryer & Speerschneider, 2016) and 'ggplot2' package (Wickham, 2016). The multivariate model analysis was analysed using SmartPLS 3.0 (Ringle et al., 2015). The alpha level of significance was set at $p < 0.05$.

5.3 Results

5.3.1 Participant demographics

Overall, 207 participants took part in the survey coaches/support staff: n = 143, male = 135, female = 8; players: n = 64, males = 63, female = 1) (Table 5.2). In the coaches and support staff group, participants were predominantly from England (n = 89), with other respondents being from the United States (n = 12); Australia (n = 10), Scotland (n = 9) and other (n = 23). In contrast, most players were from England (94%, n = 60). Most coaches/support staff worked with male football players (94%) in roles that included sport scientist (n = 46); physiotherapists (n = 26); performance analyst (n = 25); strength and conditioning coach (n = 19); technical coaches (n = 13); head coach/manager (n = 6) and other roles (n = 8) such as heads of performance, rehabilitation coaches and head of innovation and research. Coaches/support staff mostly worked with senior players (61%, n = 87) and in clubs where the first team competed in the top tier (45%, n = 64), or second tier (38%, n = 54). In contrast, most respondents for the players' survey were from first teams that competed in the second tier (72%, n = 46). The most frequently reported age for coaches/support staff was the 27-31 years category (36%, n = 51), whereas for player survey most respondents were in the 17-21 years category (53%, n = 34).

Table 5.2. Proportion and frequency of participant demographics for coaches/support staff and players. Participant demographics are included for the descriptive statistics and the PLS-SEM.

		Descriptive statistics		**PLS-SEM** [a]
Demographic	**Characteristics**	**Coaches/support staff % (n)**	**Player % (n)**	**Coaches/support staff – no access % (n)**
Gender	Male	94% (135)	98% (63)	95% (117)
	Female	6% (8)	2% (1)	5% (6)
Age	17-21	1% (2)	53% (34)	2% (2)
	22-26	21% (30)	23% (15)	19% (23)
	27-31	36% (51)	17% (11)	38% (47)
	32-36	15% (22)	5% (3)	16% (20)
	37-41	14% (20)	2% (1)	15% (19)
	42-46	4% (6)		4% (5)
	47-51	6% (8)		3% (4)
	52+	3% (4)		2% (3)
Tier	Tier 1	45% (64)	6% (4)	46% (57)
	Tier 2	38% (54)	72% (46)	34% (42)
	Tier 3	6% (9)	9% (6)	7% (8)
	Tier 4	6% (8)	13 (8)	7% (8)
	Tier 5	1% (2)		2% (2)
	National association team	4% (6)		5% (6)
Team	Senior players	61% (87)		57% (70)
	Senior academy players	27% (39)		30% (37)
	Academy players	11% (16)		12% (15)
	Junior academy players	1% (1)		1% (1)
Country	England	62% (89)	94% (60)	59% (73)
	United states of America	8% (12)		8% (10)
	Scotland	6% (9)		7% (9)
	Australia	7% (10)	1% (1)	8% (10)
	Other	17% (23)	5% (3)	18% (21)
Working gender	Male	94% (135)		93% (115)
	Female	6% (8)		7% (8)
Highest qualification	PhD	9% (13)		10% (12)
	Masters	55% (78)		57% (70)
	Bachelors	28% (40)		25% (31)
	Other	8% (12)		8% (10)

[a] PLS-SEM – Partial least squares structural equation model

5.3.2 Descriptive statistics

5.3.2.1 Awareness and experience of VR

Most coaches/support staff (94%) and players (89%) knew what VR was, based on the definitions and pictures provided within the survey. Additionally, most coaches/support staff (76%) and players (72%) were aware of VR being used within professional soccer training grounds. Most of the coaches/support staff (70%) and players (54%) had never used VR within a professional football ground, however, those that have used VR within a professional soccer training ground did so within the last year (coaches/support staff: 22%; players: 44%) (Table 5.3).

Table 5.3. Proportion and frequency of coach/support staff and player responses to awareness and experience of using VR.

Question	Answer	Coaches/support staff % (n)	Player % (n)
Have you ever heard of, or used a virtual reality system?	Yes	94% (134)	89% (57)
	No	6% (9)	11% (7)
Are you aware of VR in professional soccer	Yes	76% (108)	72% (46)
	No	24% (35)	28% (18)
Have you used virtual reality within a professional soccer club? If you have used virtual reality as part of a demonstration or trial, please answer 'yes'.	Yes	30% (43)	47% (30)
	No	70% (100)	53% (34)
When was your last experience of using virtual reality within a professional soccer club?	Never used	70% (100)	53% (34)
	Within the last year	22% (32)	44% (28)
	Between 1 - 2 years ago	6% (8)	3% (2)
	Between 3 - 5 years ago	1% (2)	
	Between 6 - 10 years ago	1% (1)	

5.3.2.2 Performance expectancy

Coaches and players were similar on how they perceived VR could improve performance (Figure 5.1). Responses by coaches and players indicated in favour of agreement that VR could improve cognition (coach: 93%; player: 80%) and tactical performance (coach: 97%; player: 75%). Similarly, approximately half of coaches (52%) and players (53%) indicated in favour of disagreement that VR could improve physical performance. For VR’s ability to improve mental wellbeing, although over half of coaches generally agree to some extent (55%), a third of respondents remained neutral (33%). As for players, there was a greater variability in responses, with most remaining neutral (39%), followed by agreement (36%) and disagreement (25%). Larger differences were observed between groups for belief that VR could improve technical ability, with players indicating greater disagreement than coaches (58% vs 34%). Additionally, whereas the largest frequency of responses for players was 2 (Disagree), the largest frequency of responses for coaches was 5 (Somewhat agree).

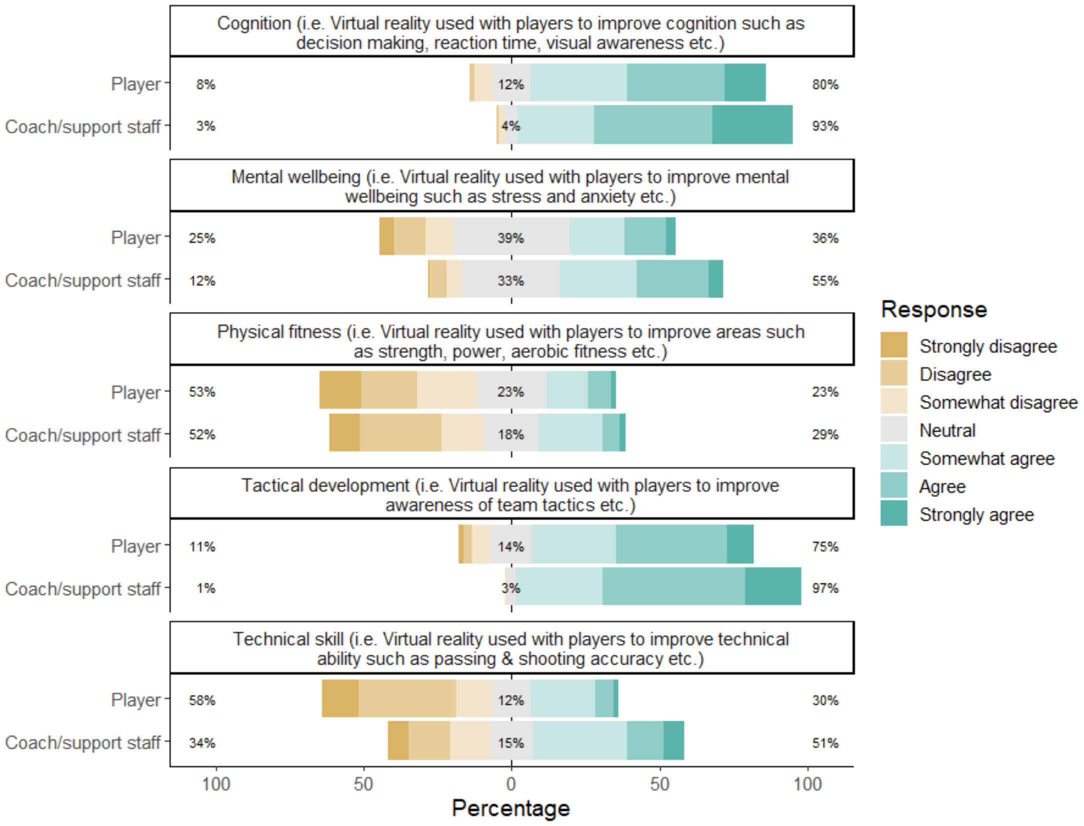


Figure 5.1 Likert bar-plot of responses by coaches/support staff and players to statements regarding what VR should be used for in order to improve. Percentages indicate overall disagreement, neutral and overall agreement, from left to right respectively.

Coaches generally agreed that VR should be used for performance analysis (93%), followed by preparation (77%) and rehabilitation (73%) (Figure 5.2). While over half of coaches generally agreed that VR should be used for testing (54%) or training (53%), a third of respondents disagreed, 29% and 31%, respectively. There was no consensus on whether VR should be used for player monitoring or talent identification, with an equal proportion of coaches agreeing (35%), disagreeing (34%) or remaining neutral (31%) in the latter.

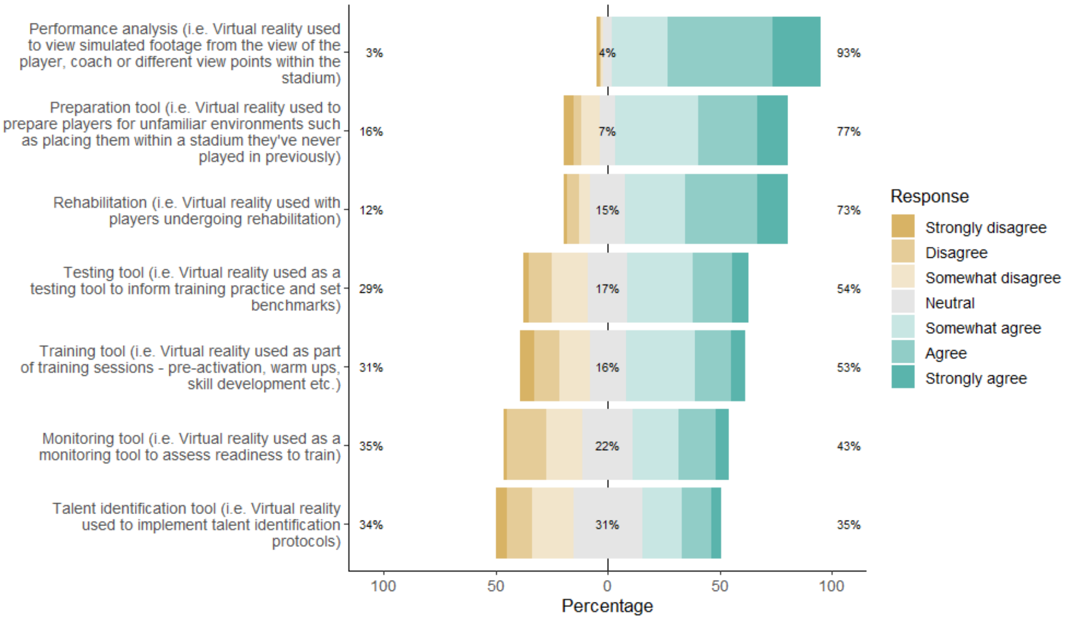


Figure 5.2. Likert bar-plot of responses by coaches/support staff to statements regarding when VR should be used. Percentages indicate overall disagreement, neutral and overall agreement, from left to right respectively.

Coaches/support staff agreed to some extent that VR should be used with senior academy players (83%), academy players (71%) or senior players (70%), with all demonstrating relatively little disagreement (6-15%). Coaches/support staff indicated greater variability for VR being used with junior academy players, with just over a half of respondents agreeing (53%), and a third of respondents disagreeing (34%) (Figure 5.3).

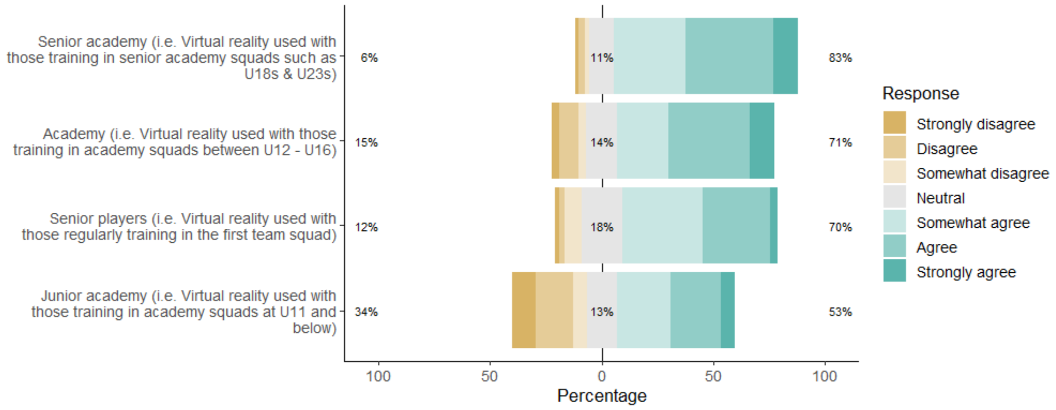


Figure 5.3. Likert bar-plot of responses by coaches/support staff to statements regarding who VR should be used with. Percentages indicate overall disagreement, neutral and overall agreement, from left to right respectively.

5.3.2.3 Social influence

Coaches and players responded similarly on influences to use VR (Figure 5.4). Coaches and players are influenced to use VR in some capacity if influential clubs use VR (coach = 78%, player = 98%) and if influential others use VR (coach = 83%, player = 78%). Being seen to use VR as an influencer also demonstrated similarities between groups, with both groups showing an equal split between being not influenced (47%) and being influenced in some capacity (53%).

Additionally, coaches responded as being very influenced to use VR if players enjoyed using it (98% overall influence) and somewhat influenced if seniority wanted it to be used (91% overall influence).

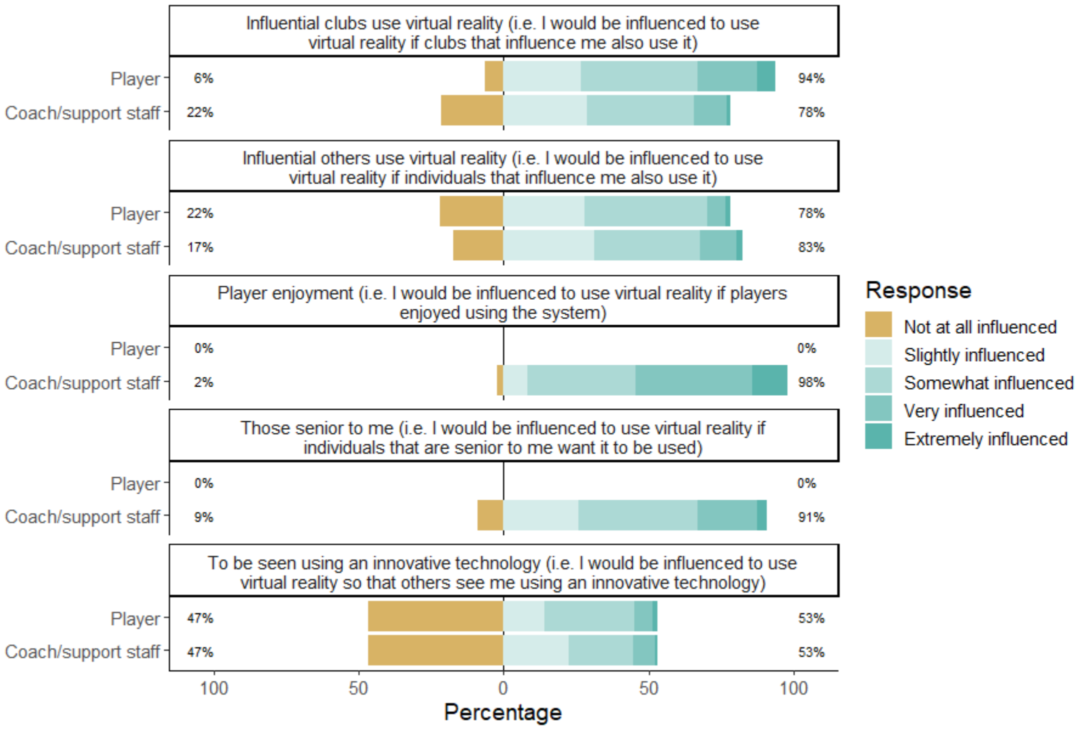


Figure 5.4. Likert bar-plot of responses by coaches/support staff and players to statements on how they perceive they are influenced to use VR. Percentages indicate not at all influenced, and overall influenced, from left to right respectively.

5.3.2.4 Facilitating conditions

Of the barriers that coaches and support staff responded to, monetary cost was seen as the largest barrier albeit in a reduced sample of 36 participants who were aware of the associated cost of VR. As such, 107 participants were not aware of the costs associated with VR. For this reason, this variable was not included in PLS-SEM. Limited research within football and time available to use VR were also seen as moderate barriers to using the technology. First impression of VR was seen as the lowest barrier, with 45% of respondents indicating it as not being a barrier (Figure 5.5).

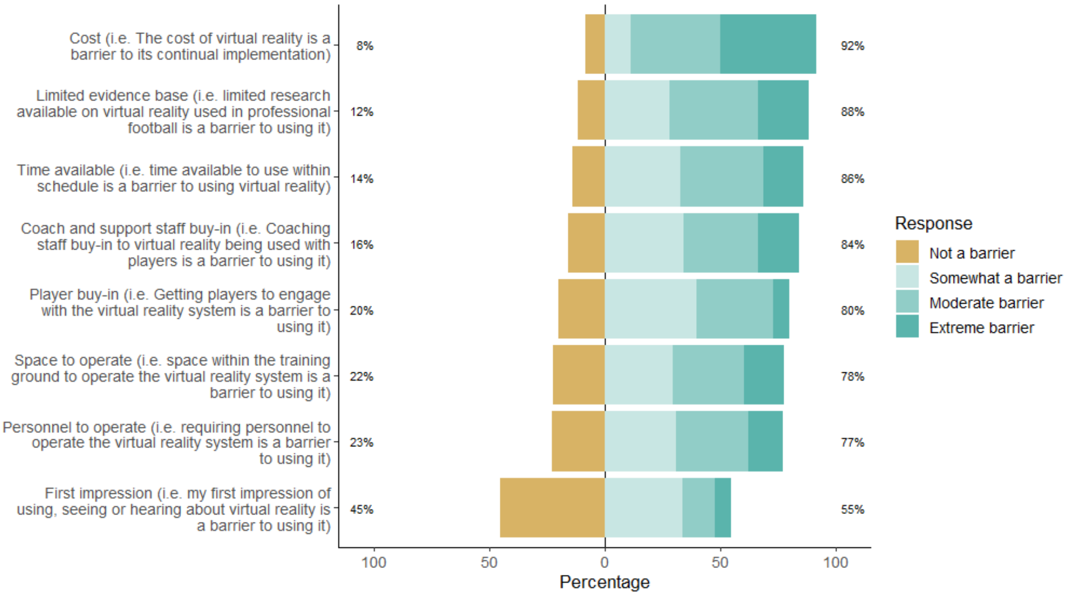


Figure 5.5. Likert bar-plot of responses by coaches/support staff to statements regarding facilitating conditions to using VR. Percentages indicate not a barrier and overall barrier, from left to right respectively. For the statement ‘cost’, only 36 participants responded as being aware of the monetary cost associated with VR.

5.3.2.5 Current access to VR

Of the 143 coach/support staff responses, only 20 had access to VR technology within their respective clubs, whereas of the 64 player responses, 32 had access to VR within their clubs. Out of the 20 coaches/support staff that had current access, 3 responded as being responsible for operating or instructing players through VR related activities. In regard to how VR was used within their respective clubs, 3 identified VR use as a training tool, 2 identified VR use as a testing tool and for video analysis, and 1 identified VR use for entertainment purposes. Regarding what coaches aimed to change through VR use, 3 specified cognition, 2 specified tactical awareness and 1 specified for changing technical skill and mental well-being. The 3 participants used VR for approximately 11-30 minutes with each player, ranging between a few times per week, to a few times per month. All 20 coaches responded with their clubs having access to VR within the past 2 years, or less.

5.3.2.6 Opinion of virtual reality

Coaches/support staff and players responded similarly on their overall opinions of VR being used within soccer training ground facilities (Figure 5.6). In both groups, half of respondents generally viewed VR as being positive, whereas 13% and 11% of coaches/support staff and players, respectively, generally viewed VR as being negative. Both groups largely remained neutral in their views of VR.

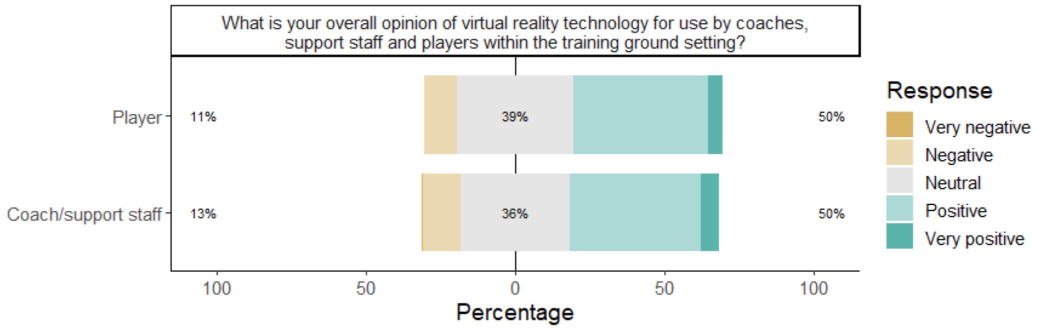


Figure 5.6. Likert bar-plot of responses by coaches/support staff and players to a statement regarding the overall opinion of VR being used within the training ground setting. Percentages indicate overall negativity, neutral and overall positivity, from left to right respectively.

5.3.3 Multivariate analysis: factors to determine acceptance of virtual reality if given access

5.3.3.1 Demographic information

Overall, 123 coaches/support staff were included in the multivariate analysis based on them not having access to VR (Table 5.2). A sensitivity analysis was conducted to determine the size of effect that the PLS-SEM could detect. Sensitivity analysis was carried out using G*Power software (Faul et al., 2007) using the F tests, linear multiple regression: Fixed model, R^2 deviation from zero option. Input parameters were set as a statistical power of .9; predictors set to 7; and alpha at .05. Based on a sample of 123, the study had enough power to detect a moderate effect size ($f^2 = 0.16$).

5.3.3.2 Assessment of the outer model

The results of the reflective constructs internal consistency reliability, convergent validity and discriminant validity are shown in Table 5.4. For the internal consistency reliability, Cronbach alpha and composite reliability were above 0.7, indicating satisfactory reliability (Peterson & Kim, 2013). For the convergent validity, outer loadings were above 0.7 (0.914, 0.904), and the average variance extracted was above 0.5 (0.826) demonstrating satisfactory convergent validity. Finally, the Fornell-Larcker criterion showed that the square root of the average variance extracted was greater than the related inter-construct correlations of the other constructs in the model, illustrating that adequate discriminant validity existed. The formative constructs were assessed for validity using the variance inflation factor, and the outer weights and loadings significance testing. In assessing collinearity, the variance inflation factor illustrates that all indicators were below the optimal threshold of 3.3, indicating no collinearity issues existed and therefore multicollinearity was not an issue for estimating the PLS path model (Table 5.5).

Table 5.4 Internal consistency reliability, convergent validity and discriminant validity of the reflective construct variable, likeliness to use.

Constructs	**Outer loadings**	**Cronbach Alpha**	**Composite reliability**	**Average variance extracted**	**Fornell-Larcker criterion**
					Likeliness to use
Likeliness to use	(0.914, 0.904)	0.79	0.905	0.826	0.909
Performance expectancy					0.620
Social influence					0.376
Facilitating conditions					-0.502

Table 5.5. Collinearity assessment of the formative construct indicators using variance inflation factor.

Performance expectancy		Social influence		Facilitating conditions	
Indicator	VIF	Indicator	VIF	Indicator	VIF
Physical	1.283	To be seen using	1.646	Player buy-in	1.298
Cognition	1.647	Influential others	3.043	Coach buy-in	1.431
Technical	1.607	Influential clubs	2.928	Space to operate	1.177
Tactical	1.542	Seniors want it used	1.777	Personnel to operate	1.241
Mental Wellbeing	1.306	Players enjoy using	1.374	Limited evidence	1.339
				Time available	1.489
				First impression	1.453

VIF – variance inflation factor

Each construct's indicator weight and loadings are outlined in Table 5.6. For the performance expectancy construct, the indicator weights were statistically significant for cognition (0.66, $p < 0.001$) and physical (0.326, $p = 0.018$), whereas they were not statistically significant for technical (0.33, $p = 0.062$), tactical (-0.083, $p = 0.565$), and mental wellbeing (0.062, $p = 0.681$). For the three indicators that were not statistically significant, outer loadings illustrated highly statistically significant contributions to the overall construct and were therefore not removed from the model. For the social influence construct, indicator weights were statistically significant for influential clubs (0.736, $p = 0.026$) and player enjoyment (0.671, $p = 0.002$), whereas the remaining three indicators for to be seen using (0.008, $p = 0.525$), influential others (-0.044, $p = 0.915$) and seniors want it used (-0.156, $p = 0.633$) were not statistically significant. However, the outer loading illustrated a highly statistically significant contribution to the construct and therefore remained. Finally, facilitating conditions constructs indicator weights were statistically significant for coach buy in (0.496, *p* = 0.011) and limited evidence base (0.56, *p* = 0.001). The remaining five indicators of player buy-in (0.022, *p* = 0.915), space to operate (-0.15, *p* = 0.373), personnel to operate (0.047, *p* = 0.781), time available (-0.031, *p* = 0.849) and first impression (0.348, *p* = 0.062) were not statistically significant. On assessment of the outer loadings, all displayed statistically significant contributions to the overall construct except for space to operate (0.115, *p* = 0.503), illustrating no absolute contribution to the overall construct. As such, the space to operate indicator was removed from the construct. This was further supported due to the low theoretical contribution to support the indicators inclusion in the construct.

Table 5.6. Formative construct indicators item weight and outer loading. Also included are the item weight and outer loading t statistic and p value, and the item weight 95% bias-corrected and accelerated confidence interval.

Formative construct	**Indicator**	**Item weight (outer loading)**	**Item weight t. stat (*p* value)**	**95% Bca Confidence intervals**	**Outer loading t.stat (*p* value)**
Performance expectancy	Physical	0.326 (0.635)	2.372 (0.018)	0.044,0.579	5.965 (0.000)
	Cognition	0.66 (0.847)	5.016 (0.000)	0.404, 0.917	10.452 (0.000)
	Technical	0.33 (0.733)	1.867 (0.062)	0.012, 0.692	3.688 (0.000)
	Tactical	-0.083 (0.485)	0.575 (0.565)	-0.37, 0.184	7.306 (0.000)
	Mental Wellbeing	0.062 (0.52)	0.411 (0.681)	-0.221,0.361	4.532 (0.000)
Social influence	Seen using	0.008 (0.525)	0.026 (0.979)	-0.539, 0.635	2.754 (0.006)
	Influential others	-0.044 (0.681)	0.106 (0.915)	-0.89,0.713	3.935 (0.000)
	Influential clubs	0.736 (0.802)	2.224 (0.026)	0.131, 1.392	5.55 (0.000)
	Seniors want it used	-0.156 (0.563)	0.478 (0.633)	-0.795,0.461	2.586 (0.01)
	Players enjoy using	0.671 (0.78)	3.068 (0.002)	0.247,1.035	4.429 (0.000)
Facilitating conditions	Player buy in	0.022 (0.376)	0.106 (0.915)	-0.365,0.452	2.217 (0.027)
	Coach buy in	0.496 (0.722)	2.559 (0.011)	0.13, 0.86	6.065 (0.000)
	Space to operate	-0.15 (0.115)	0.89 (0.373)	-0.478,0.178	0.67 (0.503)
	Personnel to operate	0.047 (0.381)	0.278 (0.781)	-0.287, 0.372	2.318 (0.021)
	Limited evidence	0.56 (0.728)	3.387 (0.001)	0.256, 0.897	6.586 (0.000)
	Time available	-0.031 (0.451)	0.19 (0.849)	-0.344,0.308	3.227 (0.001)
	First impression	0.348 (0.686)	1.865 (0.062)	-0.034, 0.691	5.093 (0.000)

Bca – Bias-corrected and accelerated confidence interval.

5.3.3.3 Assessment of the inner model

Before the assessment of the structural inner model, multicollinearity of the independent construct variables and the single item construct (technology acceptance) was assessed through the VIF. Results showed that all VIF values ranged between 1.005 – 1.326, which was below the optimal VIF value of 3.3, indicating no issues of collinearity between the independent construct variables and the dependent construct variable (Table 5.7).

Table 5.7. Collinearity assessment of the formative constructs using variance inflation factor

Independent construct variables	Likeliness to use VIF value
Performance expectancy	1.326
Social influence	1.252
Facilitating conditions	1.076
Technology acceptance	1.005

Assessment of the inner structural model is shown in Table 5.8. The path from performance expectancy to Likeliness to use was positive and statistically significant (β Performance expectancy – Likeliness to use = 0.465, t = 7.028, p = <0.001, f^2 = 0.34), whereas facilitating conditions was negative and statistically significant (β Facilitating conditions – Likeliness to use = -0.364, t = 5.164, p = <0.001, f^2 = 0.26). The path coefficients of social influence (β Social influence – Likeliness to use = .131, t = 1.924, p = 0.054, f^2 = 0.029) and technology acceptance (β Technology acceptance – Likeliness to use = 0.039, t = 0.617, p = 0.537, f^2 = 0.003) was insignificant. The explained variance (R^2) in the dependent construct by the independent constructs was 0.523, indicating moderate in-sample explanatory power of the structural model (Figure 5.7).

Table 5.8. Path coefficients between the independent and dependent construct variables. Also included are the t-value, p-value, effect size, and the dependent construct variables explained variance.

Path	Path coefficient β (95% CI)	t-value	p-value	Effect size f^2 (95% CI)	Explained variance R^2
Performance expectancy - Likeliness to use	0.465 (0.336, 0.592)	7.028	0.000	0.343 (0.173,0.674)	0.523
Social influence - Likeliness to use	0.131 (-0.042, 0.231)	1.924	0.054	0.029 (0.002,0.175)	
Facilitating conditions - Likeliness to use	-0.364 (-0.489, -0.209)	5.164	0.000	0.259 (0.099,0.591)	
Technology acceptance - Likeliness to use	0.039 (-0.082, 0.162)	0.617	0.537	0.003 (0, 0.059)	

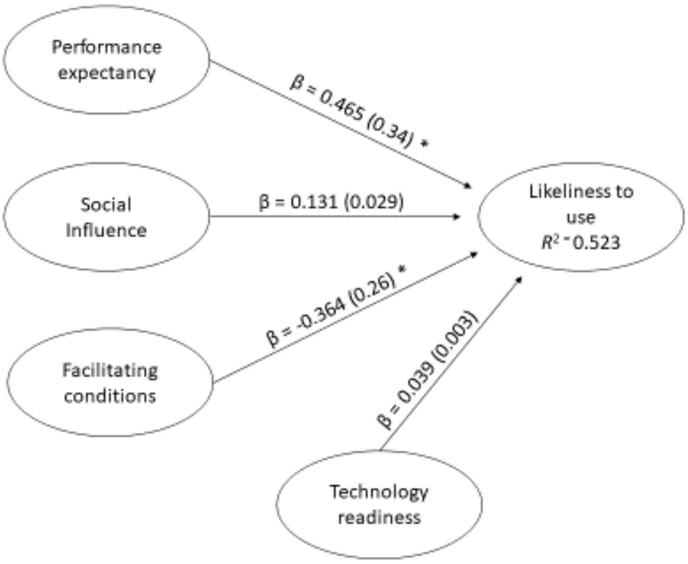


Figure 5.7. Node diagram showing the path coefficients between the independent construct variables and the dependent construct variables.
β = beta coefficients; (f^2) = path coefficient effect size; R^2 = explained variance
* statistically significant at < 0.001.

5.4 Discussion

The findings of this study add to a limited area of research on the current perceptions of VR use within professional soccer. Using a modified Unified Theory of Acceptance and Use of Technology model, the present study explored the relationships between coaches' perceptions of VR and the likelihood of using it if given the access. Our model explained a moderate amount of variance in the dependent variable (r^2 = 0.523), with performance expectancy and facilitating conditions explaining the most variance, and social influence and technology readiness explained the least variance. Finally, for each construct we explore some of the key perceptions of VR such as its use for improving cognition and tactical awareness, as well as potential barriers such as limited evidence base and time availability.

Our study showed that performance expectancy positively contributed towards likeliness to use and was the greatest overall contributor in our model (Table 5.8). Performance expectancy, in most cases, has been reported as the determinant of intention to use a technology (Venkatesh et al., 2003). This comes as no surprise given that modern technology has had a significant impact on professional sport, with many coaches considering advances in technology to be invaluable (Liebermann et al., 2002). In our model, the largest absolute contributors to performance expectancy were cognition, followed by tactical and physical (Table 5.6). Distribution of responses for performance expectancy are shown above (Figure 5.1). Overall, there was agreement that cognition could be improved through VR. Several studies have already alluded to the potential of VRs ability to improve elements of cognition (Fadde & Zaichkowsky, 2018). Further, research in baseball, tennis and darts provide evidence of perceptual cognitive improvements through use of VR (Gray, 2017; Michalski et al., 2019; Tirp et al., 2015). Similarly, this study found overall agreement for VRs ability to improve tactical awareness. Commercially available VR software from companies such as Rezzil and Beyond Sport, allow players to review performance footage from their on-field perspective (as opposed to a view from the stadium), providing an athlete-centred approach (Thatcher et al., 2021). Although studies have not examined the effect of reviewing match footage with VR in players, previous research has illustrated the effectiveness of Australian rules football umpires using VR for reviewing match footage (Kittel et al., 2020). Kittel et al (2020) reported that VR video training was just as effective as conventional video footage for improving and retaining decision making in professional Australian rules football umpires. However, VR rated higher for enjoyment and relevance compared to convention video training, which may contribute towards adherence. Our finding on coaches/support staff overall agreement for VR to be used to improve tactical awareness is consistent with those found by Thatcher et al. (2020), who reported coaches commenting that VR could be used in team meetings for position-specific units, for example with central midfielders to visualise passages of play.

There was also a disagreement shown for VR being used to improve physical abilities, such as strength, power, and aerobic fitness. Over the last decade, extensive research has contributed to developing the physical abilities of professional soccer players. As such, the need for VR to be used in this area is perhaps not a necessity. That said, in other populations, VR has been used effectively to improve strength and balance more so than conventional training methods (de Rooij et al., 2016). This may be partly due to the ability of VR to increase adherence to exercise through motivation and enjoyment. While the player sample were not included within our model, nor were inferential comparisons made between groups, the use of VR to improve technical ability revealed contrasting opinions between coaches and players, with players showing greater disagreement than coaches (Figure 5.1). As our player sample consists of those on a professional contract, it could be assumed that technical proficiency has already been achieved at that level, and that a greater emphasis is placed on tactical mastery.

The current study also aimed to understand the perceptions of how VR should be used within professional soccer. These perceptions were not included as part of the model but used to get a firmer understanding on the direction of VR within the sport. Coaches and support staff generally agreed that VR should be used for performance analysis purposes, followed by using it as a preparation tool (visiting new environments, i.e., stadiums) and for rehabilitation (Figure 5.2). Regarding rehabilitation, VR may provide players with the opportunity to continue to train mentally in the absence of soccer-specific training completed outside. This could be achieved through soccer-specific drills using a virtual ball, thereby eliminating physical contact. Alternatively, those in the early stage of rehabilitation where movement is not possible (such as post-surgery), can revisit match footage, viewing the game from their own perspective. Our results support the views expressed previously by coaches in professional soccer, who have indicated the value in VR being used during player rehabilitation (Thatcher et al., 2021). Further, this finding is supported by research showing the potential of VRs ability to promote a dissociative attentional focus, distracting the user from the exercise performed (Mestre et al., 2011) which in the case of the study by Gokeler et al. (2016) resulted in patients with anterior cruciate ligament reconstruction having a greater movement proficiency while using VR, compared to not using VR.

As for coaches' perceptions on who VR should be used with, respondents generally agreed that the technology should be used with senior academy players (U18s and U23s), whereas a greater inconsistency was shown in agreement for junior players (under 11s) (Figure 5.3). It is plausible that VR could be seen as a tool that utilises the 'marginal gain' principal, making it a valuable tool when bridging the gap from academy to first team soccer. In contrast, long-term athletic development models emphasise the learning of fundamental sport skills in children under the age of nine (Pichardo et al., 2018). This, combined with the limited contact time that coaches have with children in the younger age categories, could have contributed to the greater disagreement.

The second largest contributor in our model was facilitating conditions, which had a negative relationship with likeliness to use (Table 5.8). In terms of barriers to using VR within professional soccer clubs, coaches and support staff reported that the limited evidence base and time to use the technology were the greatest overall barriers. While there is a growing evidence base to support the efficacy of VR in sport, its applications for improving soccer performance are not available. As such, this may lead to early scepticism of VRs value, creating beliefs that the technology is a 'gimmick' or 'novelty' (Thatcher et al., 2021). Time was also regarded as a barrier to using VR (Figure 5.5). Coaches and support staff are already under time constraints, and so the question is, when do they fit in the time to use VR? Previous surveys on practitioners working within professional soccer have indicated time as a barrier to conducting applied collaborative sport science research with university academics (Malone et al., 2019). Our model also indicated first impression as a significant contributor to the facilitating conditions construct (Table 5.6). First impression bias refers to a limitation in human information processing whereby individuals are strongly influenced by the first piece of information they receive, and that future information is biasedly evaluated to fit the narrative of the original information (Lim et al., 2000). In our study, 55% of coaches and support staff indicated that hearing, seeing, or using VR was a barrier towards its future use (Figure 5.5). This finding is interesting, as the results may suggest that there is initial scepticism towards the technology without deeper considerations of how it could be beneficial within the training ground environment. While not included in our model, monetary cost was also indicated as a barrier for using VR, albeit in a reduced sample (n = 36) that were aware of the associated cost. These findings resonate with those of Segal et al. (2011), who reported monetary cost to be the greatest barrier for adoption of VR in psychotherapists.

Social influence and technology readiness had small (f^2 = 0.03) and no effect (f^2 = 0.003) on likeliness to use, respectively. This finding is consistent with a previous study reporting that social influence and technology readiness did not contribute to the behavioural intentions of recreational golfers to use smart technology, such as VR and GPS (Seol et al., 2017). Further, social influence has been reported to be a non-contributing factor in the intention to use new technologies for rehabilitation in physical and occupational therapists (Liu et al., 2015). One possible reason for social influence not being a contributing factor in VR adoption is the voluntary context in which participants anticipated they would use this technology. Previous research has shown that when technology use is explored within voluntary contexts, its contribution to technology acceptance is limited (Venkatesh et al., 2003). Opposing this, when its use is mandatory its contribution to technology acceptance is greater (Venkatesh et al., 2003). Coaches and support staff in professional soccer can deploy technologies within their clubs as they see appropriate, without any rules being dictated by governing bodies. It is therefore understandable that their likeliness to use VR if given access is independent from socially originated influences such as social conformity or how they are perceived by others. That said, an interesting insight in our study was the high overall perception of being influenced to use VR if senior staff wanted the technology used in their clubs (Figure 5.4). It is possible that should clubs impose mandatory use of the technology through senior stakeholders, coaches and support staff may look to other clubs and individuals that are using VR to influence their own practice.

Generally, coaches, support staff, and players shared a similar opinion of VR, with very few indicating negative opinions of the technology (Figure 5.6). The data illustrates a high percentage of neutrality which could lean onto either side of positivity or negativity. However, the low number of negative opinions is encouraging in the prediction for the future use and adoption of VR in soccer. That said, it is important to note that in the coaches and support staff sample, 86% of participants had no access to VR in their clubs, and 70% had never used a soccer specific VR system (Table 5.3). Their limited experience with this technology, despite half indicating positive views is an interesting finding, and one could look towards models such as the Gartner hype cycle as an explanation of this. The Gartner hype cycle model, developed by Gartner Inc in 1995, is used to explain the generalised evolutionary path that technology takes over time (Dedehayir & Steinert, 2016). In the early stages when adoption rates are low, an overly positive reaction to the technology is seen through people's attraction to novelty, social contagion and unclear attitudes towards decision making (Fenn & Raskino, 2008). However, once the period of 'hype' and overenthusiasm has passed, organisations and users of the technology experience dissatisfactory results that do not match up with their expectations, causing some users to abandon the technology (Trough of disillusionment). Further investments in the technology allow for greater applications, knowledge, and socialisation (slope of enlightenment), which finally, leads the technology to be realistically valued in the marketplace, and adoption begins to accelerate (Plateau of productivity). As of 2016, VR was recognised on the Gartner hype cycle within the slope of enlightenment (Wohlgenannt et al., 2020), indicating that VR had begun to find its place in the market. However, this has been largely due to the gamification of VR which differs to how it is used in soccer. Virtual reality in soccer might therefore be in an early stage of 'hype' where adoption rates are low, and its future success still unclear.

5.4.1 Limitations

Like any other study, especially where the UTAUT model has been applied, limitations in the research are evident. First, it is important to note that while the questions included in our survey to form the independent constructs were based on solid theoretical evidence and observations of VR use in professional soccer (and other fields), our constructs may not fully represent what they are trying to convey. For example, in the social influence construct we did not include a question on the influence of technology companies, who may provide bold claims on the benefits associated with having VR (Gamble et al., 2020). This inclusion, as an example may have revealed differing results to the construct. Second, the survey was made available to anyone working within a professional soccer club, irrespective of how many respondents may have come from the same club. This is in contrast to previous survey based research, where only one respondent is permitted to complete the survey in order to reduce respondent bias (Towlson et al., 2013). While it was always the intention to understand individual perceptions, it is possible that if respondents came from the same club and shared a club based philosophy on the use of VR, then this may have inflated the results in given direction (Thatcher et al., 2021). Third, the present study was only able to capture the responses of three coaches/support staff that reported themselves as being responsible for operating VR within their clubs. To the authors knowledge, at least ten professional soccer clubs are using VR, and so our study may not fully represent the ways in which VR is being used. Although contact was made with the ten soccer clubs, lack of response may be attributed to an attempt to protect club privacy, personal agenda or respondent fatigue while completing the survey. The fourth limitation within the study was the timing of the COVID-19 global pandemic, causing the majority of soccer leagues to suspend matches. Further, some countries imposed a nationwide lockdown preventing training ground access. During this time, we continued to collect data and it is possible that the respondents perception of VR may have been influenced because of necessity to train in isolation, a tool in which VR has been recommended (Melim et al., 2020).

5.5 Conclusion

Virtual reality use in professional soccer is still a relatively new technology that has been adopted by a small number of clubs. Prior to this study, little was known about the current perceptions, influences and barriers of VR and their contribution towards future VR adoption. Our study allows us to conclude that performance expectancy is the largest contributor towards likeliness to use VR in professional soccer. In other words, the belief that VR would improve soccer related performance was the most important factor in determining adoption of the technology. Additionally, social influence and an individual's technology readiness did not contribute to likeliness to use. However, our results revealed facilitating conditions as a significant negative contributor towards likeliness to use. In other words, the greater a person believes that barriers are in place to using VR, the less likely they are to adopt the technology.

Chapter 6 General Discussions, limitations, future research recommendations, conclusions, and practical applications.

6.1 General discussion

The central theme of this thesis was to explore the efficacy of VR use in professional soccer clubs. The first aim of this thesis was to examine the familiarisation properties of a VR soccer passing test, by assessing the evidence for (1) a learning effect, and (2) familiarity being achieved as indicated by a plateau in performance. On achieving this the second aim of the thesis was the examine the concurrent validity of a VR soccer passing test, by comparing the number of passes achieved to an equivalent test in the real-world. Further, the thesis looked to explore the subjective task demands of both VR and real-world conditions, along with the presence, ecological validity, and engagement captured through the virtual reality condition. As such, the first two studies of this thesis addressed gaps in the literature regarding the measurement properties required for testing the technical abilities of soccer players in VR, along with the perceptions of participating in VR soccer activity. The third aim of this thesis looked to explore the perceptions of coaches, support staff, and players on the use of VR in the professional soccer training ground environment. Further to this aim, the thesis looked to explore the mechanistic factors that might contribute towards VR adoption in the future.

Although VR has existed for over 50 years, it is only been in recent years that the technology experienced an exponential rise in usage across multiple industries due to an advancement in computer quality and power, reduced monetary costs, and greater accessibility (Bird, 2020). Industries including professional soccer have taken an interest in this technology, which has in part been made possible due to a number of VR companies specialising in solutions that are specific to soccer (Thatcher et al., 2021). While these companies specialise in different solutions for professional soccer (e.g., goalkeeper training, rehabilitation, performance analysis etc), they all share a commonality, being the ability to test soccer players (Thatcher et al., 2021). However, before any test is conducted, it is first important to assess if sources of systematic bias exist that might affect the data collected. One of these systematic bias's is the learning effect, and it is important that researchers and practitioners assess (1) if familiarity to a test is necessary, and (2) if it is necessary, the implement of familiarisations trials (Hopkins, 2000). However, no study had investigated if familiarity is required before completing a soccer specific test in VR, or in any other sport VR simulator for that matter. **Chapter 3** sought to investigate this, by assessing the learning effect of a soccer passing test in VR, along with determining the point at which familiarisation had been achieved.

One of the key findings from this thesis was that a learning effect did exist during a VR soccer passing test, with total goals scored rising from 13.6 (3.3) during trial 1, to 16 (3.3) in trial 2 (BF_{10} = 46.5, *d* = 0.56 [95% CI = 0.2 to 0.92]). This finding highlights the impact that the learning effect could have on a VR test, especially when the test is used as part of an intervention study to attain pre-post difference scores. For instance, Gray (2017) did not familiarise participants to a VR baseball batting test before being prescribed to a six week training period of either extra real-world baseball batting practice, VR baseball batting practice, or no additional training (control group). Following the intervention period, all groups improved VR baseball batting performance, including the control group that completed no additional training. Our findings would suggest that the improvement in performance witnessed in the study by Gray (2017) might have been contributed by participants learning how to perform better in the VR environment. Several reasons might have contributed to the learning effect found in **Chapter 3**. First, just like in the real-world, learning effects have been evident in football skill tests, even in experienced football players (Runswick et al., 2022). For instance, experienced football players completing the visual impaired football skills test were shown to improve their performance (as indicated by a reduced task completion time) between the first and second trial (*p* = 0.04), and the first and third trial (*p* = 0.004), (Runswick et al., 2022). Further, learning a task in VR has its own unique challenges for the participants when compared to learning an equivalent task in the real-world. One of these challenges was learning how to interact with a virtual ball that provided no haptic feedback during contact with the foot. Haptic feedback refers to the sense of touch and proprioception, and has been evidenced as being beneficial to performance during a VR surgical task, when compared to an identical VR surgical task without haptic feedback (Gani et al., 2022). In **Chapter 4**, several participants identified that not being able to feel the ball during a VR passing task might have contributed towards the task difficulty, when compared to a real-world equivalent task. This might have contributed to the observably higher subjective ratings of task difficulty when trying to control a ball in VR compared to the equivalent task in the real-world (Figure 4.5).

Another key finding from **Chapter 3** was that a plateau in performance was evident after completing five trials of a VR passing task, with strong evidence of equivalence shown between trials five and six (Figure 3.5). These findings are consistent with the learning curve literature in VR surgical tasks, that have shown a plateau in performance after five (Bartlett et al., 2020), and eight trials (Bharathan et al., 2013). As such, our findings show that multiple trials of a VR passing task are necessary before evidence of the learning effect appears to diminish, and soccer clubs looking to test players should conduct multiple familiarisation trials. That said, from a pragmatic point of view, **Chapter 5** identified the time available to operate VR as a key barrier towards its adoption in the professional soccer training ground environment (Figure 5.5). While familiarisation might be valuable for improving the quality of the data collected by reducing any systematic bias, it might inadvertently discourage those working in soccer from implementing VR for the purpose of testing players. To alleviate this, research should investigate strategies that would promote the learning of ball passing in VR, with the intention of establishing task familiarisation, quicker. In **Chapter 3**, although participants were informed on the rules and aims of the test before it commenced, participants took on a self-guided learning approach during the test whereby no advice was given by the researcher on how to control the ball in terms of stopping or passing it. For example, where in the real-world footballers might control the ball with the sole of the boot, this technique was observably difficult to do in VR due to the inability to apply pressure between the ball and the foot. Further, participants were not made aware that they would not feel the ball during the task, or that their field of view would be limited in comparison to the real-world. This approach of providing no verbal ques or guidance during the test was done deliberately to standardise the test between participants, but might have inadvertently prolonged the familiarisation process, as verbal ques have been shown to aid in task learning. For instance, external verbal cues have been shown to influence metrics such as jump height during the countermovement jump compared to when no verbal cues are provided (Barillas et al., 2023).

Having established the existence of a learning effect, and the point at which performance plateaued in **Chapter 3**, **Chapter 4** of this thesis sought to investigate the concurrent validity of a virtual reality soccer passing test, by comparison of passing ability between virtual reality and real-world conditions. This was achieved by assessment of the rebound 135 skill test from the UQSFA, in both real-world and VR environments. The main finding in **Chapter 4** was that the number of passes completed in VR was lower compared to the real-world (EMM = -3.9, 95% HDI = -5.1 to -2.7), and that 100% of the differences fell outside the ROPE criteria that was set *a priori* as a 2-pass difference. The finding of performance (as determined by number of passes) being lower in VR compared to a real-world equivalent task is comparable to other concurrent validity VR studies, such as that of Harris et al., (2021), who compared a 3 m golf putting task in both VR and real-world conditions, and found a reduced performance (measured by distance of the ball to the hole) in the VR condition. Several reasons might have contributed to a reduced performance in VR, including topics previously discussed in this chapter, such as a lack of haptics, a modality that was also missing in the study by Harris et al. (2021). Another potential reason could be due to the VR equipment used in the study. Participants were fitted with a head mounted display that was tethered to the PC via a wire, and many reported ‘feeling’ the wire during the completion of the task, which might have distracted the participant and potentially restricted movement. Similarly, Ferrer, Shishido, Kitahara, and Id, (2020) also indicated participants reporting the feel of the wire during a VR soccer task, and theorised that the distraction of the wire might have severed the concentration of the participant, disconnecting them momentarily from the virtual environment. However, this theory is not supported, with research showing that wireless VR headsets did not enhance presence after playing a VR game, in comparison to the same game played with a wired headset (Goncalves et al., 2020). However, in the same study, anecdotal evidence from participants indicated that the cable did interfere with the experience, such as being tangled in the wires, or fearing tripping over the wires. This would suggest the use of wireless headsets should be considered, although these also have their limitations, such as instability in signal when configured to a PC set-up, causing blackouts of the image and sound to occur (Goncalves et al., 2020). As a compromise, wired headsets with cable management systems should be considered.

A key finding from **Chapter 4**, as already discussed, was that the number of passes in VR was lower than in the real-world, and based on the ROPE it was not possible to conclude that the two conditions agreed with one another. This type of agreement can be considered as the 'absolute agreement'. However, even without absolute agreement being established, it is possible that those who perform better in the real-world test would also perform better in the VR test. This type of agreement is known as the 'relative agreement'. As such, a second aim of **Chapter 4** was to assess the relative agreement by assessing if better performing players in the real-world also performed better in the VR test. Using nine quantile distributions it was observed that the differences in number of passes increased from the 40th to the 90th percentile, such that better performing participants in the real-world had greater differences between their real-world and VR passing scores, than worse performing participants (Figure 4.3). These observed differences are due to certain participants performing better in the real-world, than others, but not necessarily performing any better in the VR test. In other words, those who performed better in the real-world did not perform better in VR. In **Chapter 4**, a sample of the population consisted of professional academy football players, and it is hypothesised that the better performing participants in the real-world came from this group. In fact, research by Wilson et al. (2016) has shown the 135 rebound board test (used in **Chapter 4**) to discriminate between players of differing ability levels, so it would be no surprise to expect academy players to perform better in the real world compared to lesser experienced soccer players, such as amateurs who were also used in the study. However, it appears that this enhanced technical ability does not transfer to VR, which could be due to the observable force an advanced player applies through the ball when passing in the real-world compared to a less advanced player. Because the equivalent force could not be applied in VR, it is likely that advanced players had to adapt their normal style of passing, thereby losing any advantage they had in the real-world. That said, because we used competitive soccer players over the age of 16 who all had observed competency kicking with both legs (to varying degrees), it is hypothesised if lesser experienced populations were included in the study design (e.g., youth footballers) that have not yet developed the coordination to pass with both legs in the real-world, that this could have also been evident through the VR test. As such, results may have shown that less competent players in the real world, also performing less competently in VR.

Practically, soccer clubs are recommended to remain cautious about using VR to distinguish between soccer players based on their ability to pass the ball continuously at short distances. While this might not be in the interest of those working in professional soccer based on notable disagreement for VR to be used for technical skills such as passing (Figure 5.1), it is possible that there could be implications if dynamic passing is used as a method of response in VR tests that are assessing other components of soccer related performance, such as decision making. To illustrate, Ferrer et al. (2020) used VR to measure visual exploratory activity of soccer players during a football specific match scenario where players had to decide which team mate to pass the ball following observing the scenario played around them. However, to decide which player to pass to, the participant had to fixate on the player of their choosing and complete a kicking gesture to respond. As such, although a kicking action was involved, it had no impact on the direction the ball travelled and so lacks intuitive and natural behaviours that are a key feature of VR (Slater & Sanchez-Vives, 2016). Therefore, it could be argued that a more natural kicking response is required whereby foot-to-ball contact is made, without fixation on a player, as not all ball passes (in the real-world) would require this (i.e., players might be looking at the ball, or the opposite direction when making a pass). Yet, for this argument to have meaning, it is important that players can kick the ball accurately and naturally, at an ability level that is expected for that individual in the real-world. The results of **Chapter 4** would suggest that this is not possible, therefore, if players are required to pass a ball as a method of response in a decision-making test, their decision-making ability might be inaccurately recorded due to the ball not accurately moving in the intended direction.

In addition to assessing the absolute, and relative agreement (validity) between the two conditions, **Chapter 4** also aimed to assess the representativeness of the virtual environment through measures of psychological fidelity. One of the findings was the observably lower subjective physical demands in the VR condition by comparison to the real-world condition (Figure 4.5). This observation is consistent with a study by Harris et al. (2021) who also observed greater subjective physical demands in the real-world compared to the VR condition, albeit non-significant. In the current study it is likely that the subjective physical demands were lower in VR due to the reduced space in which players could move around the virtual environment, which was restricted to a 1 m radius (see chapter 4.2.3). This was a default safety feature as set by the software, and although the player could move out of the playing circle and still interact with the virtual ball without any interference, doing so would cause the screen to dim, warning the participant to move back into the circle. As such, it is possible that this might have restricted the movement of the participant, relative to the real-world condition. Additionally, the physical demands might have been lower in VR due to a lack of a physical ball, which removes impact forces during foot to ball contact, that are estimated to be approximately 1 – 1.1 kN during maximal effort ball kicking (Tsaousidis & Zatsiorsky, 1996). As a result of the removed contact forces, it is hypothesised that kicking in VR requires a reduced muscular activation to create ball momentum during the swing phase of the kick. Furthermore, it is likely that during the VR test, alternations in muscular activation existed in the stance leg during ball contact, which in the real-world acts to stabilise the lower extremity through flexion at the knee through a slow contraction velocity, that enables muscles to generate their highest forces (Lees et al., 2010). Consequentially, these biomechanical alterations could have been a reason as to why the number of passes was lower in VR compared to the real-world. While this might be a disadvantage for the purpose of testing soccer players on their passing ability, reduced physical demands (by a result of less force) might be advantageous during the rehabilitation of football players where the management of external 'load' is important (Taberner et al., 2019). In fact, 73% of coaches, and support staff agreed in some capacity that VR should be used with players undergoing rehabilitation (Figure 5.2). As such, reduced force on the body might allow soccer specific stimulus (e.g., kicking a ball) to be introduced into the return to play protocol earlier than is usually possible. However, reducing external 'load' on the body is not the only reason as to why those at soccer clubs might believe VR to be a useful part of player rehabilitation. Another reason could be the ability of VR to influence alterations in joint biomechanics following an injury by promoting a dissociative focus of attention, whereby the environment acts as a distractor to the exercise performed (Gokeler et al., 2016). This could be through the ability of VR to decrease internal focus of attention of joint motion (i.e., flexion at the hip) and more on an external focus (i.e., kicking a ball) that allows for a more efficient movement pattern, and shown to be advantageous for rehabilitation and return to play (Gokeler et al., 2015). Alternatively, VR's perceived usefulness for player rehabilitation might be unrelated to physiological adaptations but might be more psychologically orientated, such as increased engagement and therefore adherence to the rehabilitation program. However, it must be acknowledged that subjective perceptions of physical exertion to compare real-world and VR conditions might be misleading, especially given that VR has been shown to reduce perceived pain exertion (Ioannou et al., 2020). As such, it is possible that the VR condition was equally as physically demanding as the real-world, but perceived exertion was lower in VR due to potential factors such as the virtual environment acting as a distractor. To address this, future studies should examine the acute physiological response to physical activity in VR, such as heart rate or oxygen consumption, and compare these against the real-world.

Chapter 4 also examined the participants' sense of presence during the VR condition, which is defined as the person's illusion of being in the scene displayed by the VR system (Slater et al., 2010), and is often referred to as a psychological sense of 'being there' (Slater et al., 1995; Steuer, 1992). Understanding the sense of presence is important in the context of using VR in soccer because with greater levels of presence comes more opportunity for users to behave physically and emotionally as they would in the real-world (Slater & Wilbur, 1997). Results from **Chapter 4** showed participants to have a high sense of presence based on their spatial presence, engagement, and ecological validity scores, and are comparable to research by Le Noury et al. (2020) who also reported high sense of presence during a VR tennis simulator using the same sense of presence questionnaire. However it should be noted that the three constructs of presence were observably higher in the study by Le Noury et al. (2020). For example, engagement was moderately higher (*d* = 0.69) in the tennis VR study (M = 3.89, SD = 0.59) compared to our own study (M = 3.48, SD = 0.6). This might be explained by the competitive nature of the tennis environment, whereby participants had to compete against a virtual opponent. Previous research on VR rowing has demonstrated greater motivation when the VR environment includes an on-screen competitor to challenge the participant (Parton & Neumann, 2019). Spatial presence, described as the sense of 'being' in the virtual environment, was also greater (*d* = 0.39) in the tennis environment (M = 3.6 SD = 0.58) compared to our own study (M = 3.39, SD = 0.47), which might be explained by the aesthesis that were incorporated such as a scoreboard, crowds, and large buildings that surrounded a court. Such detail might have captured the attention of the participants, distracting them from the real-world, and provided a sense of what it would be like to compete in a major tournament. This is especially so given the sample age (12-17 years) used in the study by Le Noury et al. (2020), as many are unlikely to have competed in an equivalent environment in the real-world. In contrast, the environment used in **Chapter 4** resembled an indoor soccer training centre used by professional soccer teams, that featured no crowd, noise, or scoreboard. Yet, it is unknown if the sense of presence scores captured in **Chapter 4** had any influence on the participants performance in the VR condition. To which, it is unclear if those who had higher sense of presence scores, also performed better in the VR condition. Further, it is unclear if adapting the VR environment to have features more akin to Le Noury et al. (2020) would have improved these scores, and subsequently, if performance would have improved as a result.

Having assessed the concurrent validity and subjective response to a soccer passing test in VR (**Chapter 4**), **Chapter 5** of this thesis looked to explore VR in the professional football training ground in a broader sense by measuring the perceptions of how it should be used, its influences, and barriers towards its use, in a sample of coaches, support staff, and players currently employed at professional soccer teams. Further, **Chapter 5** aimed to establish the extent that these factors would have on the future adoption of VR in soccer coaches and support staff who had no access to VR in their clubs. One of the key findings of **Chapter 5** was the lack of consensus for VR being used as a testing tool (Figure 5.2). While a marginal majority agreed in some capacity that VR should be used to test players (54%), the majority of those that agreed only did so 'somewhat'. Furthermore, 29% disagreed with this statement, and 17% remained neutral which might be perceived in either direction of agreement. This finding is interesting given numerous VR companies that specialise in football are promoting the technology on the basis of testing soccer players (Thatcher et al., 2021). On the side of agreement, VR might be seen as beneficial at providing a controlled and repeatable environment to test characteristics of footballers such as visual exploratory activity during a simulated soccer match (Ferrer et al., 2020; Wirth et al., 2021). Such characteristics would not be possible to measure in the real-world due to the lack of controllability in the environment (e.g., weather constraints), plus the impracticality of requiring multiple outfield players to create each scenario. In contrast, those on the side of disagreement might see VR's use as an added burden to the coaches and players time, especially given the number of tests that are already applied in professional soccer on a daily or weekly basis (Svensson & Drust, 2005).

Chapter 5 also identified that half of coaches, support staff, and players viewed VR as a positive technology, for use in the professional soccer training ground setting, with less than 15% identifying negative opinions of the technology (Figure 5.6). This finding is particularly interesting given that 86% of participants had no access to VR in their clubs, and 70% had never used a soccer specific VR system (Table 5.3). As such, it is possible that many of the participants provided their opinions of VR based on socially orientated influences such as what others (clubs, individuals) are doing, information attained through social media or directly from VR companies themselves. This information provided might have given an idea of what VR is capable of, leading to a sense of 'hype' and enthusiasm seen in the early stages of technology adoption which can be visualised through models such as the Gartner Hype Cycle (Dedehayir & Steinert, 2016). Following the hype stage, organisations and users of the technology experience dissatisfactory results that do not match up with their expectations, causing some users to abandon the technology. After the decline in use, technology begins to find its realistic value in marketplace which leads to accelerated growth, or a steady inclination of further adoption. As of 2016, VR was recognised on the Gartner Hype Cycle within the slope of enlightenment (Wohlgenannt et al., 2020), although this is a generalist view of VR and not its specific use in sport. As such, it is possible that sports such as soccer are going through a stage of 'hype' and its future success is yet to be known. This theory also has implications for some of the findings already reported in **Chapter 4,** where participants identified the VR condition to be engaging and illicit a high sense of presence (Figure 4.6). Of the 29 participants that took part, 24 of them had never used VR before, and it is possible that these results could be influenced by the 'hype' and novelty of experiencing VR for the first time.

6.2 Limitations

The studies that form this thesis have their limitations which are addressed in this section. One of the key limitations of **Chapter 3** and **Chapter 4,** is the limited generalisation, in that the findings of these studies only relate to dynamic soccer passing ability and its relation to how the skill is learned, and validated in VR, respectively. As such, the findings from these studies can only be applied where VR includes a dynamic ball passing component, and other technical components of soccer such as heading, or throwing a ball require investigation. Further, it is acknowledged that both tests required participants to pass a ball that was in motion towards them, and in **Chapter 4** this required participants to control the ball with the non-passing foot. Because we did not consider controlling a ball to be an individual technical component of soccer and was considered to be a part of passing ability, it is unknown if the findings from **Chapter 4** would have differed had the ball been stationary, such as found during a goal-kick, free-kick, or penalty situation. Future studies should look to assess these two technical components independently.

Another limitation of **Chapter 3** and **Chapter 4** was the lack of validation of the ball physics and the physics of the environment (such as the rebound boards) in relation to the real-world before either study commenced. This is an important first step for VR that requires physical interaction with an object, as demonstrated in studies exploring table tennis (Oagaz et al., 2022), and billiards (Haar et al., 2021). For example, Haar et al. (2021) validated the physics of the cue ball in VR by merging the virtual and real-world environment so that interaction with the environment was simultaneously aligned. In other words, a real pool table and cue ball was linked to a virtual pool table so that interactions with the cue ball in VR corresponded to interactions with a cue ball in the real-world. The results showed with-in shot comparisons between conditions to be highly consistent for the direction of travel (Pearson correlation r = 0.99) and velocities (Pearson correlation r = 0.83). If a similar approach had been established *a priori* **Chapter 3 and 4**, it is possible that alterations could have been made to the coding language that might have, for example, made the ball control more representative of the real-world. This in turn might have led to alternate findings such as improved task learning (Chapter 3), or agreement in passes between conditions (Chapter 4).

It is also worth acknowledging the way in which the VR was introduced to the participants in **Chapter 3**, as perceptions of the technologies use in soccer might have influenced how participants engaged with the task. In **Chapter 3**, a professional cohort of soccer players were recruited to take part in the study which took place in the soccer training ground environment. At the time of the study, the VR equipment was not used for purposes other than research and was not implemented as a part of the soccer training program. As such, players might have not bought-in to the technology and lacked motivation to perform in the test to the best of their capabilities. This is a possibility given that 11% of players perceived VR negatively, and 39% were indecisive when gauging its overall use in professional football (**Chapter 5**). In contrast, other technologies such as force plates might be looked upon more favourably in the training ground environment knowing its potential to inform training programs to improve performance, or monitor fatigue (Brownlee et al., 2018; Loturco et al., 2018)

Another limitation to keep in mind is how we determined the familiarisation equivalence interval for the test in **Chapter 3** (within a two goal difference between trials 4, 5, and 6), and what we considered as a practical equivalence (and therefore evidence for validity) between conditions in **Chapter 4** (within a two pass difference). For both chapters, unpublished data was used as a justification for the equivalence zones, which was then proceeded by logical reasoning to examine if these equivalence zones made sense in the eyes of a coach or support staff who would interpret the results. For example, **Chapter 4** used unpublished data by independent researchers that had collected data from repeated trials in the real-world soccer passing test used in our study. Based on the repeated measures means (trial 1 and 2), and the pooled standard deviation, a two-pass difference was calculated to equate to a standardised effect size of 0.4 either side of 0. Additionally, a two-pass difference was determined as an acceptable parameter of equivalence because a passing difference of this magnitude would probably not be considered worthy of intervention by a coach. However, we did not seek the input of coaches to verify our parameter, to which alternative equivalence zones might have been suggested and deemed more appropriate. For instance, whilst a smaller equivalence zone (1 pass difference) wouldn’t have changed the findings of the chapter, a larger equivalence zone (e.g., 4 pass difference) would have had considerable ramifications on the study’s findings.

Finally, it is acknowledged that a limitation in the PLS-SEM (**Chapter 5**) was the exclusion of the barrier ‘cost’ for the formative construct ‘facilitating barriers’. Initially, this barrier was not included in the model on the basis that only a small proportion of coaches and backroom staff

(36/143, 25%) were aware of the associated costs of VR in professional soccer. Had this data been included, it would have only represented those with knowledge of the cost, as participants without knowledge of the associated cost were not asked if they perceived it to be a barrier. At the time, it was argued that a small representative sample of data could have biased the results, and therefore this variable was removed. In hindsight, it was not considered that even if participants were not aware of the monetary costs, their perceptions might have been valuable based on their knowledge of technology (in general) being paid for by the clubs/teams they were associated with. For example, a coach with a lower budget (e.g. working in a lower tier) might have perceived cost to be a barrier on the basis that most technology acquirement is subject to budget constraints. This point of view might have differed in a team with a larger budget.

6.3 Future research recommendations

The main aim of this thesis was to investigate the efficacy of virtual reality being used in a professional soccer training ground environment. Some of the key findings and limitations of this thesis have led to several research questions that could be addressed in future. As discussed in **Chapter 3**, clear evidence was shown for a learning effect between the first two trials of a VR passing test, and potentially multiple trials are required before a plateau in performance is reached. Although multiple factors are likely to have contributed towards the familiarisation of the task, a key contribution could have been the lack of verbal cues and guidance on how to interact with the virtual ball. This is important, especially given the perceived difficulties of interacting with the VR ball observed in **Chapter 4**. Given the effectiveness of verbal cues at improving learning and technique (Barillas et al., 2023), researchers should look to assess what verbal cues could effectively contribute to participants ability to control the ball in VR, as such cues could increase the rate of learning, thereby reducing the time needed to familiarise to the test.

Chapter 4 of the thesis identified a high sense of presence as measured through the ITC-SOPI presence questionnaire (Figure 4.6). While this is reassuring for the purpose of this study, it is unknown what influence it had on participants' performance. While some research has alluded to higher presence being linked with improved task performance in VR (Grassini et al., 2020), others have suggested that higher presence does not directly lead to improved performance, but instead has an indirect effect by increasing the likeliness of those in the VR environment responding naturally to stimulus using their own natural skills and senses as they would in the real-world (Slater & Wilbur, 1997). For instance, Le Noury et al. (2020) assessed the lower body stance of participants during a VR tennis task, and found that it corresponded with the same stance adopted in the real-world. It is possible that those who were identified as having a lower sense of presence could have responded differently (in terms of their stance) to how they would in the real-world. Future research should aim to explore this by either correlating presence with performance variables of a soccer VR test, or response-actions from the VR environment that correspond to natural skills and senses found in the real-world. Future research might also look to assess the influence that haptic feedback, and /or wireless has on task performance, as those technological components could contribute to user experience and performance.

Chapter 5 of the thesis identified generally positive attitudes towards VR despite limited adoption or even use of a soccer specific VR system. This led to the theory of 'hype' and overenthusiasm towards this technology in the early stages of adoption, which based on the Gartner Hype Cycle is followed up by disappointment and abandonment due to a lack in expectations (Dedehayir & Steinert, 2016). While this theory could be true, our study did not validate this by exploring perceptions of those who might have had previously adopted VR in their clubs, but at the time the survey was conducted, they did no longer do so. Such information, if it existed, would provide greater transparency on how the technology has been adopted. Further, research might look to longitudinally explore the perceptions of those currently using VR, to see if their perceptions change over time.

6.4 Concluding remarks

Commercially available VR solutions, designed to test characteristics of soccer players have appeared in recent years (Thatcher et al., 2021), but limited research exists to support its use in this area. The ability to test 'performance' in VR is advantageous for numerous reasons outlined in detail earlier in this thesis, such as to replicate controlled environments, or to create testing environments that are not physically possible in the real-world (See chapter 2.3.5). However, in the rush to sell, use, or research soccer specific VR testing solutions, many have forgotten the fundamental measurement properties of a test. To that end, the thesis first looked to explore some of the fundamental measurement properties of soccer related passing skill tests, that are necessary before being used in professional soccer. The main findings were (1) evidence of a learning effect from repeated trials of a VR soccer passing test (**Chapter 3**), (2) several repeated trials are necessary for the number of passes to plateau in VR (**Chapter 3**), and (3) lack of evidence to support the concurrent validity of a VR passing test, when compared to a real-world equivalent (**Chapter 4**). These findings have several implications for the way in which VR is used in professional soccer environment. The thesis then looked to explore VR in a more generalist sense, by attaining the perceptions of professional coaches, support staff, and players (**Chapter 5**). Main findings showed that intention to use was primarily driven by 'performance enhancement' beliefs, in contrast to socially orientated ones, yet many barriers exist that might prevent its wider adoption. Further, perceptions of soccer coaches and support staff revealed that its desirability of use is not evident in VR's ability to test soccer players, but for use in other areas such as rehabilitation, and performance analysis.

6.5 Practical applications

1. Coaches and support staff who plan on using VR for the purpose of testing players should be mindful of systematic bias such as the learning effect which might affect the results if familiarisation trials are not implemented. At the very least, one familiarisation trial should be implemented before the test commences, although it is possible that multiple familiarisation trials might be necessary before performance in the task begins to plateau.
2. Coaches and support staff should remain cautious of VR tests that intend to assess the technical passing abilities of soccer players. Furthermore, it is important to consider that if a VR test is implemented to assess other qualities of soccer players, such as decision-making ability, the results could be affected by a player's inability to pass effectively by comparison to how they would pass in the real-world.
3. Virtual reality companies who specialise in soccer should be considerate of how professional soccer teams perceive VR should be used, and what barriers might prevent VR from being implemented in practice. By doing so, companies can develop VR solutions that meet the needs of the end-user and provide greater buy-in during the technology's early stages of adoption. Notably, this might include applications for rehabilitation and performance analysis, and focusing on improving cognitive components of soccer players. Further, VR companies might alleviate some of the important barriers to VR adoption such as the limited evidence base by working closely with universities to upscale academic research and develop VR applications that are effective to use in the limited time that soccer coaches and support staff have available to use the technology.

Chapter 7 Reflective account of an embedded researcher in a professional soccer club

7.1 The embedded researcher

Embedded research typically describes the mutually beneficial relationship between academics and a host organisation, whereby research is conducted in a public, private, or third sector organisation (McGinity & Salokangas, 2014). For the academic researcher, greater access to the host organisation comes with benefits such as data collection privileges (e.g., access to participants) which might be inaccessible without the relationship. For the host, the relationship provides access to academic resources and knowledge needed to answer research questions, and ultimately develop organisational policy or practice (McGinity & Salokangas, 2014). Embedded research comes with a number of other perceived benefits such as encouraging researchers to be more mindful of the hosts needs, and equally, encouraging the host and employees within it, to be more mindful of scientific evidence (Vindrola-Padros et al., 2017). Perceived benefits also include answering research questions that then have an immediate impact on the organisation, and the opportunity for the researcher to attain ‘insider knowledge’ of the organisation that can be used to inform the research process (Vindrola-Padros et al., 2017). However, embedded research is not without its challenges, as embedded research arrangements can be complex in nature (McGinity & Salokangas, 2014). For instance, politico-economic circumstances have implications for these arrangements, whereby embedded research is affected by external factors such as budgets cut, and organisational restructuring (i.e., employees, departments). Further, organisational-specific power-relationships, cultures, and interests might also affect embedded research, and the researcher might be torn between adhering to the desires of the host, even if it conflicts with scientific principles used in academia. For instance, internal interests might discourage scientific objectivity, whereby the researcher has to consider how negative or undesirable findings from research (should there be any) might directly or indirectly affect the organisation (Vindrola-Padros et al., 2017).

Embedded research programs between universities and professional soccer clubs have been established for several years (Lacome, 2023). In soccer, the embedded researcher is typically involved in collaborating with stakeholders in the soccer club to identify, design, and conduct research projects, and share such findings with the organisation, with a view of implementing change or new practice (Vindrola-Padros et al., 2017). As part of the arrangement, the embedded researcher might be required to provide day-to-day support in a specific role such as a sport scientist, or performance analyst. Often, these two responsibilities (academic research and daily support) converge with each other, whereby data collected in the training ground environment (e.g., through GPS) used to inform training, is also used to answer research questions. Such practice contrast with the traditional research model, whereby data is collected for the specific purpose of answering a research question. While there can be advantages to this practice, such as improving the ecological validity of the study’s findings, it is inherited with methodological and ethical issues that are unique to professional soccer. For instance, many professional soccer players are contractually obliged to take part in research projects (directly or indirectly) which raises questions about how the embedded researcher manages topics such as informed consent, and the handling of data. Such issues have been reported previously by an embedded researcher in professional soccer, who highlighted some of the moral and ethical questioning that coincides with it (Champ et al., 2020). This knowledge is important, as understanding the challenges that embedded researchers face in professional soccer, might allow for strategies to be put in place that strengthens the relationship between academia and host organisations. However, even as acknowledged by Champ et al., (2020), more research is required as the forementioned study is an account of one individual in one club and so might not be a representative view of other embedded researchers in similar positions.

To that end, in January 2019, an embedded PhD research program between Hull City Football Club, and the University of Hull commenced to explore the efficacy of VR in professional soccer. As part of the program, a practitioner-researcher role was undertaken whereby academic research was conducted at the soccer club, in combination with day-to-day sport science duties, and at times, an overlap was attained between the two. In this chapter of the thesis, a reflective account is presented on the interactions made with coaches, support staff, and players in relation to the personal, moral, and ethical challenges faced as an embedded researcher in professional soccer. In this chapter, three core challenges are discussed through a reflective account, those being: informed consent, confidentiality, and athlete data. That said, it is acknowledged that many other challenges could have been discussed in this chapter. Throughout this chapter, pseudonyms are used to maintain the anonymity of coaches, support staff, and players.

7.2 Informed consent

Informed consent is common practice in university ethics applications and can be defined as the attainment of written and/or verbal agreement from a participant to take part in a study following them receiving information on the aims of the study, their involvement, associated risks, and how the data will be used (del Carmen & Joffe, 2005). By obtaining such information, participants can make an informed decision on their willingness to take part in the study, and as such, it is a voluntary process. Yet, this topic in the context of embedded research in soccer is complicated by the fact that consent, even if provided, might not always be voluntary. In soccer, many players are contractually obliged to take part in research activity. At the time of a player agreeing to a contract, it might specify the involvement in research, in a general sense, but not with enough depth (as outlined early) for a player to truly be informed. Only when the project begins would a player be informed of what is required of them, but then the question is asked, if they agree to take part are they doing so because they are happy to do so, or because they are required to as part of their terms of employment. This situation is further complicated by the power-relationship seen in soccer, where a history of conformity exists between coach and the player (i.e., a level of expectation from players to participate in training as set out by the coaching staff). As such, in this reflection I detail an event that took place that highlights the challenges relating to embedded research and informed consent.

One of the benefits of embedded research in the soccer club was access to participants. The second study (Chapter 3) of the PhD began 18 months into the embedded research program, at the training ground. Recruitment of participants was a relatively simple process, given that I was working with them daily, and was made simpler having built a positive relationship with them. Ethical approval was granted, and informed consent was obtained verbally and by signature for each participant. On one occasion, I was sitting with Fred (soccer coach) during lunch who asked me about my study and what I hoped to get out of it. We talked, and it was nice for one of the coaches to take an interest in my research. As we finished the conversation, his last words on the topic were, ‘make sure you tell me if any of the lads (players) give you a hard time and I’ll make sure they do it’. In other words, I perceived this to be, that if they said no to the study participation, then Fred would reinforce that they had to do the research. I smiled and thanked him for his support. In the moment it was reassuring to have a coach looking out for me. However, shortly after I realised the difficult situation I had been put in. What would happen if a player decided not to take part? Would Fred have followed through with his comment and speak to the player? Could there have been consequences for the player? While it was unlikely that the topic of the player’s contract would have come into question, it was likely that the player would have been monetarily fined (as is a commonality in soccer for not obeying the rules) for not taking part. If Fred made the player take part, even if consent was provided verbally or through writing, could this still be classed as voluntary? Of course, had a player declined to take part in the study I could have just not told Fred. But there lies a personal constraint that I would have faced. Although contractually I did not report to Fred, what would the consequences have been had I withheld this information? A hegemonic culture exists within professional soccer, with the expectation of conforming to authority (Gearing, 1999). Fred was certainly in a more senior position than I was, so it would have felt inappropriate to not keep him informed on my progress given that he explicitly asked me to. Fortunately, I was not put in a position where a player decided they did not wish to take part, meaning I never had to approach Fred. Yet, this line of question highlights some of the moral and ethical issues on embedded research in professional soccer.

Perhaps, part of the issue here in this event is the lack of understanding from the coach regarding my role as an academic. Champ et al., (2020) discussed in their study the reluctance of the soccer club to acknowledge the authors identify as a researcher, with perceptions of lack of interest and ignorance towards their research. Whilst I do not feel the coach displayed any ignorance towards my research, I believe more education about informed consent might have been beneficial. For instance, before the study took place it could have been helpful to inform the staff at the soccer club, not only about the scope of the study but also how the process of recruitment and participation in academia works, and how it differs to that of the day-to-day general activity in the soccer club.

7.3 Confidentiality

Another commonality seen in university ethics applications is maintaining confidentiality of participants. Confidentiality refers to the separation or modification of participants personal identifying information from the data collected during the study (Allen, 2017). For instance, changing names to identification codes or pseudonyms are common practice, and are used to protect the privacy of the participant. This is especially important to keep participants safe from harm, embarrassment, or repercussions from employers (Allen, 2017). However, like the section above on informed consent, the nature of the environment (the soccer club) presented unique challenges in relation to handling the confidentiality of participants' data. In soccer, daily conversations take place between coaches and support staff relating to player data, including physical, tactical, and medical (Nosek et al., 2021). A variety of decisions are made using this data, such as match selection through to recruitment decisions. This topic (confidentiality) has been discussed previously with professional soccer coaches such as in a reflective account by Champ et al. (2020), where one coach refers to confidentiality of players as 'not the way it is around here …. everyone needs to be kept in the loop about everything'. However, as an embedded researcher a challenge exists whereby one part of the role involves the adherence to ethics guidelines as set by the university, and the other part involves working in an environment whereby player data are shared openly amongst staff. In this reflective account, I share my own experiences of handling the confidentiality of participant data following a study conducted at the soccer club.

As in most human participant research, information of the study was provided verbally and in writing, detailing the aims, requirements and importantly how confidentiality is adhered to. On the latter, I would explicitly make clear that the study data would be used for my research only, and not shared with staff at the soccer club. Most players did not express any concern or even ask any questions, they were just happy to proceed with participation based on the information provided. However, one incident stood out from the rest. In the morning before training commenced, I had recruited Simon, a young player in the soccer club to take part in my study. All information was provided as routine. Halfway through the study, Simon stopped and asked, "will the coaches see how I performed?". I reiterated the information I previously provided, yet his further comments suggested he was not at ease. He replied, "I just don't think I did well, and I don't want them (the coaches) to know". I reassured him of his concern and proceeded to finish the study data collection. Following this incident, I was never questioned by staff at the club on this (or any other) players regarding how they performed. Yet, what would have happened if a staff member did question me. It's easy to acclaim that the research has had to undergo ethics approval which requires me to maintain player confidentiality, and therefore I can not disclose the information. But at the same time, my role was to optimise the technology's (virtual reality) use in the soccer club, an environment that acknowledges the importance of sharing player information. Furthermore, the club had invested a significant monetary investment in this technology. Could I have told them that I would not be able to disclose the data that they are effectively paying for? On this subject, there was always a worry in my mind that the club could stop funding the resources required to complete the PhD.

On reflection, it would have been useful to clearly distinguish between the two roles I held at the club, the academic and the practitioner. As the academic, some research needs to maintain ethical standards to ensure data is captured with full confidence of the participant. Unfortunately, the issue remains when data collected is used as part of normal day-to-day soccer related activity as well as for research.

7.4 Athlete data

Recently, a discussion paper by the Australian Academy of Science (2022) was published, that aimed to create conversations regarding the complexity and scale of athlete data that is collected in team sports, while acknowledging the legal, ethical and data practices that are notably lacking. As an embedded researcher in a professional soccer environment, it was clear to see the vast amounts of data that was collected from players daily. Data collected from GPS, athlete management systems, physiological measurements, and game statistics are all used with best intentions of supporting the individual athlete and team during competition. However, quite often, vast amounts of data collected is limited, that provides tenuous benefits to the athlete, or the team (Australian Academy of Science, 2022). This is problematic, as information (through data) should only be collected if it is ‘reasonably necessary’, which is not always the case in professional soccer. Further, questions exist on athlete data ownership. Does it belong to the athlete or the organisation? Although there are bound to be legislation in place to govern data ownership, this reflective account examines this topic from the perspective of a novice embedded researcher, with sparce knowledge on the subject. Therefore, the topic of athlete data, it’s collection and ownership are discussed from the perspective of an individual applying academic research practices in a soccer environment.

On starting the embedded research program, I never asked questions regarding data collection, particularly in relation to ‘what’ is collected and ‘how’ it would be used. As an example, GPS data obtained from training and competition would produce over 50 variables, albeit only a small proportion of these variables would be used in reports for coaches and support staff to review. All this data would then be compiled into a database and stored for future use, if used at all. Although this practice is not as debatable as biometric data surveillance which constitutes to around-the-clock monitoring of athletes (Karkazis & Fishman, 2017), the practice of collecting data unnecessarily raises concerns. As a researcher, such practice would be deemed unethical, and for my own studies I had to make it very clear, what data was being collected and why. However, part of the problem is due to the technologies themselves, which produce large quantities of data irrespective of whether they are required or not. For example, in two of my studies, data was collected on variables that did not contribute towards the studies design and objectives, and simply had to be collected because there was not a way of opting out on specific variables. The concern of this was, could some of this additional data be used in a way that could consequentially impact the player? Although the data was anonymised and safely stored as per the data management plan outlined in my ethics application, what would I have done if management had asked to review the data I had collected on a specific player? I discuss this topic (confidentiality) in a previous paragraph. However, in the examples above the concern was directed towards data intentionally collected for the purpose of research. Now, it is the subject of additional data not used in my research but potentially viewed and interpreted by others in the organisation. On reflection, being an embedded researcher in the football club perhaps led me to believe that it is acceptable to collect vast amounts of data even if it is not to be used. But on reflection, such practices are unwarranted and as a researcher in a future situation I would make sure only data needed for the study is stored.

Another issue that the embedded researcher in soccer might face is the discussion of data ownership. From an ethical standpoint, participants are entitled to request access to the data that is collected on them (Australian Academy of Science, 2022). In soccer, technology might provide useful insights into the physiological and cognitive capabilities of a player which might be used to provide the team with a competitive edge against the opposition (Windt et al., 2020). In this sense, a player requesting their information obtained from research is not problematic. However, what happens if a player that has undertaken research at the club moves to an opposing team, in the same league. As an academic, a participant's occupation should not come into question if they request access to their data. But in soccer, data could provide useful information and hold a competitive advantage. From my perspective as an embedded researcher, it is unclear if I would need to ask the clubs permission first, especially as they are the ones who have paid for the technology and perhaps hold some rights to that data. If I refused the participants access to data, even if only temporarily, I would be acting unethically and not abiding by good ethical practice as required by me as a doctoral researcher. Likewise, would sharing data with a competitor be grounds for dismissal at the club I worked for?

7.5 Conclusion

This reflective account highlights some of the ethical challenges that the embedded researcher in football might encounter. Through real examples, challenges relating to informed consent, confidentiality, and the management and ownership of participant data are discussed. It is clear, applying academic research within a soccer environment is complex, with a number of decisions that are required to best serve the organisation and the academic field (Champ et al., 2020). Embedded researchers should engage in conversations early in the research program, involving key stakeholders from the university and soccer club to address some of these challenges.

Reference list

Abdi, H., Chin, W. W., Vinzi, V. E., Russolillo, G., & Trinchera, L. (2013). New Perspectives in Partial Least Squares and Related Methods. In *Springer Proceedings in Mathematics and Statistics* (Vol. 56). https://doi.org/10.1007/978-1-4614-8283-3

Abt, G., Boreham, C., Davison, G., Jackson, R., Nevill, A., Wallace, E., & Williams, M. (2020). Power, precision, and sample size estimation in sport and exercise science research. *Journal of Sports Sciences*, *38*(17), 1933–1935. https://doi.org/10.1080/02640414.2020.1776002

Abt, G., Jobson, S., Morin, J., Passfield, L., Sampaio, J., Sunderland, C., & Twist, C. (2022). Raising the bar in sports performance research. *Journal of Sports Sciences*, *40*(2), 125–129. https://doi.org/10.1080/02640414.2021.2024334

Adnan, H., William J, R., & Anthony W, B. (2018). Enhancing the Physiology and Effectiveness of Whole-Body Cryotherapy Treatment for Sports Recovery by Establishing an Optimum Protocol: A Review of Recent Perspectives. *Journal of Physical Medicine*, *1*(1), 41–52. https://doi.org/10.36959/942/338

Aksum, K. M., Magnaguagno, L., Bjørndal, C. T., & Jordet, G. (2020). What Do Football Players Look at? An Eye-Tracking Analysis of the Visual Fixations of Players in 11 v 11 Elite Football Match Play. *Frontiers in Psychology*, *11*(562995). https://doi.org/10.3389/fpsyg.2020.562995

Aksum, K. M., Pokolm, M., Bjørndal, C. T., Rein, R., Memmert, D., & Jordet, G. (2021). Scanning activity in elite youth football players. *Journal of Sports Sciences*, *39*(21), 2401–2410. https://doi.org/10.1080/02640414.2021.1935115

Alaker, M., Wynn, G. R., & Arulampalam, T. (2016). Virtual reality training in laparoscopic surgery: A systematic review & meta-analysis. *International Journal of Surgery*, *29*, 85–94. https://doi.org/10.1016/j.ijsu.2016.03.034

Albrecht, T., Nikendei, C., & Praetorius, M. (2022). Face, Content, and Construct Validity of a Virtual Reality Otoscopy Simulator and Applicability to Medical Training. *Otolaryngology–Head and Neck Surgery*, *166*(4), 753–759. https://doi.org/10.1177/01945998211032897

Alexandre, D., da Silva, C. D., Hill-Haas, S., Wong, D. P., Natali, A. J., De Lima, J. R. P., Filho, M. G. B. B., Marins, J. J. C. B., Garcia, E. S., & Karim, C. (2012). Heart Rate Monitoring in Soccer. *Journal of Strength and Conditioning Research*, *26*(10), 2890–2906.

https://doi.org/10.1519/JSC.0b013e3182429ac7

Ali, A., Williams, C., Hulse, M., Strudwick, A., Reddin, J., Howarth, L., Eldred, J., Hirst, M., & McGregor, S. (2007). Reliability and validity of two tests of soccer skill. *Journal of Sports Sciences*, *25*(13), 1461–1470. https://doi.org/10.1080/02640410601150470

Allen, M. (2017). Confidentiality and Anonymity of Participants. *The SAGE Encyclopedia of Communication Research Methods*, 228–230. https://doi.org/10.4135/9781483381411.n86

Almulla, J., Takiddin, A., & Househ, M. (2020). The use of technology in tracking soccer players' health performance: A scoping review. *BMC Medical Informatics and Decision Making*, *20*(1), 1–10. https://doi.org/10.1186/s12911-020-01156-4

Anzanello, M. J., & Fogliatto, F. S. (2011). Learning curve models and applications: Literature review and research directions. *International Journal of Industrial Ergonomics*, *41*(5), 573–583. https://doi.org/10.1016/j.ergon.2011.05.001

Appelbaum, L. G., & Erickson, G. (2018). Sports vision training: A review of the state-of-the-art in digital training techniques. *International Review of Sport and Exercise Psychology*, *11*(1), 160–189. https://doi.org/10.1080/1750984X.2016.1266376

Aquino, R., Puggina, E. F., Alves, I. S., & Garganta, J. (2017). Skill-related performance in soccer: A systematic review. *Human Movement*, *18*(5), 3–24. https://doi.org/10.1515/humo-2017-0042

Australian Academy of Science. (2022). *Getting Ahead of the Game: Athlete Data in Professional Sport*.

Ballhaus, W., Chow, W., & Rivet, E. (2022). *Perspectives from the Global Entertainment & Media Outlook 2022–2026*. PwC Global. https://www.pwc.com/gx/en/industries/tmt/media/outlook/outlook-perspectives.html

Baños, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpiña, C., & Alcañiz, M. (2000). Presence and Reality Judgment in Virtual Environments: A Unitary Construct? *CyberPsychology & Behavior*, *3*(3), 327–335. https://doi.org/10.1089/10949310050078760

Barillas, S. R., Oliver, J. L., Lloyd, R. S., & Pedley, J. S. (2023). Kinetic Responses to External Cues Are Specific to Both the Type of Cue and Type of Exercise in Adolescent Athletes. *Journal* *of Strength and Conditioning Research*, *37*(3), 597–605. https://doi.org/10.1519/JSC.0000000000004307

Barris, S., & Button, C. (2008). A review of vision-based motion analysis in sport. *Sports Medicine*, *38*(12), 1025–1043. https://doi.org/10.2165/00007256-200838120-00006

Bartlett, J. D., Lawrence, J. E., Yan, M., Guevel, B., Stewart, M. E., Audenaert, E., & Khanduja, V. (2020). The learning curves of a validated virtual reality hip arthroscopy simulator. *Archives of Orthopaedic and Trauma Surgery*, *140*(6), 761–767. https://doi.org/10.1007/s00402-020-03352-3

Bauer, M. S., Damschroder, L., Hagedorn, H., Smith, J., & Kilbourne, A. M. (2015). An introduction to implementation science for the non-specialist. *BMC Psychology*, *3*(1), 1–12. https://doi.org/10.1186/S40359-015-0089-9

Bauer, M. S., & Kirchner, J. A. (2020). Implementation science: What is it and why should I care? In *Psychiatry Research* (Vol. 283). https://doi.org/10.1016/j.psychres.2019.04.025

Be Your Best. (2020). *Be Your Best*. https://www.beyourbest.com/#Our-Purpose

Bekris, E., Kounalakis, S., Ispirlidis, I., & Katis, A. (2020). Evaluation of ball passing and space detection skill in soccer: implementation of two new soccer tests. *Research in Sports Medicine*, *28*(4), 518–528. https://doi.org/10.1080/15438627.2020.1789133

Berkman, M. I. (2018). History of Virtual Reality. *Encyclopedia of Computer Graphics and Games*, 1–9. https://doi.org/10.1007/978-3-319-08234-9_169-1

Bharathan, R., Vali, S., Setchell, T., Miskry, T., Darzi, A., & Aggarwal, R. (2013). Psychomotor skills and cognitive load training on a virtual reality laparoscopic simulator for tubal surgery is effective. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, *169*(2), 347–352. https://doi.org/10.1016/j.ejogrb.2013.03.017

Bird, J. M. (2020). The use of virtual reality head-mounted displays within applied sport psychology. *Journal of Sport Psychology in Action*, *11*(2), 115–128. https://doi.org/10.1080/21520704.2018.1563573

Bishop, G., & Fuchs, H. (1992). Research directions in virtual environments. *ACM SIGGRAPH Computer Graphics*, *26*(3), 153–177. https://doi.org/10.1145/142413.142416

Bowen, L., Gross, A. S., Gimpel, M., Bruce-low, S., & Li, F. (2019). Spikes in acute : chronic workload ratio (ACWR) associated with a 5 – 7 times greater injury rate in English Premier League football players : a comprehensive 3-year study. *British Journal of Sports Medicine*, 1–9. https://doi.org/10.1136/bjsports-2018-099422

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, *40*(7), 36–43. https://doi.org/10.1109/MC.2007.257

Bradley, P. S., Lago-Peñas, C., Rey, E., & Gomez Diaz, A. (2013). The effect of high and low percentage ball possession on physical and technical profiles in English FA Premier League soccer matches. *Journal of Sports Sciences*, *31*(12), 1261–1270. https://doi.org/10.1080/02640414.2013.786185

Brault, S., Kulpa, R., Duliscouët, L., Marin, A., & Bideau, B. (2015). Virtual kicker vs. real goalkeeper in soccer: a way to explore goalkeeper’s performance. *Movement & Sport Sciences - Science & Motricité*, *88*(89), 79–88. https://doi.org/10.1051/sm/2015026

Brownlee, T. E., Murtagh, C. F., Naughton, R. J., Whitworth-Turner, C. M., O’Boyle, A., Morgans, R., Morton, J. P., Erskine, R. M., & Drust, B. (2018). Isometric maximal voluntary force evaluated using an isometric mid-thigh pull differentiates English Premier League youth soccer players from a maturity-matched control group. *Science and Medicine in Football*, *2*(3), 209–215. https://doi.org/10.1080/24733938.2018.1432886

Bryer, J., & Speerschneider, K. (2016). *Likert: Analysis and Visualization Likert Items* (R package version 1.3.5).

Buchheit, M. (2008). The 30-15 Intermittent Fitness Test: Accuracy for Individualizing Interval Training of Young Intermittent Sport Players. *Journal of Strength and Conditioning Research*, *22*(2), 365–374. https://doi.org/10.1519/JSC.0b013e3181635b2e

Buckley, C. E., Nugent, E., Ryan, D., & Neary, P. C. (2012). Virtual Reality – A New Era in Surgical Training. In *Virtual Reality in Psychological, Medical and Pedagogical Applications*. InTech. https://doi.org/10.5772/46415

Burg, D., & Ausubel, J. H. (2021). Moore’s Law revisited through Intel chip density. *PLOS ONE*, *16*(8), e0256245. https://doi.org/10.1371/journal.pone.0256245

Bürkner, P.-C. (2017). brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Cambridge Dictionary. (n.d.). *Technology*. Retrieved March 11, 2023, from https://dictionary.cambridge.org/dictionary/english/technology

Carling, C. (2011). Influence of opposition team formation on physical and skill-related performance in a professional soccer team. *European Journal of Sport Science*, *11*(3), 155–164. https://doi.org/10.1080/17461391.2010.499972

Carroll, L. S. L. (2017). A comprehensive definition of technology from an ethological perspective. *Social Sciences*, *6*(4). https://doi.org/10.3390/socsci6040126

Caulfield, T., & Condit, C. (2012). Science and the sources of hype. *Public Health Genomics*, *15*(3–4), 209–217. https://doi.org/10.1159/000336533

Champ, F., Ronkainen, N., Nesti, M. S., Tod, D., & Littlewood, M. (2020). 'Through the lens of ethnography': Perceptions, challenges, and experiences of an early career practitioner-researcher in professional football. *Qualitative Research in Sport, Exercise and Health*, *12*(4), 513–529. https://doi.org/10.1080/2159676X.2019.1638444

Chang, E., Kim, H. T., & Yoo, B. (2020). Virtual Reality Sickness: A Review of Causes and Measurements. *International Journal of Human-Computer Interaction*, *36*(17), 1–25. https://doi.org/10.1080/10447318.2020.1778351

Chin, W. W., & Newsted, P. R. (1999). Structural Equation Modeling Analysis with Small Samples using Partial Least Squares. In R. H. Hoyle (Ed.), *Statistical Strategies for Small Sample Research* (pp. 307–341). Thousand Oaks: Sage.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Collet, C. (2013). The possession game? A comparative analysis of ball retention and team success in European and international football, 2007-2010. *Journal of Sports Sciences*, *31*(2), 123–136. https://doi.org/10.1080/02640414.2012.727455

Colliver, J. A., Conlee, M. J., & Verhulst, S. J. (2012). From test validity to construct validity... and back? *Medical Education*, *46*(4), 366–371. https://doi.org/10.1111/j.1365-2923.2011.04194.x

Cotterill, S. T. (2018). Virtual Reality and Sport Psychology: Implications for Applied Practice.

Case Studies in Sport and Exercise Psychology, *2*, 21–22. https://doi.org/https://doi.org/10.1123/cssep.2018-0002

Craig, C. (2014). Understanding perception and action in sport : how can virtual reality technology help? *Sports Technology*. https://doi.org/10.1080/19346182.2013.855224

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM*, *35*(6), 64–72. https://doi.org/10.1145/129888.129892

D'Orazio, T., & Leo, M. (2010). A review of vision-based systems for soccer video analysis. *Pattern Recognition*, *43*(8), 2911–2926. https://doi.org/10.1016/j.patcog.2010.03.009

Datta, D. (2017). *blandr: a Bland-Altman Method Comparison package for R*. https://doi.org/10.5281/zenodo.824514

Davis, F. D. (1986). A technology acceptance model for empirically testing new end-user information systems: Theory and results. In *Doctoral dissertation, Sloan School of Management, Massachusetts Institute of Technology*. https://doi.org/oclc/56932490

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

de Rooij, I. J. M., van de Port, I. G. L., & Meijer, J.-W. G. (2016). Effect of Virtual Reality Training on Balance and Gait Ability in Patients With Stroke: Systematic Review and Meta-Analysis. *Physical Therapy*, *96*(12), 1905–1918. https://doi.org/10.2522/ptj.20160054

Dedehayir, O., & Steinert, M. (2016). The hype cycle model: A review and future directions. *Technological Forecasting and Social Change*, *108*, 28–41. https://doi.org/10.1016/j.techfore.2016.04.005

del Carmen, M. G., & Joffe, S. (2005). Informed Consent for Medical Treatment and Research: A Review. *The Oncologist*, *10*(8), 636–641. https://doi.org/10.1634/theoncologist.10-8-

636

Dellaserra, C. L., Gao, Y., & Ransdell, L. (2014). Use of Integrated Technology in Team Sports: A Review of Opportunities, Challenges, and Future Directions for Athletes. *Journal of Strength and Conditioning Research*, *28*(2), 556–573. https://doi.org/10.1519/JSC.0b013e3182a952fb

Dessing, J. C., & Craig, C. M. (2010). Bending it like Beckham: How to visually fool the goalkeeper. *PLoS ONE*, *5*(10). https://doi.org/10.1371/journal.pone.0013161

DiCesare, C. A., Kiefer, A. W., Bonnette, S., & Myer, G. D. (2020). High-Risk Lower-Extremity Biomechanics Evaluated in Simulated Soccer-Specific Virtual Environments. *Journal of Sport Rehabilitation*, *29*(3), 294–300. https://doi.org/10.1123/jsr.2018-0237

Discombe, R. M., Bird, J. M., Kelly, A., Blake, R. L., Harris, D. J., & Vine, S. J. (2022). Effects of traditional and immersive video on anticipation in cricket: A temporal occlusion study. *Psychology of Sport and Exercise*, *58*, 102088. https://doi.org/10.1016/j.psychsport.2021.102088

Donaldson, A., Leggett, S., & Finch, C. F. (2012). Sports policy development and implementation in context: Researching and understanding the perceptions of community end-users. *International Review for the Sociology of Sport*, *47*(6), 743–760. https://doi.org/10.1177/1012690211422009

Dubrowski, A. (2005). Performance vs. learning curves: what is motor learning and how is it measured? *Surgical Endoscopy*, *19*(9), 1290–1290. https://doi.org/10.1007/s00464-004-8261-y

Düking, P., Holmberg, H.-C., & Sperlich, B. (2018). The Potential Usefulness of Virtual Reality Systems for Athletes: A Short SWOT Analysis. *Frontiers in Physiology*, *9*(128), 1–4. https://doi.org/10.3389/fphys.2018.00128

Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, *1*(1), 1–3. https://doi.org/10.1186/1748-5908-1-1

Eversbusch, A., & Grantcharov, T. P. (2004). Learning curves and impact of psychomotor training on performance in simulated colonoscopy: A randomized trial using a virtual reality endoscopy trainer. *Surgical Endoscopy and Other Interventional Techniques*, *18*(10), 1514–1518. https://doi.org/10.1007/s00464-003-9264-9

Fadde, P. J., & Zaichkowsky, L. (2018). Training perceptual-cognitive skills in sports using technology. *Journal of Sport Psychology in Action*, *9*(4), 239–248. https://doi.org/10.1080/21520704.2018.1509162

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feldman, L. S., Cao, J., Andalib, A., Fraser, S., & Fried, G. M. (2009). A method to characterize the learning curve for performance of a fundamental laparoscopic simulator task: Defining "learning plateau" and "learning rate." *Surgery*, *146*(2), 381–386. https://doi.org/10.1016/j.surg.2009.02.021

Fenn, J., & Raskino, M. (2008). *Mastering the Hype Cycle: How to Choose the Right Innovation at the Right Time*. Harvard Business Press.

Ferrer, R., Shishido, H., Kitahara, I., & Id, Y. K. (2020). Read-the-game : System for skill-based visual exploratory activity assessment with a full body virtual reality soccer simulation. *PLoS ONE*, *15*(3), 1–29. https://doi.org/10.1371/journal.pone.0230042

Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research.* Addison-Wesley. https://doi.org/10.2307/2065853

Foskett, A., Ali, A., & Gant, N. (2009). Caffeine enhances cognitive function and skill performance during simulated soccer activity. *International Journal of Sport Nutrition and Exercise Metabolism*, *19*(4), 410–423. https://doi.org/10.1123/ijsnem.19.4.410

Gallagher, A. G., & Satava, R. M. (2002). Virtual reality as a metric for the assessment of laparoscopic psychomotor skills: Learning curves and reliability measures. *Surgical Endoscopy and Other Interventional Techniques*, *16*(12), 1746–1752. https://doi.org/10.1007/s00464-001-8215-6

Gamble, P., Chia, L., & Allen, S. (2020). The illogic of being data-driven: reasserting control and restoring balance in our relationship with data and technology in football. *Science and Medicine in Football*, *4*(4), 338–341. https://doi.org/10.1080/24733938.2020.1854842

Gani, A., Pickering, O., Ellis, C., Sabri, O., & Pucher, P. (2022). Impact of haptic feedback on surgical training outcomes: A Randomised Controlled Trial of haptic versus non-haptic immersive virtual reality training. *Annals of Medicine & Surgery*, *83*, 104734. https://doi.org/10.1016/j.amsu.2022.104734

Gavgani, A. M., Walker, F. R., Hodgson, D. M., & Nalivaiko, E. (2018). A comparative study of cybersickness during exposure to virtual reality and “classic” motion sickness: Are they different? *Journal of Applied Physiology*, *125*(6), 1670–1680. https://doi.org/10.1152/japplphysiol.00338.2018

Gearing, B. (1999). Narratives of identity among former professional footballers in the United Kingdom. *Journal of Aging Studies*, *13*(1), 43–58. https://doi.org/10.1016/S0890-4065(99)80005-X

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems*, *4*(7). https://doi.org/10.17705/1cais.00407

Gianaros, P. J., Muth, E. R., Mordkoff, J. T., Levine, M. E., & Stern, R. M. (2001). A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation Space and Environmental Medicine*, *72*(2), 115–119.

Gibbons, T. (2016). *English National Identity and Football Fan Culture*. Routledge. https://doi.org/10.4324/9781315579788

Gigante, M. A. (1993). Virtual Reality: Definitions, History and Applications. In *Virtual Reality Systems* (pp. 3–14). Elsevier. https://doi.org/10.1016/B978-0-12-227748-1.50009-3

Gokeler, A., Benjaminse, A., Welling, W., Alferink, M., Eppinga, P., & Otten, B. (2015). The effects of attentional focus on jump performance and knee joint kinematics in patients after ACL reconstruction. *Physical Therapy in Sport*, *16*(2), 114–120. https://doi.org/10.1016/j.ptsp.2014.06.002

Gokeler, A., Bisschop, M., Myer, G. D., Benjaminse, A., Dijkstra, P. U., van Keeken, H. G., van Raay, J. J. A. M., Burgerhof, J. G. M., & Otten, E. (2016). Immersive virtual reality improves movement patterns in patients after ACL reconstruction: implications for enhanced criteria-based return-to-sport rehabilitation. *Knee Surgery, Sports Traumatology, Arthroscopy*, *24*(7), 2280–2286. https://doi.org/10.1007/s00167-014-3374-x

Goncalves, G., Monteiro, P., Melo, M., Vasconcelos-Raposo, J., & Bessa, M. (2020). A Comparative Study between Wired and Wireless Virtual Reality Setups. *IEEE Access*, *8*, 29249–29258. https://doi.org/10.1109/ACCESS.2020.2970921

Goto, H., Morris, J. G., & Nevill, M. E. (2015). Match Analysis of U9 and U10 English Premier League Academy Soccer Players Using a Global Positioning System. *Journal of Strength* *and Conditioning Research*, *29*(4), 954–963. https://doi.org/10.1519/JSC.0b013e3182a0d751

Grassini, S., & Laumann, K. (2020). Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00349

Grassini, S., Laumann, K., & Rasmussen Skogstad, M. (2020). The Use of Virtual Reality Alone Does Not Promote Training Performance (but Sense of Presence Does). *Frontiers in Psychology*, *11*(1743). https://doi.org/10.3389/fpsyg.2020.01743

Gray, R. (2017). Transfer of training from virtual to real baseball batting. *Frontiers in Psychology*, *8*(2183), 1–11. https://doi.org/10.3389/fpsyg.2017.02183

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, *31*(4), 337–350. https://doi.org/10.1007/s10654-016-0149-3

Grgic, J., Lazinica, B., & Pedisic, Z. (2021). Test–retest reliability of the 30–15 Intermittent Fitness Test: A systematic review. *Journal of Sport and Health Science*, *10*(4), 413–418. https://doi.org/10.1016/j.jshs.2020.04.010

Gurusamy, K., Aggarwal, R., Palanivelu, L., & Davidson, B. R. (2008). Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. *British Journal of Surgery*, *95*(9), 1088–1097. https://doi.org/10.1002/bjs.6344

Haar, S., Sundar, G., & Faisal, A. A. (2021). Embodied virtual reality for the study of real-world motor learning. *PLOS ONE*, *16*(1), e0245717. https://doi.org/10.1371/journal.pone.0245717

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, J., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2013). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). In *Sage*.

Hair, J., Risher, J., Sarstedt, M., & Ringle, C. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Handley, M. A., Gorukanti, A., & Cattamanchi, A. (2016). Strategies for implementing implementation science: a methodological overview. *Emergency Medicine Journal*, *33*(9), 660–664. https://doi.org/10.1136/emermed-2015-205461

Harris, D., Bird, J., Smart, P., Wilson, M., & Vine, S. (2020). A Framework for the Testing and Validation of Simulated Environments in Experimentation and Training. *Frontiers in Psychology*, *11*(March), 1–10. https://doi.org/10.3389/fpsyg.2020.00605

Harris, D., Buckingham, G., Wilson, M., Brookes, J., Mushtaq, F., Mon-Williams, M., & Vine, S. (2021). Exploring sensorimotor performance and user experience within a virtual reality golf putting simulator. *Virtual Reality*, *25*(3), 647–654. https://doi.org/10.1007/s10055-020-00480-4

Harris, D., Wilson, M., & Vine, S. (2018). A systematic review of commercial cognitive training devices: Implications for use in sport. *Frontiers in Psychology*, *9*(709). https://doi.org/10.3389/fpsyg.2018.00709

Harris, D., Wilson, M., & Vine, S. (2020). Development and validation of a simulation workload measure: the simulation task load index (SIM-TLX). *Virtual Reality*, *24*(4), 557–566. https://doi.org/10.1007/s10055-019-00422-9

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Power Technology and Engineering* (Vol. 43, Issue 5, pp. 139–183). https://doi.org/10.1016/S0166-4115(08)62386-9

Held, R. M., & Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators and Virtual Environments*, *1*(1), 109–112. https://doi.org/10.1162/pres.1992.1.1.109

Hoffmann, C. P., Filippeschi, A., Ruffaldi, E., & Bardy, B. G. (2014). Energy management using virtual reality improves 2000-m rowing performance. *Journal of Sports Sciences*, *32*(6), 501–509. https://doi.org/10.1080/02640414.2013.835435

Hogle, N. J., Briggs, W. M., & Fowler, D. L. (2007). Documenting a Learning Curve and Test-Retest Reliability of Two Tasks on a Virtual Reality Training Simulator in Laparoscopic Surgery. *Journal of Surgical Education*, *64*(6), 424–430. https://doi.org/10.1016/j.jsurg.2007.08.007

Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Medicine*, *30*(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001

Im, G. H., Shin, D., & Cheng, L. (2019). Critical review of validation models and practices in language testing: their limitations and future directions for validation research. *Language Testing in Asia*, *9*(1). https://doi.org/10.1186/s40468-019-0089-4

Impellizzeri, F. M., & Marcora, S. M. (2009). Test Validation in Sport Physiology: Lessons Learned From Clinimetrics. *International Journal of Sports Physiology and Performance*, *4*(2), 269–277. https://doi.org/10.1123/ijspp.4.2.269

Incisiv. (2020). *Incisiv*. https://cleansheet.incisiv.tech/?utm_source=incisiv&utm_campaign=main_website_referal

Ioannou, A., Papastavrou, E., Avraamides, M. N., & Charalambous, A. (2020). Virtual Reality and Symptoms Management of Anxiety, Depression, Fatigue, and Pain: A Systematic Review. *SAGE Open Nursing*, *6*, 1–13. https://doi.org/10.1177/2377960820936163

Jacobsen, M. E., Andersen, M. J., Hansen, C. O., & Konge, L. (2015). Testing basic competency in knee arthroscopy using a virtual reality simulator exploring validity and reliability. *Journal of Bone and Joint Surgery - American Volume*, *97*(9), 775–781. https://doi.org/10.2106/JBJS.N.00747

Janse, J. A., Goedegebuure, R. S. A., Veersema, S., Broekmans, F. J. M., & Schreuder, H. W. R. (2013). Hysteroscopic sterilization using a virtual reality simulator: Assessment of learning curve. *Journal of Minimally Invasive Gynecology*, *20*(6), 775–782. https://doi.org/10.1016/j.jmig.2013.04.016

JASP team. (2021). *JASP (Version 0.14.1)[Computer software]*.

Jokinen, E., Mikkola, T. S., & Härkki, P. (2020). Simulator training and residents’ first laparoscopic hysterectomy: a randomized controlled trial. *Surgical Endoscopy*, *34*(11), 4874–4882. https://doi.org/10.1007/s00464-019-07270-3

Jordet, G., Aksum, K. M., Pedersen, D. N., Walvekar, A., Trivedi, A., McCall, A., Ivarsson, A., & Priestley, D. (2020). Scanning, Contextual Factors, and Association With Performance in English Premier League Footballers: An Investigation Across a Season. *Frontiers in Psychology*, *11*(553813), 1–16. https://doi.org/10.3389/fpsyg.2020.553813

Kanumuri, P., Ganai, S., Wohaibi, E. M., Bush, R. W., Grow, D. R., & Seymour, N. E. (2008). Virtual reality and computer-enhanced training devices equally improve laparoscopic surgical skill in novices. *Journal of the Society of Laparoendoscopic Surgeons*, *12*(3), 219–226.

Karkazis, K., & Fishman, J. R. (2017). Tracking U.S. Professional Athletes: The Ethics of Biometric Technologies. *American Journal of Bioethics*, *17*(1), 45–60. https://doi.org/10.1080/15265161.2016.1251633

Kelly, N., Stafford, J., Craig, C., Herring, M. P., & Campbell, M. (2022). Using a virtual reality cricket simulator to explore the effects of pressure, competition anxiety on batting performance in cricket. *Psychology of Sport and Exercise*, *63*(102244). https://doi.org/10.1016/j.psychsport.2022.102244

Kennedy, R. S., Lane, N. E., Kevin, S., & Lilienthal, M. G. (1993). The International Journal of Aviation Psychology Simulator Sickness Questionnaire : An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. https://doi.org/10.1207/s15327108ijap0303

Keshavarz, B., & Hecht, H. (2011). Validating an efficient method to quantify motion sickness. *Human Factors*, *53*(4), 415–426. https://doi.org/10.1177/0018720811403736

Khanduja, V., Lawrence, J. E., & Audenaert, E. (2017). Testing the Construct Validity of a Virtual Reality Hip Arthroscopy Simulator. *Arthroscopy - Journal of Arthroscopic and Related Surgery*, *33*(3), 566–571. https://doi.org/10.1016/j.arthro.2016.09.028

Kim, D., & Ko, Y. J. (2019). The impact of virtual reality (VR) technology on sport spectators' flow experience and satisfaction. *Computers in Human Behavior*, *93*, 346–356. https://doi.org/10.1016/j.chb.2018.12.040

Kittel, A., Larkin, P., Elsworthy, N., Lindsay, R., & Spittle, M. (2020). Effectiveness of 360° virtual reality and match broadcast video to improve decision-making skill. *Science and Medicine in Football*, *4*(4), 255–262. https://doi.org/10.1080/24733938.2020.1754449

Kittel, A., Larkin, P., Elsworthy, N., & Spittle, M. (2019). Using 360° virtual reality as a decision-making assessment tool in sport. *Journal of Science and Medicine in Sport*, *22*(9), 1049–1053. https://doi.org/10.1016/j.jsams.2019.03.012

Koch, A., Pfandler, M., Stefan, P., Wucherer, P., Lazarovici, M., Navab, N., Stumpf, U., Schmidmaier, R., Glaser, J., & Weigl, M. (2019). Say, What Is on Your Mind? Surgeons'

Evaluations of Realism and Usability of a Virtual Reality Vertebroplasty Simulator. *Surgical Innovation*, *26*(2), 234–243. https://doi.org/10.1177/1553350618822869

Köckerling, F. (2014). Robotic vs. Standard Laparoscopic Technique - What is Better? *Frontiers in Surgery*, *1*(15), 10–13. https://doi.org/10.3389/fsurg.2014.00015

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Krosnick, J. A., & Presser, S. (2010). Question and Questionnaire Design. In *Handbook of Survey Research* (2nd ed., pp. 263–314). Emerald Group Publishing.

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin and Review*, *25*(1), 178–206. https://doi.org/10.3758/s13423-016-1221-4

Lacome, M. (2023). *Insights from an Embedded PhD*. https://www.mathlacome.com/blog

Lambourne, K., & Tomporowski, P. (2010). The effect of exercise-induced arousal on cognitive task performance: A meta-regression analysis. *Brain Research*, *1341*, 12–24. https://doi.org/10.1016/j.brainres.2010.03.091

Larkin, P., & O'Connor, D. (2017). Talent identification and recruitment in youth soccer: Recruiter's perceptions of the key attributes for player recruitment. *PLoS ONE*, *12*(4), 1–15. https://doi.org/10.1371/journal.pone.0175716

LaViola, J. J. (2000). A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, *32*(1), 47–56. https://doi.org/10.1145/333329.333344

Le Noury, P., Buszard, T., Reid, M., & Farrow, D. (2021). Examining the representativeness of a virtual reality environment for simulation of tennis performance. *Journal of Sports Sciences*, *39*(4), 412–420. https://doi.org/10.1080/02640414.2020.1823618

Le Noury, P., Polman, R., Maloney, M., & Gorman, A. (2022). A Narrative Review of the Current State of Extended Reality Technology and How it can be Utilised in Sport. *Sports Medicine*, *52*(7), 1473–1489. https://doi.org/10.1007/s40279-022-01669-0

Lee, M. D., & Wagenmakers, E. J. (2013). Bayesian cognitive modeling: A practical course. *Bayesian Cognitive Modeling: A Practical Course*, 1–264.

https://doi.org/10.1017/CBO9781139087759

Lees, A., Asai, T., Andersen, T. B., Nunome, H., & Sterzing, T. (2010). The biomechanics of kicking in soccer: A review. *Journal of Sports Sciences*, *28*(8), 805–817. https://doi.org/10.1080/02640414.2010.481305

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (R package version 1.6.1.).

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-sense of presence inventory. *Presence: Teleoperators and Virtual Environments*, *10*(3), 282–297. https://doi.org/10.1162/105474601300343612

Levac, D. E., & Sveistrup, H. (2014). Motor Learning and Virtual Reality. In *Virtual Reality for Physical and Motor Rehabilitation* (pp. 25–46). Springer. https://doi.org/10.1007/978-1-4939-0968-1_3

Li, H., Gail, M. H., Scott Braithwaite, R., Gold, H. T., Walter, D., Liu, M., Gross, C. P., & Makarov, D. V. (2014). Are hospitals “keeping up with the Joneses”?: Assessing the spatial and temporal diffusion of the surgical robot. *Healthcare*, *2*(2), 152–157. https://doi.org/10.1016/j.hjdsi.2013.10.002

Liebermann, D. G., Katz, L., Hughes, M. D., Bartlett, R. M., McClements, J., & Franks, I. M. (2002). Advances in the application of information technology to sport performance. *Journal of Sports Sciences*, *20*(10), 755–769. https://doi.org/10.1080/026404102320675611

Lim, K. H., Benbasat, I., & Ward, L. M. (2000). The Role of Multimedia in Changing First Impression Bias. *Information Systems Research*, *11*(2), 115–136. https://doi.org/10.1287/isre.11.2.115.11776

Liu, L., Miguel Cruz, A., Rios Rincon, A., Buttar, V., Ranson, Q., & Goertzen, D. (2015). What factors determine therapists’ acceptance of new technologies for rehabilitation-a study using the Unified Theory of Acceptance and Use of Technology (UTAUT). *Disability and Rehabilitation*, *37*(5), 447–455. https://doi.org/10.3109/09638288.2014.923529

Lohr, K. N. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research*, *11*(3), 193–205. https://doi.org/10.1023/A:1015291021312

Loturco, I., Pereira, L. A., Kobal, R., Cal Abad, C. C., Fernandes, V., Ramirez-Campillo, R., & Suchomel, T. (2018). Portable force plates: A viable and practical alternative to rapidly and accurately monitor elite sprint performance. *Sports*, *6*(3). https://doi.org/10.3390/sports6030061

Madden, J., Pandita, S., Schuldt, J. P., Kim, B., Won, A. S., & Holmes, N. G. (2020). Ready student one: Exploring the predictors of student learning in virtual reality. *PLoS ONE*, *15*(3), 1–27. https://doi.org/10.1371/journal.pone.0229788

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Malone, J. J., Harper, L. D., Jones, B., Perry, J., Barnes, C., Malone, J. J., Harper, L. D., Jones, B., Perry, J., Malone, J. J., Harper, L. D., Jones, B. E. N., Barnes, C., & Towlson, C. (2019). Perspectives of applied collaborative sport science research within professional team sports professional team sports. *European Journal of Sport Science*, *19*(2), 147–155. https://doi.org/10.1080/17461391.2018.1492632

Malone, J. J., Lovell, R., Varley, M. C., & Coutts, A. J. (2017). Unpacking the Black Box: Applications and Considerations for Using GPS Devices in Sport. *International Journal of Sports Physiology and Performance*, *12*, S2-18-S2-26. https://doi.org/10.1123/ijspp.2016-0236

Maneuvrier, A., Decker, L. M., Ceyte, H., Fleury, P., & Renaud, P. (2020). Presence Promotes Performance on a Virtual Spatial Cognition Task: Impact of Human Factors on Virtual Reality Assessment. *Frontiers in Virtual Reality*, *1*(571713), 1–17. https://doi.org/10.3389/frvir.2020.571713

Marescaux, J., Clément, J. M., Tassetti, V., Koehl, C., Cotin, S., Russier, Y., Mutter, D., Delingette, H., & Ayache, N. (1998). Virtual reality applied to hepatic surgery simulation: The next revolution. *Annals of Surgery*, *228*(5), 627–634. https://doi.org/10.1097/00000658-199811000-00001

Marris, J., Barrett, S., Abt, G., & Towlson, C. (2022). Quantifying technical actions in professional soccer using foot-mounted inertial measurement units. *Science and Medicine in Football*, *6*(2), 203–214. https://doi.org/10.1080/24733938.2021.1910333

Mascret, N., Montagne, G., Devrièse-Sence, A., Vu, A., & Kulpa, R. (2022). Acceptance by athletes of a virtual reality head-mounted display intended to enhance sport performance. *Psychology of Sport and Exercise*, *61*(102201). https://doi.org/10.1016/j.psychsport.2022.102201

McGinity, R., & Salokangas, M. (2014). Introduction: “embedded research” as an approach into academia for emerging researchers. *Management in Education*, *28*(1), 3–5. https://doi.org/10.1177/0892020613508863

McMenemy, K., & Ferguson, R. S. (2007). A Hitchhiker’s Guide to Virtual Reality. In *A Hitchhiker’s Guide to Virtual Reality*. A K Peters/CRC Press. https://doi.org/10.1201/b10677

Melim, A. M., Petiot, G. H., Clemente, F. M., Nakamura, F. Y., & Aquino, R. (2020). Home training recommendations for soccer players during the COVID - 19 pandemic. *SportRxiv [Preprint]*, Available from: https://osf.io/preprints/sportrxiv. https://doi.org/10.31236/osf.io/ycwvj

Mérienne, F. (2017). Virtual Reality : Principles and Applications. In *Encyclopedia of Computer Science and Technology* (pp. 1–11). Taylor and Francis. https://doi.org/10.1081/E-ECST2-140000194

Mesquida, C., Murphy, J., Lakens, D., & Warne, J. (2022). Replication concerns in sports science: a narrative review of selected methodological issues in the field. *SportRχiv*. https://doi.org/10.51224/SRXIV.127

Messick, S. (1990). VALIDITY OF TEST INTERPRETATION AND USE. *ETS Research Report Series*, *1*, 1487–1495. https://doi.org/10.1002/j.2333-8504.1990.tb01343.x

Mestre, D. R., Ewald, M., & Maiano, C. (2011). Virtual reality and exercise: Behavioral and psychological effects of visual feedback. *Annual Review of CyberTherapy and Telemedicine*, *9*(1), 99–103.

Michalski, S. C., Szpak, A., Saredakis, D., Ross, T. J., Billinghurst, M., & Loetscher, T. (2019). Getting your game on: Using virtual reality to improve real table tennis skills. *PLOS ONE*, *14*(9), 1–14. https://doi.org/10.1371/journal.pone.0222351

Miles, H. C., Pop, S. R., Watt, S. J., Lawrence, G. P., & John, N. W. (2012). A review of virtual environments for training in ball sports. *Computers and Graphics*, *36*(6), 714–726. https://doi.org/10.1016/j.cag.2012.04.007

Miller, H. L., & Bugnariu, N. L. (2016). Level of Immersion in Virtual Environments Impacts the Ability to Assess and Teach Social Skills in Autism Spectrum Disorder. *Cyberpsychology, Behavior, and Social Networking*, *19*(4), 246–256. https://doi.org/10.1089/cyber.2014.0682

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & De Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Møllerløkken, N. E., Lorås, H., & Pedersen, A. V. (2017). A Comparison of Players' and Coaches' Perceptions of the Coach-Created Motivational Climate within Youth Soccer Teams. *Frontiers in Psychology*, *8*(109), 1–10. https://doi.org/10.3389/fpsyg.2017.00109

Moore, G. E. (2006). Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff. *IEEE Solid-State Circuits Society Newsletter*, *11*(3), 33–35. https://doi.org/10.1109/N-SSC.2006.4785860

Muth, E. R., Stern, R. M., Thayer, J. F., & Koch, K. L. (1996). Assessment of the multiple dimensions of nausea: The nausea profile (NP). *Journal of Psychosomatic Research*, *40*(5), 511–520. https://doi.org/10.1016/0022-3999(95)00638-9

Myhill, N., Weaving, D., Robinson, M., Barrett, S., & Emmonds, S. (2023). Concurrent validity and between-unit reliability of a foot-mounted inertial measurement unit to measure velocity during team sport activity. *Science and Medicine in Football*, *00*(00), 1–9. https://doi.org/10.1080/24733938.2023.2237493

Nagahara, R., Botter, A., Rejc, E., Koido, M., Shimizu, T., Samozino, P., & Morin, J. B. (2017). Concurrent validity of GPS for deriving mechanical properties of sprint acceleration. *International Journal of Sports Physiology and Performance*, *12*(1), 129–132. https://doi.org/10.1123/ijspp.2015-0566

Nalivaiko, E., Davis, S. L., Blackmore, K. L., Vakulin, A., & Nesbitt, K. V. (2015). Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time. *Physiology & Behavior*, *151*, 583–590. https://doi.org/10.1016/j.physbeh.2015.08.043

Neretin, E. S., Kolokolnikov, P. A., & Mitrofanov, S. Y. (2021). Prospect for the application of augmented and virtual reality technologies in the design, production, operation of aircraft and training of aviation personnel. *Journal of Physics: Conference Series*, *1958*(1). https://doi.org/10.1088/1742-6596/1958/1/012030

Neumann, D. L., Moffitt, R. L., Thomas, P. R., Loveday, K., Watling, D. P., Lombard, C. L., Antonova, S., & Tremeer, M. A. (2018). A systematic review of the application of interactive virtual reality to sport. *Virtual Reality*, *22*(3), 183–198. https://doi.org/10.1007/s10055-017-0320-5

Nichols, S., Haldane, C., & Wilson, J. R. (2000). Measurement of presence and its consequences in virtual environments. *International Journal of Human Computer Studies*, *52*(3), 471–491. https://doi.org/10.1006/ijhc.1999.0343

Nitzl, C. (2016). The use of partial least squares structural equation modelling (PLS-SEM) in management accounting research: Directions for future theory development. *Journal of Accounting Literature*, *37*, 19–35. https://doi.org/10.1016/j.acclit.2016.09.003

Nortwick, S. S. Van, Lendvay, T. S., Jensen, A. R., Wright, A. S., Horvath, K. D., & Kim, S. (2009). Methodologies for establishing validity in surgical simulation studies. *Surgery*, *147*(5), 622–630. https://doi.org/10.1016/j.surg.2009.10.068

Nosek, P., Brownlee, T. E., Drust, B., & Andrew, M. (2021). Feedback of GPS training data within professional English soccer: a comparison of decision making and perceptions between coaches, players and performance staff. *Science and Medicine in Football*, *5*(1), 35–47. https://doi.org/10.1080/24733938.2020.1770320

Oagaz, H., Schoun, B., & Choi, M.-H. (2022). Performance Improvement and Skill Transfer in Table Tennis Through Training in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, *28*(12), 4332–4343. https://doi.org/10.1109/TVCG.2021.3086403

Padrón-Cabo, A., Rey, E., Pérez-Ferreirós, A., & Kalén, A. (2019). Test–Retest Reliability of Skill Tests in the F-MARC Battery for Youth Soccer Players. *Perceptual and Motor Skills*, *126*(5), 1006–1023. https://doi.org/10.1177/0031512519866038

Pagé, C., Bernier, P.-M., & Trempe, M. (2019). Using video simulations and virtual reality to improve decision-making skills in basketball. *Journal of Sports Sciences*, *37*(21), 2403–2410. https://doi.org/10.1080/02640414.2019.1638193

Palet, M. J., Antúnez-Riveros, M., & Barahona, M. (2021). Construct Validity of a Virtual Reality Simulator for Surgical Training in Knee Arthroscopy. *Cureus*, *13*(5).

https://doi.org/10.7759/cureus.15237

Palter, V. N., Orzech, N., Aggarwal, R., Okrainec, A., & Grantcharov, T. P. (2010). Resident perceptions of advanced laparoscopic skills training. *Surgical Endoscopy*, *24*(11), 2830–2834. https://doi.org/10.1007/s00464-010-1058-2

Parasuraman, A. (2000). Technology Readiness Index (Tri): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies. *Journal of Service Research*, *2*(4), 307–320. https://doi.org/10.1177/109467050024001

Parasuraman, A., & Colby, C. L. (2015). An Updated and Streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research*, *18*(1), 59–74. https://doi.org/10.1177/1094670514539730

Parton, B. J., & Neumann, D. L. (2019). The effects of competitiveness and challenge level on virtual reality rowing performance. *Psychology of Sport and Exercise*, *41*, 191–199. https://doi.org/10.1016/j.psychsport.2018.06.010

Patel, A. D., Gallagher, A. G., Nicholson, W. J., & Cates, C. U. (2006). Learning Curves and Reliability Measures for Virtual Reality Simulation in the Performance Assessment of Carotid Angiography. *Journal of the American College of Cardiology*, *47*(9), 1796–1802. https://doi.org/10.1016/j.jacc.2005.12.053

Patrick, D. L., & Beery, W. L. (1991). Measurement Issues: Reliability and Validity. *American Journal of Health Promotion*, *5*(4), 305–310. https://doi.org/10.4278/0890-1171-5.4.305

Persky, S., Kaphingst, K. A., McCall, C., Lachance, C., Beall, A. C., & Blascovich, J. (2009). Presence Relates to Distinct Outcomes in Two Virtual Environments Employing Different Learning Modalities. *CyberPsychology & Behavior*, *12*(3), 263–268. https://doi.org/10.1089/cpb.2008.0262

Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology*, *98*(1), 194–198. https://doi.org/10.1037/a0030767

Pichardo, A. W., Oliver, J. L., Harrison, C. B., Maulder, P. S., & Lloyd, R. S. (2018). Integrating models of long-term athletic development to maximize the physical development of youth. *International Journal of Sports Science and Coaching*, *13*(6), 1189–1199. https://doi.org/10.1177/1747954118785503

Plattfaut, R., & Koch, J. (2021). Preserving the legacy – Why do professional soccer clubs (not) adopt innovative process technologies? A grounded theory study. *Journal of Business Research*, *136*, 237–250. https://doi.org/10.1016/j.jbusres.2021.07.024

Pruna, R., & Bahdur, K. (2016). Cognition in Football. *Journal of Novel Physiotherapies*, *6*(6), 1–5. https://doi.org/10.4172/2165-7025.1000316

QSR International Pty Ltd. (2020). *NVivo* (released in March 2020).

Ramsay, C. R., Grant, A. M., Wallace, S. A., Garthwaite, P. H., Monk, A. F., & Russell, I. T. (2001). Statistical assessment of the learning curves of health technologies. *Health Technology Assessment*, *5*(12). https://doi.org/10.3310/hta5120

Redwood-Brown, A. (2008). Passing patterns before and after goal scoring in FA Premier League Soccer. *International Journal of Performance Analysis in Sport*, *8*(3), 172–182. https://doi.org/10.1080/24748668.2008.11868458

Rein, R., & Memmert, D. (2016). Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. *SpringerPlus*, *5*(1). https://doi.org/10.1186/s40064-016-3108-2

Reinartz, W., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, *26*(4), 332–344. https://doi.org/10.1016/j.ijresmar.2009.08.001

Renshaw, I., Davids, K., Araújo, D., Lucas, A., Roberts, W. M., Newcombe, D. J., & Franks, B. (2019). Evaluating weaknesses of “perceptual-cognitive training” and “brain training” methods in sport: An ecological dynamics critique. *Frontiers in Psychology*, *9*(2468). https://doi.org/10.3389/fpsyg.2018.02468

Rezzil. (2020). *Rezzil*. https://rezzil.com/learn-more/

Riach, J. (2013). *Virtual reality simulators could end England’s penality shootout woe*. The Guardian. https://www.theguardian.com/football/2013/aug/23/virtual-reality-england-penalty-shootout

Rigdon, E. E., Sarstedt, M., & Ringle, C. M. (2017). On Comparing Results from CB-SEM and PLS-SEM: Five Perspectives and Five Recommendations. *Marketing ZFP*, *39*(3), 4–16. https://doi.org/10.15358/0344-1369-2017-3-4

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). *"SmartPLS 3."* SmartPLS.

Ringuet-Riot, C. J., Hahn, A., & James, D. A. (2013). A structured approach for technology innovation in sport. *Sports Technology*, *6*(3), 137–149. https://doi.org/10.1080/19346182.2013.868468

Robertson, S., Gupta, S., Kremer, P., & Burnett, A. F. (2015). Development and measurement properties of a putting skill test for high-level golf. *European Journal of Sport Science*, *15*(2), 125–133. https://doi.org/10.1080/17461391.2014.932014

Robertson, S., Kremer, P., Aisbett, B., Tran, J., & Cerin, E. (2017). Consensus on measurement properties and feasibility of performance tests for the exercise and sport sciences: a Delphi study. *Sports Medicine - Open*, *3*(2), 1–10. https://doi.org/10.1186/s40798-016-0071-y

Rogers, E. M. (1962). *Diffusion of innovations* (First Ed). Free Press.

Rogers, E. M. (1995). *Diffusion of innovations* (5th ed.). Free Press.

Rosch, D., Hodgson, R., Peterson, L., Graf-Baumann, T., Junge, A., Chomiak, J., & Dvorak, J. (2000). Assessment and Evaluation of Football Performance. *The American Journal of Sports Medicine*, *28*(5), 29–39. https://doi.org/10.1177/28.suppl_5.s-29

Rose, T., Nam, C. S., & Chen, K. B. (2018). Effect of levels of immersion on performance and presence in virtual occupational tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 2079–2083. https://doi.org/10.1177/1541931218621469

RStudio Team. (2020). *RStudio: Integrated Development Environment for R* (1.3.959). RStudio, PBC.

Runswick, O. R., Rawlinson, A., Datson, N., & Allen, P. M. (2022). A valid and reliable test of technical skill for vision impaired football. *Science and Medicine in Football*, *6*(1), 89–97. https://doi.org/10.1080/24733938.2021.1885725

Russell, M., Benton, D., & Kingsley, M. (2011). The effects of fatigue on soccer skills performed during a soccer match simulation. *International Journal of Sports Physiology and Performance*, *6*(2), 221–233. https://doi.org/10.1123/ijspp.6.2.221

Russell, M., Rees, G., & Kingsley, M. I. C. (2013). Technical Demands of Soccer Match Play in the English Championship. *Journal of Strength and Conditioning Research*, *27*(10), 2869–

2873. https://doi.org/10.1519/JSC.0b013e318280cc13

Sal de Rellán-Guerra, A., Rey, E., Kalén, A., & Lago-Peñas, C. (2019). Age-related physical and technical match performance changes in elite soccer players. *Scandinavian Journal of Medicine and Science in Sports*, *29*(9), 1421–1427. https://doi.org/10.1111/sms.13463

Satava, R. M. (1993). Virtual reality surgical simulator. *Surgical Endoscopy*, *7*(3), 203–205. https://doi.org/10.1007/BF00594110

Schloerb, D. W. (1995). A questionnaire measure of telepresence. In *Presence* (Vol. 4, Issue 1, pp. 64–80).

Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322–339. https://doi.org/10.1037/met0000061

Schwartzman, D., Segal, R., & Drapeau, M. (2012). Perceptions of virtual reality among therapists who do not apply this technology in clinical practice. *Psychological Services*, *9*(3), 310–315. https://doi.org/10.1037/a0026801

Segal, R., Bhatia, M., & Drapeau, M. (2011). Therapists’ perception of benefits and costs of using virtual reality treatments. *Cyberpsychology, Behavior, and Social Networking*, *14*(1–2), 29–34. https://doi.org/10.1089/cyber.2009.0398

Seol, S. H., Ko, D. S., & Yeo, I. S. (2017). UX analysis based on TR and UTAUT of sports smart wearable devices. *KSII Transactions on Internet and Information Systems*, *11*(8), 4162–4179. https://doi.org/10.3837/tiis.2017.08.024

Shen, S., Chen, H.-T., Raffe, W., & Leong, T. W. (2021). Effects of Level of Immersion on Virtual Training Transfer of Bimanual Assembly Tasks. *Frontiers in Virtual Reality*, *2*(597487), 1–14. https://doi.org/10.3389/frvir.2021.597487

Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence: Teleoperators and Virtual Environments*, *1*(1), 120–126. https://doi.org/10.1162/pres.1992.1.1.120

Shimi, A., Tsestou, V., Hadjiaros, M., Neokleous, K., & Avraamides, M. (2021). Attentional skills in soccer: Evaluating the involvement of attention in executing a goalkeeping task in virtual reality. *Applied Sciences (Switzerland)*, *11*(19), 1–17. https://doi.org/10.3390/app11199341

Slater, M. (1999). Measuring Presence: A Response to the Witmer and Singer Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *8*(5), 560–565. https://doi.org/10.1162/105474699566477

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI*, *3*(74), 1–47. https://doi.org/10.3389/frobt.2016.00074

Slater, M., Spanlang, B., & Corominas, D. (2010). Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. *ACM SIGGRAPH 2010 Papers, SIGGRAPH 2010*, *29*(4), 1–9. https://doi.org/10.1145/1778765.1778829

Slater, M., Usoh, M., & Steed, A. (1995). Taking Steps: The Influence of a Walking Technique on Presence in Virtual Reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *2*(3), 201–219. https://doi.org/10.1145/210079.210084

Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Souza, V., Maciel, A., Nedel, L., & Kopper, R. (2021). Measuring Presence in Virtual Environments: A Survey. *ACM Computing Surveys*, *54*(8). https://doi.org/10.1145/3466817

Staunton, C. A., Abt, G., Weaving, D., & Wundersitz, D. W. T. (2022). Misuse of the term 'load' in sport and exercise science. *Journal of Science and Medicine in Sport*, *25*(5), 439–444. https://doi.org/10.1016/j.jsams.2021.08.013

Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, *42*(4), 73–93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x

Stevens, J. A., & Kincaid, J. P. (2015). The Relationship between Presence and Performance in Virtual Simulation Training. *Open Journal of Modelling and Simulation*, *03*(02), 41–48. https://doi.org/10.4236/ojmsi.2015.32005

Stølen, T., Chamari, K., Castagna, C., & Wisløff, U. (2005). Physiology of Soccer. *Sports Medicine*, *35*(6), 501–536. https://doi.org/10.2165/00007256-200535060-00004

Stone, J. A., Strafford, B. W., North, J. S., Toner, C., & Davids, K. (2018). Effectiveness and efficiency of virtual reality designs to enhance athlete development: an ecological dynamics perspective. *Movement & Sport Sciences - Science & Motricité*, *102*, 51–60. https://doi.org/10.1051/sm/2018031

Sutherland, L. M., Middleton, P. F., Anthony, A., Hamdorf, J., Cregan, P., Scott, D., & Maddern, G. J. (2006). Surgical simulation: A systematic review. *Annals of Surgery*, *243*(3), 291–300. https://doi.org/10.1097/01.sla.0000200839.93965.26

Svensson, M., & Drust, B. (2005). Testing soccer players. *Journal of Sports Sciences*, *23*(6), 601–618. https://doi.org/10.1080/02640410400021294

Sweet, J. (2016). *A model for improving survey outcomes by reducing cognitive load*. University of Illinois.

Taberner, M., Allen, T., & Cohen, D. D. (2019). Progressing rehabilitation after injury: consider the ‘control-chaos continuum.’ *British Journal of Sports Medicine*, *53*(18), 1132–1136. https://doi.org/10.1136/bjsports-2018-100157

Taberner, M., Van Dyk, N., Allen, T., Richter, C., Howarth, C., Scott, S., & Cohen, D. D. (2019). Physical preparation and return to sport of the football player with a tibia-fibula fracture: Applying the control-chaos continuum’. *BMJ Open Sport and Exercise Medicine*, *5*(1), 1–7. https://doi.org/10.1136/bmjsem-2019-000639

Thatcher, B., Ivanov, G., Szerovay, M., & Mills, G. (2021). Virtual Reality Technology in Football Coaching: Barriers and Opportunities. *International Sport Coaching Journal*, *8*(2), 234–243. https://doi.org/10.1123/iscj.2020-0011

Theodoropoulos, J. S., Bettle, J., & Kosy, J. D. (2020). The use of GPS and inertial devices for player monitoring in team sports: A review of current and future applications. *Orthopedic Reviews*, *12*(1), 1–7. https://doi.org/10.4081/or.2020.7863

Thompsett, B., Harland, A., & Roberts, J. (2016). Investigating the Relationship between Physical Properties of a Football and Player Perceptions. *Procedia Engineering*, *147*, 519–525. https://doi.org/10.1016/j.proeng.2016.06.231

Tirp, J., Steingröver, C., Wattie, N., Baker, J., & Schorer, J. (2015). Virtual realities as optimal learning environments in sport – A transfer study of virtual and real dart throwing. *Psychological Test and Assessment Modeling*, *57*(1), 57–69.

Towlson, C., Midgley, A. W., & Lovell, R. (2013). Warm-up strategies of professional soccer players: practitioners' perspectives. *Journal of Sports Sciences*, *31*(13), 1393–1401. https://doi.org/10.1080/02640414.2013.792946

Tsaousidis, N., & Zatsiorsky, V. (1996). Two types of ball-effector interaction and their relative contribution to soccer kicking. *Human Movement Science*, *15*(6), 861–876. https://doi.org/10.1016/S0167-9457(96)00027-9

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments*, *9*(5), 497–503. https://doi.org/10.1162/105474600566989

van Baren, J., & IJsselsteijn, W. (2004). Deliverable 5 Measuring Presence : A Guide to Current Measurement Approaches. In *Measurement*.

Vater, C., Gray, R., & Holcombe, A. O. (2021). A critical systematic review of the Neurotracker perceptual-cognitive training tool. *Psychonomic Bulletin and Review*, *28*(5), 1458–1483. https://doi.org/10.3758/s13423-021-01892-2

Venkatesh, Morris, Davis, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Vickers, J. (1996). Visual control when aiming at a far target. *Journal of Experimental Psychology. Human Perception and Performance*, *22*(2), 342–354.

Vindrola-Padros, C., Pape, T., Utley, M., & Fulop, N. J. (2017). The role of embedded research in quality improvement: A narrative review. *BMJ Quality and Safety*, *26*(1), 70–80. https://doi.org/10.1136/bmjqs-2015-004877

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin and Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and Cybersickness in Virtual Reality Are Negatively Related: A Review. *Frontiers in Psychology*, *10*(158), 1–19. https://doi.org/10.3389/fpsyg.2019.00158

Westen, D., & Rosenthal, R. (2003). Quantifying Construct Validity: Two Simple Measures.

Journal of Personality and Social Psychology, *84*(3), 608–618. https://doi.org/10.1037/0022-3514.84.3.608

Weston, M. (2018). Training load monitoring in elite English soccer : a comparison of practices and perceptions between coaches and practitioners perceptions between coaches and practitioners. *Science and Medicine in Football*, *2*(3), 216–224. https://doi.org/10.1080/24733938.2018.1427883

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag.

Williams, A. M., & Davids, K. (1998). Visual search strategy, selective attention, and expertise in soccer. *Research Quarterly for Exercise and Sport*, *69*(2), 111–128. https://doi.org/10.1080/02701367.1998.10607677

Wilson, R. S., James, R. S., David, G., Hermann, E., Morgan, O. J., Niehaus, A. C., Hunter, A., Thake, D., & Smith, M. D. (2016). Multivariate analyses of individual variation in soccer skill as a tool for talent identification and development: utilising evolutionary theory in sports science. *Journal of Sports Sciences*, *34*(21), 2074–2086. https://doi.org/10.1080/02640414.2016.1151544

Windt, J., MacDonald, K., Taylor, D., Zumbo, B. D., Sporer, B. C., & Martin, D. T. (2020). "'To tech or not to tech?'" A critical decision-making framework for implementing technology in sport. *Journal of Athletic Training*, *55*(9), 902–910. https://doi.org/10.4085/1062-6050-0540.19

Wirth, M., Kohl, S., Gradl, S., Farlock, R., Roth, D., & Eskofier, B. M. (2021). Assessing Visual Exploratory Activity of Athletes in Virtual Reality Using Head Motion Characteristics. *Sensors*, *21*(11), 3728. https://doi.org/10.3390/s21113728

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240. https://doi.org/10.1162/105474698565686

Wohlgenannt, I., Simons, A., & Stieglitz, S. (2020). Virtual Reality. *Business and Information Systems Engineering*, *62*(5), 455–461. https://doi.org/10.1007/s12599-020-00658-9

Wood, G., Wright, D. J., Harris, D., Pal, A., Franklin, Z. C., & Vine, S. J. (2021). Testing the construct validity of a soccer-specific virtual reality simulator using novice, academy, and professional soccer players. *Virtual Reality*, *25*(1), 43–51. https://doi.org/10.1007/s10055-020-00441-x

Yildirim, C. (2020). Don’t make me sick: investigating the incidence of cybersickness in commercial virtual reality headsets. *Virtual Reality*, *24*(2), 231–239. https://doi.org/10.1007/s10055-019-00401-0

‘You can’t score if you don’t shoot.’ (2016, March). ‘You can’t score if you don’t shoot’: Johan Cruyff, Ajax and Barca legend, dies at 68’. *The National News*.

Zaal, F. T. J. M., & Bootsma, R. J. (2011). Virtual reality as a tool for the study of perception-action: The case of running to catch fly balls. *Presence: Teleoperators and Virtual Environments*, *20*(1), 93–103. https://doi.org/10.1162/pres_a_00037

Appendix 1: HTC Vive trackers fitted to the trainers, chapter 3 and 4

Appendix 2: Room set up for Chapter 3. Blue circles indicate location of HTC lighthouses

II

Appendix 3: Poster used to illustrate the VR task in Chapter 3

RONDO SCAN

A Scanning Drill providing data about a player's anticipation and decision-making process while passing and receiving the ball.

This drill helps a player improve on the ball composure and situational awareness, whilst enhancing their anticipation and body form.

01 Scan left and right, Look for the ball and be ready to shoot into the highlighted net.

02 A green light will appear, a ball released then a net will be highlighted.

03 Look for the highlighted net and position your body to receive the ball.

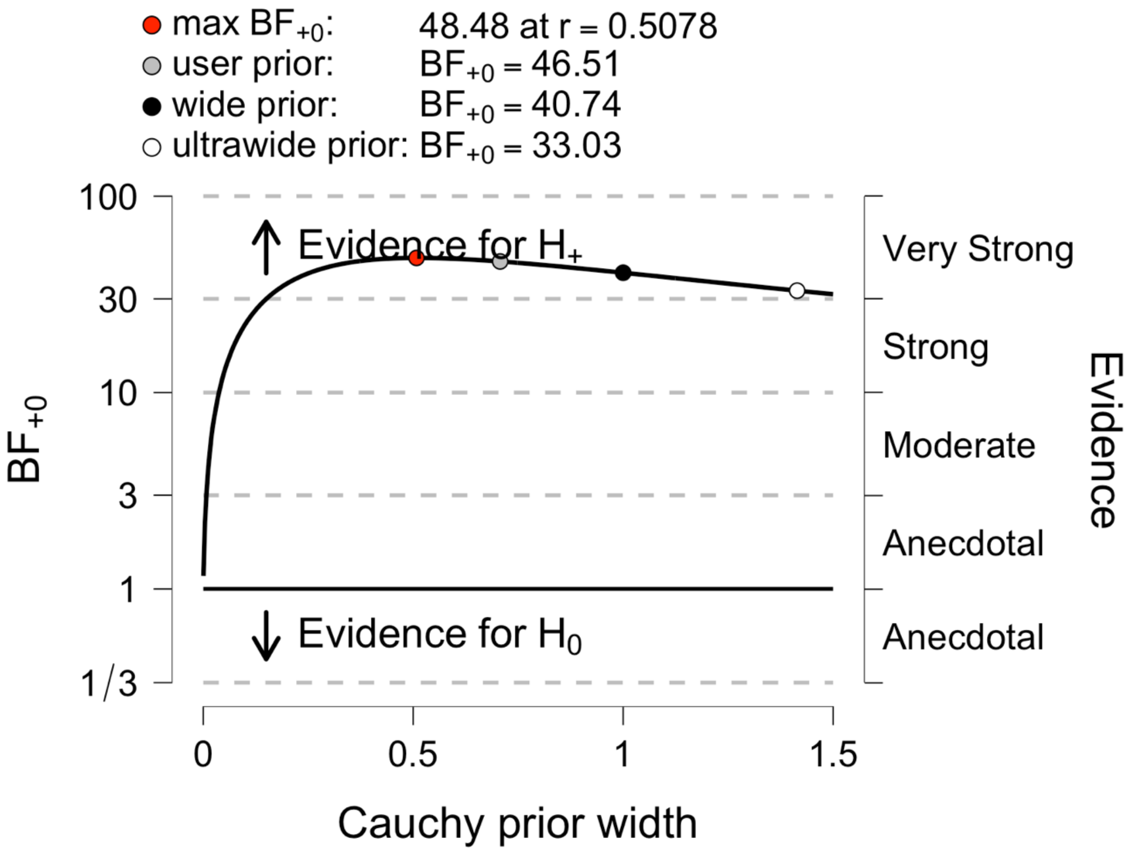
04 Kick the ball into the highlighted net, you can stop the ball first if needed.

05 Once you have kicked the ball keep scanning and look for the next light.

REZZIL®
POWERED BY MIHIEPA

T: +1 201 753-5403 E: info@rezzil.com www.rezzil.com

Appendix 4: Bayes Factors Robustness test for H_1, Chapter 3



Appendix 5: Recruitment poster used in Chapter 4

UNIVERSITY OF Hull

Interested in virtual reality and football?

Help us to investigate a football passing test in virtual reality. For more details scan the QR code

1 Between the age of 16-37?

2 Participate in football?

3 Healthy and injury free?

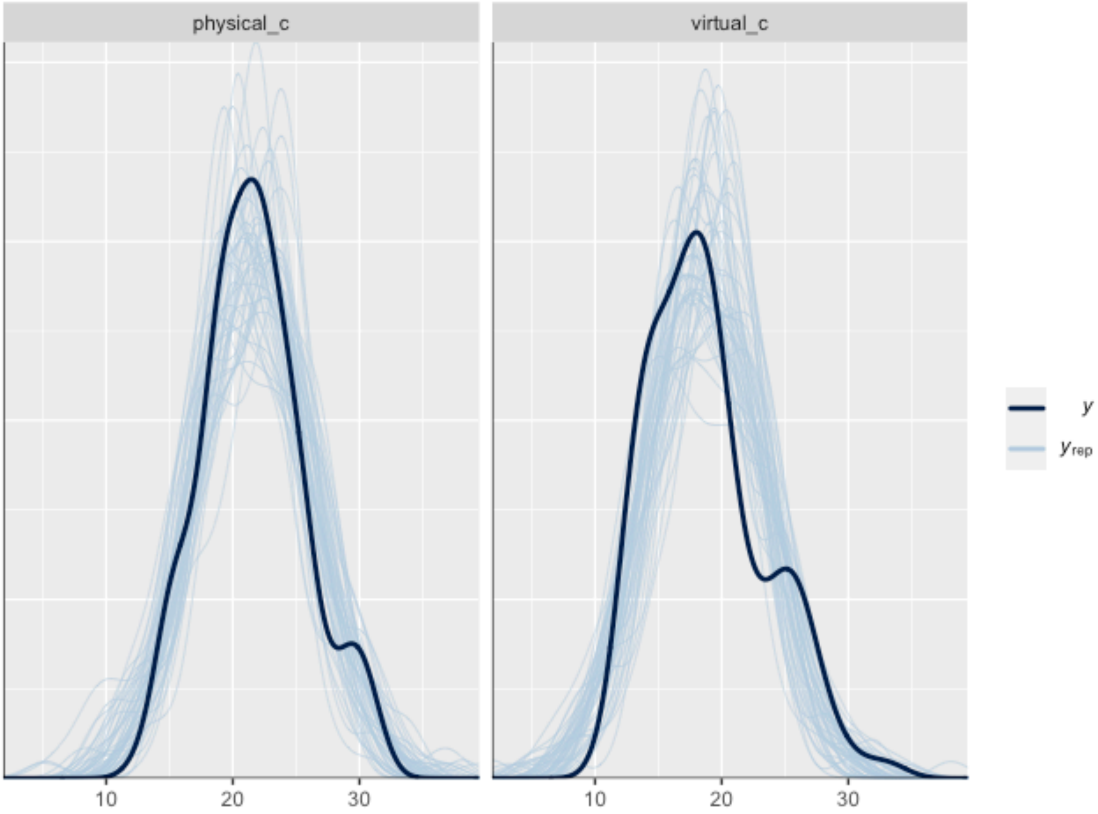
This study will require a single 45-minute visit to the University of Hull, Sport, Health and Exercise Science lab, HU6 7RX

If interested or want to know more, scan to QR code or contact Ben:

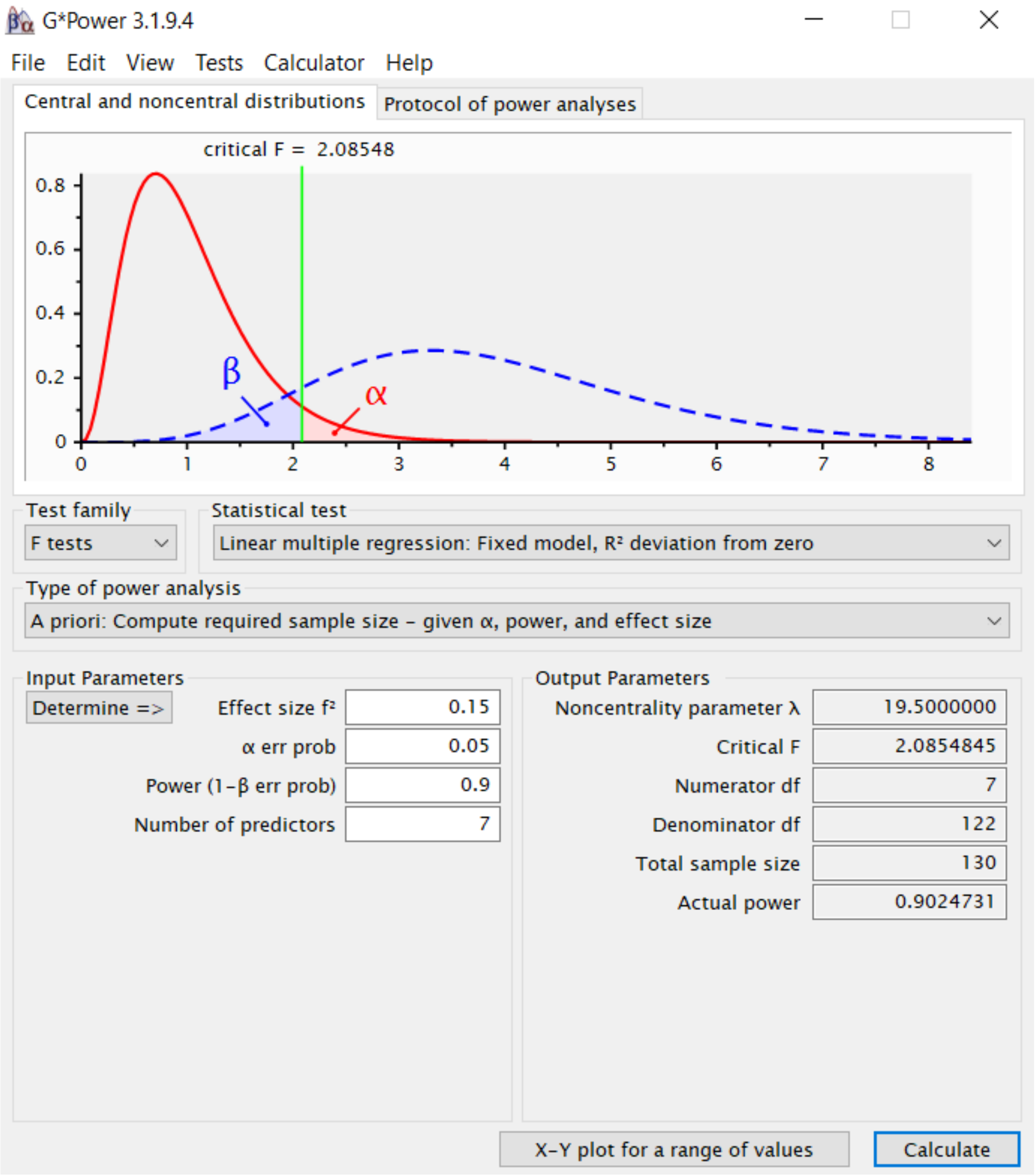
ben.greenhough@hulltigers.com

@btgreenhough

Appendix 6: Posterior predictive checks for confirmatory analysis, Chapter 4



Appendix 7: G power output for sample size calculation, Chapter 5



Test family: F tests

Statistical test: Linear multiple regression: Fixed model, R^2 deviation from zero

Type of power analysis: A priori: Compute required sample size – given α, power, and effect size

Input Parameters

Effect size f^2: 0.15
α err prob: 0.05
Power (1–β err prob): 0.9
Number of predictors: 7

Output Parameters

Noncentrality parameter λ: 19.5000000
Critical F: 2.0854845
Numerator df: 7
Denominator df: 122
Total sample size: 130
Actual power: 0.9024731